

Implementationserfolg von Schulreformen in der Berliner Schulanfangsphase

Dissertation

zur Erlangung des akademischen Grades einer Doktorin der
Philosophie (Dr. phil.)

am Fachbereich Erziehungswissenschaft und Psychologie
der Freien Universität Berlin

Dekan: Prof. Dr. Hauke Heekeren

vorgelegt von

Katharina Thoren, M. A.

Berlin, März 2017

Gutachter/in: 1. Prof. Dr. Martin Brunner

2. Prof. Dr. Hans Anand Pant

Datum der Einreichung: 31. März 2017

Datum der Disputation: 09. Juni 2017

Diese Arbeit entstand im Rahmen des von der Senatsverwaltung für Bildung, Jugend und Wissenschaft (SenBJW) in Berlin geförderten Projekts „ALTERJÜL: Wie hängen jahrgangsübergreifendes Lernen und Einschulungsalter mit der schulischen Entwicklung zusammen?“ unter der Leitung von Prof. Dr. Martin Brunner. Es handelt sich dabei um ein Gemeinschaftsprojekt in Kooperation mit Prof. Dr. Petra Stanat und Prof. Dr. Hans Anand Pant.

Danksagung

Die Danksagung ist in der Online-Version
aus Gründen des Datenschutzes nicht enthalten.

Inhaltsverzeichnis

Zusammenfassung	13
Summary	16
Einleitung	19
1. Implementationserfolg von Schulreformen	23
1.1 Begriffsklärung	23
1.2 Der Implementationsprozess	25
1.3 Implementationsstrategien.....	27
1.4 Implementationserfolg.....	30
1.5 Fazit	38
2. Schulbeginn in Berlin und Deutschland	41
2.1 Schulbeginn in Deutschland	41
2.2 Historische und aktuelle Entwicklungen der Schulanfangsphase in Deutschland	45
2.3 Die Schulanfangsphase in Berlin.....	48
2.4 Einführung des Jahrgangsübergreifenden Lernens (JÜL)	52
2.5 Senkung des Schuleintrittsalters	54
2.6 Zusammenfassung	55
3. Zielsetzung der Arbeit	59
3.1 Studien zur Einführung des Jahrgangsübergreifenden Lernens (JÜL).....	59
3.2 Studien zur Senkung des Schuleintrittsalters	60
4. Studien zur Einführung des Jahrgangsübergreifenden Lernens	63
4.1 Studie I: Wie wurde das Jahrgangsübergreifende Lernen an Berliner Grundschulen implementiert?	66
4.2 Studie II: Unterscheiden sich Schulen entsprechend ihrer Zuordnung zu einem Implementationstypen in ihrer Schul- und Unterrichtsqualität?	87
4.3 Fazit der Studien zum JÜL	98
5. Studien zur Senkung des Schuleintrittsalters	101
5.1 Studie III: Wie wurde die Senkung des Schuleintrittsalters an Berliner Grundschulen umgesetzt?.....	103
5.2 Studie IV: Wie wirkt sich die Senkung des Schuleintrittsalters auf die Schüler*innenleistung aus?	116
5.3 Studie V: Relative Age Effects in Mathematics and Reading: Investigating the Generalizability across Students, Time, and Classes	137
5.4 Fazit der Studien zum Schuleintrittsalter	159
6. Gesamtdiskussion	161

6.1 Zusammenfassende Darstellung und Diskussion der Ergebnisse	161
6.2 Diskussion vor dem Hintergrund implementationstheoretischer Erkenntnisse.....	168
6.3 Limitationen der Arbeit und Ausblick.....	171
6.4 Implikationen für die Bildungspolitik	173
6.5 Implikationen für die Forschung	174
7. Literaturverzeichnis.....	177
Anhang	197
A.1.Übersicht der Schulgesetze und -verordnungen in den Bundesländern	197
A.2.,,Reformen auf einen Blick“ (SenBJW)	200
A.3.Übersicht zur Berichterstattung regionaler und überregionaler Medien	206
A.4.Ergänzende Tabellen Studie II	213
A.5.Ergänzende Tabellen zu Studie IV	218
Erklärung.....	225

Abbildungsverzeichnis

Abbildung 1: Anwenderkategorien entsprechend ihrer Innovationsfreude (Rogers, 2003, S. 281).....	36
Abbildung 2: Ökosystemisches Modell der Schulfähigkeit (Nickel & Schmidt-Denter, 1995, S. 227)	43
Abbildung 3: Ausschnitt der Faxabfrage für Umsetzung von JÜL	73
Abbildung 4: Verteilung der Schulen nach der Anzahl ihrer JüL-Klassen	78
Abbildung 5: Verteilung der JüL-2 und JüL-3-Klassen innerhalb der Schulen, die JüL in mindestens einer ihrer Klassen der Schulanfangsphase umsetzen.....	79
Abbildung 6: Typen der JÜL-Implementation	81
Abbildung 7: Anteile der jahrgangsübergreifend unterrichteten Stunden an den Jahresunterrichtsstunden für das Fach Deutsch	83
Abbildung 8: Anteile der jahrgangsübergreifend unterrichteten Stunden an den Jahresunterrichtsstunden für das Fach Mathematik	83
Abbildung 9: Streudiagramm zu den jahrgangsübergreifend unterrichteten Stunden in den Fächern Deutsch und Mathematik in der zweiten und dritten Jahrgangsstufe; Korrelationskoeffizient r	84
Abbildung 10: Strukturmodell zu Evaluation und Qualitätssicherung im Bildungswesen (Ditton, 2000; 2002)	88
Abbildung 11: Ergebnisse der ANOVAs zu Unterschieden der Schul- und Unterrichtsqualität zwischen den Typen der JÜL-Implementation.....	93
Abbildung 12: Ergebnisse der Unterschiedsanalysen (t-Tests, η^2 , Cohens d) zu Unterschieden der Schul- und Unterrichtsqualität zwischen den JÜL-2- und JÜL-3-Schulen	96
Abbildung 13: Multikohorten-Sequenz-Design der Studien zur Senkung des Schuleintrittsalters.....	104
Abbildung 14: Ausschnitt der Klassenliste aus dem ISQ-Portal	108
Abbildung 15: Entwicklung der vorzeitigen und verzögerten Einschulungen in Deutschland und Berlin (2000-2014).....	114
Abbildung 16: Altersgruppenverteilungen der Hauptkohorte	124
Abbildung 17: Altersgruppenverteilungen der Vergleichskohorte I im Vergleich zur Hauptkohorte..	126
Abbildung 18: Altersgruppenverteilungen der Vergleichskohorte II im Vergleich zur Hauptkohorte	126
Abbildung 19: Altersgruppenverteilungen der Hauptkohorte, nach Schularten unterteilt	127
Abbildung 20: Mittlere Leistung (z-standardisiert) der Altersgruppen der Hauptkohorte in Deutsch / Lesen	128
Abbildung 21: Mittlere Leistung (z-standardisiert) der Altersgruppen der Hauptkohorte in Mathematik	129
Abbildung 22: Mittlere Leistung (z-standardisiert) der Altersgruppen der Vergleichskohorte I in Deutsch / Lesen.....	131
Abbildung 23: Mittlere Leistung (z-standardisiert) der Altersgruppen der Vergleichskohorte I in Mathematik	131

Abbildung 24: Mittlere Leistung (z-standardisiert) der Altersgruppen der Vergleichskohorte II in Deutsch / Lesen.....	132
Abbildung 25: Mittlere Leistung (z-standardisiert) der Altersgruppen der Vergleichskohorte II in Mathematik	132
Abbildung 26: Leistungsverteilung (Boxplots) der Hauptkohorte nach Einschulungsalter (in Monaten)	133
Abbildung 27: Gegenüberstellung der Anwendungskategorien mit den JÜL- Implementationstypen	169

Tabellenverzeichnis

Tabelle 1: Tätigkeiten im Implementationsprozess nach Goldenbaum (2012, S. 91)	26
Tabelle 2: Einflussfaktoren auf die Implementation schulischer Innovationen (Goldenbaum, 2012, S. 115, 2013, S. 50)	31
Tabelle 3: Erfolgsindikatoren von Implementationsprozessen nach Coburn (2003).....	34
Tabelle 4: Entwicklungsverlauf der JÜL-Implementation in Berlin	67
Tabelle 5: Übersicht zu den Daten der Schulstatistik für die Schulanfangsphase (Schulen in öffentlicher Trägerschaft).....	72
Tabelle 6: Übersicht zur Erweiterungsstudie.....	74
Tabelle 7: Anteil der JÜL-Klassen in Berliner Schulen und Klassen.....	77
Tabelle 8: Fit-Statistiken der latenten Profilanalyse zur JÜL-Implementation	80
Tabelle 9: t-Tests für abhängige Stichproben über die Jahrgangsstufen 2 und 3.....	83
Tabelle 10: Übersicht zur den jahrgangübergreifend unterrichteten Stunden.....	84
Tabelle 11: Deskriptive Übersicht: Qualitätsmerkmale der Schulinspektion.....	90
Tabelle 12: Gegenüberstellung der Daten der Schulstatistik und Schulinspektion	91
Tabelle 13: Verteilung der JÜL-2- und JÜL-3-Schulen über die Implementationstypen „Überzeugte“ und „Pioniere“ (absolute Zahlen).....	95
Tabelle 14: Übersicht zu den Orientierungsarbeiten – Vergleichskohorte I und Hauptkohorte ...	106
Tabelle 15: Übersicht zu den Vergleichsarbeiten der 3. Jahrgangsstufe (VERA 3) – Hauptkohorte und Vergleichskohorte II.....	107
Tabelle 16: Übersicht zu den Vergleichsarbeiten der 8. Jahrgangsstufe (VERA 8) - Vergleichskohorte I und Hauptkohorte.....	107
Tabelle 17: Übersicht zu den Erweiterungsstudien	108
Tabelle 18: Übersicht der Datengrundlage	109
Tabelle 19: erwartete und tatsächliche Verteilung der Geburten über die Monate eines Jahres, sowie der jeweilige χ^2 -Wert	111
Tabelle 20: erwartete und tatsächliche Verteilung der Schulanfänger*innen entsprechend ihres Geburtsmonats, sowie der jeweilige χ^2 -Wert.....	112
Tabelle 21: Mittleres Schuleintrittsalter nach Jahrgangsstufe und Kohorte (Mittelwert und Standardabweichung).....	113
Tabelle 22: Intraklassenkorrelationen (ICC) für das Schuleintrittsalter der drei Kohorten.....	115
Tabelle 23: Zuordnung der Schüler*innen zu den Altersgruppen entsprechend ihrem Einschulungsalter (in Monaten).....	121
Tabelle 24: Anteil der durch das Alter eines Kindes aufgeklärten Varianz an der Varianz schulischer Leistungen.....	134

Zusammenfassung

Reformen stellen auch im Bildungswesen einen wesentlichen Motor zur Sicherung von Qualität und Weiterentwicklung dar. Berlin führte im Jahr 2004 u. a. als Antwort auf die Ergebnisse internationaler Schulleistungsstudien ein neues Schulgesetz ein, welches umfassende Überarbeitungen in allen Bereichen des Schulsystems beinhaltete. Ein wesentliches Anliegen im Bereich der Grundschule war hier die Überarbeitung des Schulanfangs. Zwei Reformen, die in Zusammenhang mit der neuen sogenannten Schuleingangsphase (SAPH) besondere öffentliche und politische Aufmerksamkeit bekamen, waren die verpflichtende Einführung des Jahrgangsübergreifenden Lernens (JÜL) zum Schuljahr 2008/09 sowie die Senkung des Schuleintrittsalters (Früheinschulung) um sechs Monate zum Schuljahr 2005/06. Ziel der beiden Reformen war zum einen die Einführung von JÜL in allen Klassen der Schulanfangsphase in Berlin und somit die Sicherung des Anspruchs eines integrativen Schulanfangs und zum anderen die Senkung des mittleren Einschulungsalters als Beitrag (a) zu einer langfristigen Senkung des Berufseintrittsalters sowie (b) hin zu einer kindgerechten Schule anstelle der Erwartung einer Schulfähigkeit des Kindes.

Ziel der vorliegenden Dissertation ist die Prüfung des Implementationserfolgs dieser beiden Reformen entlang ihrer Ziele. Für die Operationalisierung des Implementationserfolgs wird auf die vier Erfolgsindikatoren (Identifikation, Nachhaltigkeit, Tiefe, Verbreitung) nach Coburn (2003) sowie die im Rahmen von Rogers Diffusionstheorie (2003) beschriebenen Anwendungskategorien zurückgegriffen. Dabei sollen folgende Forschungsfragen beantwortet werden:

Studie I: Wie wurde das Jahrgangsübergreifende Lernen an Berliner Grundschulen umgesetzt?

Studie II: Unterscheiden sich Schulen entsprechend ihrer Zuordnung zu einem Implementationstypen in ihrer Schul- und Unterrichtsqualität?

Studie III: Wie wurde die Senkung des Schuleintrittsalters an Berliner Grundschulen umgesetzt?

Studie IV: Wie wirkt sich die Senkung des Schuleintrittsalters auf die Schüler*innenleistung aus?

Studie V: Lassen sich systematische Leistungsunterschiede zwischen vergleichsweise jüngeren und älteren Schüler*innen (relative Alterseffekte) über verschiedene Schüler*innengruppen, Jahrgänge, Kohorten und Klassen verallgemeinern?

Für die Beantwortung der Forschungsfragen dienen Daten der Schulstatistik, der Schulinspektion, der Orientierungsarbeiten der zweiten Jahrgangsstufe und der Vergleichsarbeiten der dritten und achten Jahrgangsstufe sowie daran anschließende repräsentative Erweiterungsstudien als Grundlage.

Die Ergebnisse zur Einführung des Jahrgangsübergreifenden Lernens zeigen, dass JÜL während des Verpflichtungszeitraums (2008/09 bis 2012/13) in einem Großteil der Schulen und Klassen verbreitet war, aber dass es auch immer Schulen gab, die keine ihrer SAPH-Klassen jahrgangsübergreifend organisiert haben (Studie I). Weiterhin lassen sich die untersuchten Schulen in Anlehnung an Rogers Diffusionstheorie (2003) in sechs Typen der JÜL-Implementation unterscheiden („Pioniere“, „Überzeugte“, „Pflichtbewusste“, „nichtüberzeugte Skeptiker“, „Nicht-Überzeugte“ und „Verweigerer“). Die Typen liefern Hinweise, dass JÜL in etwa der Hälfte der Schulen nachhaltig umgesetzt wird und diese Schulen sich mit JÜL als Form der Unterrichtsorganisation identifizieren (Coburn, 2003). Im Hinblick auf Unterschiede in der Schul- und Unterrichtsqualität zwischen den Implementationstypen (Studie II) wird vor allem auf der Ebene der Schule eine höhere Qualität für die Schulen festgestellt, die JÜL bereits vor der Reform umgesetzt haben. Auf der Unterrichtsebene zeigen sich nur wenige Unterschiede in der Qualität zwischen den Typen. Werden die Typen, die JÜL bereits langfristig umsetzen („Pioniere“ und „Überzeugte“), nochmals danach unterschieden, wie viele Jahrgänge sie in einer JÜL-Klasse zusammenfassen (JÜL-2 vs. JÜL-3) so zeigen sich insbesondere auf der Ebene des Unterrichts bedeutsame Qualitätsunterschiede zugunsten der Schulen, die in einer JÜL-Klasse drei Jahrgangsstufen zusammenfassen (JÜL-3). Insgesamt kann jedoch nur ein vergleichsweise geringer Anteil der Varianz in der Schul- und Unterrichtsqualität auf die Zuordnung zu einem Implementationstypus zurückgeführt werden.

In Bezug auf die Studien zur Senkung des Schuleintrittsalters wurde festgestellt, dass das mittlere Schuleintrittsalter fünf Schuljahre nach Einführung der Reform um fünf Monate gesenkt war, was einem Monat weniger entspricht als vorgesehen. Außerdem konnten in der Verbreitung der Reform keine systematischen Unterschiede der Umsetzung über Klassen hinweg festgestellt werden (Studie III). Bezüglich der Auswirkungen auf die Schüler*innenleistung (Studie IV), zeigten sich in der zweiten Jahrgangsstufe systematische

Benachteiligungen (relative Alterseffekte) der vergleichsweise jüngeren Schüler*innen, die bereits in der dritten Jahrgangsstufe deutlich geringer ausfielen. In der achten Jahrgangsstufe waren keine systematischen Leistungsunterschiede zwischen vergleichsweise jüngeren und älteren Schüler*innen festzustellen. Da sich gleichzeitig die Anteile der vergleichsweise jüngeren und älteren Schüler*innen nicht bedeutsam veränderten, ist nicht davon auszugehen, dass diese Leistungsannäherung aufgrund positiver Selektionsprozesse (z.B. durch erhöhte Klassenwiederholungen der vergleichsweise Jüngeren) stattgefunden hat. Im Hinblick auf den Besuch der weiterführenden Schule sind keine Unterschiede festzustellen: vergleichsweise Jüngere besuchen demnach genauso häufig ein Gymnasium, wie ihre älteren Mitschüler*innen. Darüber hinaus konnte gezeigt werden, dass lediglich maximal 1,1 % der Varianz der Schüler*innenleistungen mit Hilfe des Einschulungsalters erklärt werden kann. Abschließend wurde festgestellt, dass die in Studie IV beschriebenen Erkenntnisse bezüglich relativer Alterseffekte über verschiedene Schüler*innengruppen, Jahrgänge, Kohorten und Klassen hinweg verallgemeinert werden können (Studie V).

Insgesamt weisen die Ergebnisse der Studien darauf hin, dass die Implementation von JÜL auf der Basis der Indikatoren von Coburn (2003) und Rogers als nicht erfolgreich eingeschätzt werden kann, da JÜL weder in allen Schulen umgesetzt wurde, noch die Schulen die JÜL umsetzten, dies tiefgreifend taten. Die Senkung des Schuleintrittsalters hingegen wurde an allen Schulen umgesetzt und es ist größtenteils gelungen, die vergleichsweise jüngeren Schüler*innen bezüglich ihrer schulischen Leistungen zu integrieren. Ob die Senkung des Schuleintrittsalters auch Auswirkungen auf die Lehr-Lern-Prozesse mit sich brachte, konnte im Rahmen der Arbeit nicht untersucht werden.

Die Befunde der Dissertation liefern wertvolle Hinweise für die Umsetzung von Bildungsreformen und die Einschätzung des Implementationserfolgs. Die Arbeit stellt somit eine sinnvolle Erweiterung des aktuellen bildungspolitischen und implementationstheoretischen Diskurses dar.

Summary

Reforms are one important driving force for development and quality improvement in the education system. As a response to the results of international high stakes tests, the state of Berlin introduced a new school law. The law included a comprehensive revision of all parts of the Berlin school system. The revision of the first two grades (school entrance phase – SEP) was one major concern in the reform of elementary education. Two reforms within that revision received major political and public attention: the compulsory introduction of multi-age teaching (MAT) with the beginning of the 2008/09 school year and the reduction of the school entrance age by six months by the 2005/06 school year (early enrolment). The aims of both reforms were on the one hand, the long-range introduction of multi age teaching in all SEP- classrooms ensuring the claim of an integrative school start. On the other hand, starting school earlier was to contribute to (a) reduce the age at entry into the labor market as well as (b) support the development towards a school appropriate for its children instead of children being appropriate for their school.

The goal of the present dissertation is to study the implementation success of these two reforms according to their aims. In order to operationalize implementation success, I draw on Coburns (2003) four dimensions of success (depth, shift in reform ownership, spread, and sustainability) as well as Rogers (2003) five adopter categories of his diffusion of innovations theory. In doing so, the following questions are to be answered:

Study I: How was multi-age teaching implemented in Berlin elementary schools?

Study II: Are there differences in school and teaching quality to be identified between different types of implementation?

Study III: How was the reduction of the school entrance age implemented in Berlin elementary schools?

Study IV: How did the reduction of the school entrance age effect student achievement?

Study V: Can systematic achievement differences between relatively younger and older children (relative age effects) be generalized across students, time, and classes?

In order to solve these questions data from the official school statistic, the school inspection, state-wide proficiency assessment programs where student participation was compulsory in second, third, and eighth grade as well as subsequent representative additional subsamples were used.

The results on the introduction of multi-age teaching show that while it was compulsory (2008/09 – 2012/13) MAT was distributed in the majority of Berlin elementary schools and classes. However, at each time point, there were schools that had implemented MAT in neither of their SEP-classes (Study I). Further, following Rogers diffusion of innovations theory (2003), schools were distinguished into six types of MAT-implementation (“pioneers”; “the convinced”; “the dutiful”, “non-convinced skeptics”, “Non-convinced”, and “deniers”). These types provide evidence that approximately half of the schools had sustainably implemented and identified with the MAT-reform (Coburn, 2003). With regard to differences in school and teaching quality amongst the implementation types (Study II), especially school quality is rated higher for schools who had implemented MAT prior to the reform. Concerning teaching quality the results show only few differences. However, if those schools who sustainably implemented MAT were further divided according to the number of grades they had summarized in their MAT-classes (MAT-2 vs. MAT-3), differences in teaching quality were identified in favor of those schools who include three grades in their MAT-classes (MAT-3). Overall, only a relatively small part of the variance in school and teaching quality can be explained by the attribution to one of the six implementation types.

In relation to the studies on the reduction of the school entrance age I find that the average school entrance age was reduced by five months, five years after the reform was introduced. The goal of the reform was a reduction by six months. Furthermore, no systematic differences in the implementation were found across classes (Study III). Concerning the effects on student achievement (Study IV), systematic disadvantages for the relatively younger students (relative age effects) were found in second grade. By the time the students reached third grade, these effects had already substantially reduced and no relative age effects were identified in grade 8. At the same time, the proportions of the different age groups did not substantially change across grades, which indicates no selection bias (e.g. due to grade retention) within the age effects. With regard to secondary school types, no differences between students of different age were found. Indicating that a relatively younger student visiting a Gymnasium is as common as for a relatively older student. Moreover, the results showed that a maximum of 1.1 % of the variance in student’s achievement was explained by school entrance age. Finally, results of the relative age effects were found to be generalizable across students, time, and classes (Study V).

In conclusion, the results of the studies indicate that, on the basis of the indicators by Coburn (2003) and Rogers (2003), the implementation of MAT cannot be evaluated as being successful because neither was MAT implemented in all Berlin elementary schools at any time nor did those schools who had implemented MAT seem to realize the reform in a profound matter. On

the contrary, the reduction of the school entrance age was implemented in all Berlin elementary schools. Further, concerning academic achievement, it was largely accomplished to integrate the relatively younger. Whether the enrolment reform had an impact on teaching and learning processes could not be addressed within this thesis.

The findings of this dissertation provide valuable indications for the implementation of education reforms and the assessment of implementation success. The work thus contributes to the current discourse in education policy and implementation theory.

Einleitung

Die Reformbedürftigkeit des deutschen Schulsystems zeigte sich um die Jahrtausendwende in den unterschiedlichsten Bereichen: die Anteile der Kinder, die von der Einschulung zurückgestellt wurden, stiegen kontinuierlich an (Autorengruppe Bildungsberichterstattung, 2016); jedes fünfte Kind hat im Laufe seiner Schullaufbahn eine oder mehrere Klassen wiederholt (Stanat et al., 2002), Hauptschulen wurden als „Resteschulen“ verstanden (z. B. Breyvogel, 2012) und 7,3 % der Schüler*innen verließen im Jahr 1999 die Schule gänzlich ohne Abschluss (Konsortium Bildungsberichterstattung, 2006). Nicht zuletzt die Ergebnisse der internationalen Schulleistungsstudien (vgl. u. a. Mullis, Martin, Gonzalez & Chrostowski, 2004; OECD, 2002) waren Anlass der Bildungspolitik für eine weitreichende Überarbeitung des Schulwesens und die sogenannte empirische Wende (Helmke, 2014). Das Ausmaß dieser Wende lässt sich an der in 2006 von der Kultusministerkonferenz der Länder (KMK) verabschiedeten Gesamtstrategie zum Bildungsmonitoring verdeutlichen (KMK, 2006): neben der weiteren regelmäßigen Teilnahme an internationalen Schulleistungsstudien (PIRLS/IGLU, TIMSS-Grundschule, PISA) einigten sich die Kultusminister*innen auf die gemeinsame Bildungsberichterstattung von Bund und Ländern, die Überprüfung bzw. Umsetzung von Bildungsstandards für die Primarstufe, die Sekundarstufe I und die Allgemeine Hochschulreife (IQB, 2017b) sowie den Einsatz von Verfahren zur Qualitätssicherung auf Ebene der Schulen, die sogenannten Vergleichsarbeiten (u. a. ISQ, 2017).

Eine logische Folge der empirischen Wende ist die Outputorientierung in der Bewertung von Lehr- und Lernprozessen. So sind beispielsweise Lehrkräfte dazu angehalten, mithilfe der Rückmeldungen aus diesen Monitoringinstrumenten ihren Unterricht verstärkt datengestützt zu evaluieren und weiterzuentwickeln (KMK, 2014). Mit der Einführung der Vergleichsarbeiten werden auf der anderen Seite die Kompetenzstände der Schüler*innen standardisiert und nicht mehr, wie im Falle von Schulnoten, auf eine soziale, sondern auf eine kriteriale Norm bezogen (ISQ, 2017). Ewald Terhart fasst diese für Deutschland neue Richtung in der Bildungspolitik wie folgt zusammen (Terhart, 2002):

*Entscheidend und in gewisser Weise tatsächlich revolutionär für den Schulbereich ist es, sich bei der Steuerung nicht länger nur am Prinzip einer immer detaillierteren Vorgabe von Inputs (Gesetze, Lehrpläne, Erlasse, Stundentafeln, Ordnungen), sondern verstärkt an den Outputs bzw. Outcomes, also an **tatsächlich erreichten Effekten und Wirkungen**, zu orientieren – und diese mit gesetzten Standards zu vergleichen. So etwas hat Folgen: Nur*

zu behaupten, dass ein Mehr an Investitionen hier oder dort eben hier oder dort dann schon gesteigerte Effekte nach sich ziehen werde, ist nicht mehr ausreichend – es geht um tatsächlich zustande kommende Wirkungen, und zwar Wirkungen auf der Seite der Schüler, denn die Schule ist letztlich für die Schüler da. (S. 104; Hervorhebung im Original, d. Verf.)

Darüber hinaus galt es für Reformen, die in vielen Bereichen des öffentlichen Lebens einen wesentlichen Motor für die Weiterentwicklung von Organisationen, Prozessen und Akteur*innen darstellen, sich künftig an den ihnen zugeschriebenen Outputs und Outcomes messen zu lassen. Und nicht nur das, auch sollte die Outputorientierung bereits Eingang in die Überlegungen und Konzeptualisierungen möglicher Reformanstrengungen finden (Gräsel, 2014).

Auch das Land Berlin hat den in PISA und ähnlichen Studien konstatierten Optimierungsbedarf zum Anlass genommen, eine umfassende Reform des Berliner Schulwesens und damit des Berliner Schulgesetzes in die Wege zu leiten. Einige Bereiche dieser Reform wurden wissenschaftlich begleitet und ausgewertet (Maaz, Baumert, Neumann, Becker & Dumont, 2013), andere sollten mithilfe der Erkenntnisse aus wissenschaftlich begleiteten Modellvorhaben (Kucharz & Wagener, 2007; Ramseger, 2004) in die Breite gebracht werden. Zu den zwei wichtigen und kontrovers diskutierten Neuerungen im Bereich des Schulanfangs in Berlin, der Einführung des Jahrgangsübergreifenden Lernens (JÜL) und der Senkung des Schuleintrittsalters (Früheinschulung), gibt es jedoch keine umfassenden Untersuchungen. Die vorliegende Dissertation befasst sich mit diesen zwei Reformen. Anhand des Implementationserfolgs sollen sie ex post (Wottawa & Thierau, 2003) evaluiert werden. Im Fokus stehen dabei die summative Bewertung der Umsetzung der Reformen und die daraus resultierenden Auswirkungen. Zu diesem Zweck ist diese empirische Arbeit wie folgt aufgebaut: zunächst wird der theoretische Rahmen für die Evaluation abgesteckt. Dabei beschreibt Kapitel 1 auf Grundlage implementationstheoretischer Überlegungen Merkmale und mögliche Operationalisierungen des Implementationserfolgs. Der zweite Teil des theoretischen Rahmens (Kapitel 2) befasst sich mit Bedingungen und Erkenntnissen zum Schulanfang in Deutschland und führt in die spezifische Situation in Berlin ein. Anschließend werden aus dieser Rahmung die Ziele der Arbeit formuliert und konkrete Forschungsfragen abgeleitet (Kapitel 3). Die folgenden beiden Kapitel beschreiben die im Rahmen dieser Arbeit durchgeführten Untersuchungen zum jahrgangsübergreifenden Lernen (Kapitel 4) und der Senkung des Schuleintrittsalters (Kapitel 5). Die Studien werden dabei jeweils theoretisch eingeführt, die konkreten Forschungsfragen und Methoden abgeleitet sowie ihre Ergebnisse im

Detail beschrieben. Ein abschließendes Fazit fasst die wichtigsten Erkenntnisse für jedes Kapitel zusammen. Abschließend werden die Erkenntnisse aus den Untersuchungen in Kapitel 6 vor dem Hintergrund des aktuellen Forschungsstandes und der bildungspolitischen Entwicklung diskutiert. Zudem werden die Grenzen der hier vorgelegten Arbeit aufgeführt und aus den Studienergebnissen Implikationen für Bildungspolitik und Forschung formuliert.

1. Implementationserfolg von Schulreformen

Für die Einschätzung des Erfolgs eines Implementationsprozesses, nicht nur im Bildungswesen, ist es zunächst notwendig, sich mit den Merkmalen und Bedingungen einer Reform und ihrer Umsetzung zu befassen, da diese unmittelbare Auswirkungen auf die Ziele und Ergebnisse haben (Goldenbaum, 2012). Das folgende Kapitel stellt deshalb die Merkmale und Bedingungen aus implementationstheoretischer Sicht dar und bildet somit die Grundlage für die Beschreibung des Implementationserfolgs der Reformen in der Berliner Schulanfangsphase (SAPH). Neben einer kurzen Klärung der zugrunde liegenden Begriffe (Abschnitt 1.1) werden die Entwicklung und der aktuelle Diskurs der bildungswissenschaftlichen Implementationsforschung zusammengefasst. Dabei wird zunächst der Implementationsprozess (Abschnitt 1.2) und relevante Implementationsstrategien beschrieben (Abschnitt 1.3). Der anschließende Teil (Abschnitt 1.4) bildet den eigentlichen Schwerpunkt dieses Kapitels und betrachtet zunächst verschiedene Einflüsse auf den Implementationserfolg (Abschnitt 1.4.1), bevor Möglichkeiten der Operationalisierung von Implementationserfolg (Abschnitt 1.4.2) vorgestellt werden. Abschnitt 1.5. fasst die grundlegenden Aspekte zusammen.

1.1 Begriffsklärung

Unter Reformen werden im Rahmen dieser Arbeit, im Anschluss an Goldenbaum (2012), „bewusst intendierte, geplante und zielgerichtete Veränderungen“ verstanden, „die als Reaktion auf als defizitär eingeschätzte soziale oder gesellschaftliche Gegebenheiten durchgeführt werden“ (ebd., 2012, S. 75). Demnach können Reformen als eine Art Korrektiv gesehen werden, die im Falle des Bildungswesens die Rückgewinnung von dessen Funktionalität bzw. Verbesserung, aber auch die Sicherung der Qualität des Schulsystems zum Ziel haben (Goldenbaum, 2012; Hasselhorn, Köller, Maaz & Zimmer, 2014) und sowohl Handlungen als auch Leistungen schulischer Akteur*innen beeinflussen und verändern sollten. In der aktuellen Debatte stellen Reformen vor allem „formale Absichten zur Veränderung respektive Verbesserung dar, die vom politisch-administrativen System ausgehen und durch Handlungen konkretisiert werden müssen“ (Goldenbaum, 2012, S. 77). Reformen sind folglich politisch legitimiert und definieren den rechtlichen Rahmen für schulische Veränderungen (ebd.).

Die konkreten Handlungen zur Umsetzung von Bildungsreformen können als Innovationen bezeichnet werden. Mithilfe von Innovationen sollen Reformen auf allen Ebenen¹ umgesetzt werden. Der Begriff der Innovation steht hier häufig für eine Veränderung im Sinne von Progression (Goldenbaum, 2012) oder einen geplanten Prozess „mit dem Ziel einer wünschenswerten Veränderung bzw. Verbesserung, die zu einer signifikanten Änderung im Status quo“ führt (Hunneshagen, 2005, S. 17) und ist demnach positiv konnotiert (Bormann, 2011). So einig man sich zu sein scheint über die grundlegende Bedeutung dieses Begriffs, so sehr differieren die Ausführungen zur Konkretisierung: z. B. wird bei Rogers und Shoemaker (1971) bereits bei jedem neuen Einfall, Gedanken oder Verhalten, der bzw. das sich qualitativ von den bisher bestehenden Formen unterscheidet, von einer Innovation ausgegangen (Hunneshagen, 2005). Goldenbaum spricht daher von einem „Begriff mit hoher semantischer Unschärfe“ (2012, S. 69) und Bormann fasst zusammen, dass es sich hierbei um keinen „wissenschaftlichen Substanzbegriff“ (2011, S. 43) handelt. Insgesamt lassen sich Reform und Innovation nicht immer eindeutig voneinander abgrenzen und werden daher häufig synonym verwendet. Im Rahmen dieser Arbeit werden die untersuchten Reformen, die Einführung des Jahrgangsübergreifenden Lernens sowie die Senkung des Schuleintrittsalters, als einzelne Reformen oder Neuerungen verstanden, die Teil der umfassenden Reform des Schulgesetzes in Berlin sind und deren einzelne Komponenten, z. B. Kooperatives Lernen, Heterogenität als Wert, als Innovationen, also konkrete operative Veränderungen (Hunneshagen, 2005), bezeichnet werden können.

Unabhängig von der Einbettung einer Innovation in einen bestimmten Kontext (z. B. ein Reformvorhaben) wird der Prozess der Übertragung einer Innovation in die Praxis als Implementation bezeichnet². Eine Implementation geht dabei „über die Entscheidung für eine Innovation hinaus und beginnt mit deren Realisierung in der Praxis“ (Goldenbaum, 2013, S. 152) und endet erst, wenn die Innovation als Standardpraktik übernommen wurde (Hasselhorn et al., 2014; Klein & Knight, 2005). Dieser letzte Schritt wird häufig auch als

¹ Zur Beschreibung dieser Ebenen wird in der Bildungsforschung häufig, die u. a. von Helmut Fend (2009) vorgeschlagene Unterscheidung in Makro-, Meso- und Mikroebene verwendet. Dabei wird unter der Makroebene die grobe Strukturierung langer und differenzierter Lernprozesse, wie z. B. auf der Ebene des Bildungssystems oder der Bildungspolitik, unter der Mesoebene z. B. ein lokales Schulprogramm oder ein Schulverwaltungsbezirk und unter der Mikroebene schließlich die konkrete Vorbereitung einer Unterrichtsstunde durch eine Lehrkraft verstanden (ebd., S. 31).

² Der Begriff der Implementation wird auch für die Umsetzung von Ideen, Programmen oder zusammenhängenden Aktivitäten und Strukturen verwendet (u. a. Fullan, 2007, S. 87).

Institutionalisierung bezeichnet (vgl. u.a. Altrichter & Wiesinger, 2005; Bormann, 2011). Wie dieser Prozess im Einzelnen beschrieben werden kann, wird im folgenden Abschnitt dargestellt.

1.2 Der Implementationsprozess

Die Erforschung der Implementation von Innovationen und Reformen hat ihren Ursprung einerseits in den Wirtschaftswissenschaften, wobei erforscht wurde, wie sich verschiedenartige Innovationen in ein Unternehmen übertragen lassen (Goldenbaum, 2012), andererseits wurde in den Politikwissenschaften die erfolgreiche Verankerung größerer und kleinerer Reformen in das politische System untersucht (Luchte, 2005). Doch auch in den Bildungswissenschaften gab es seit den 1990er Jahren verstärkt Überlegungen, wie der Erfolg von Reformbemühungen sichergestellt werden kann (Luchte, 2005). Bildungspraktiker*innen und -politiker*innen hatten bereits seit Längerem Schwierigkeiten mit der konsistenten und verlässlichen Verbreitung von Reformansätzen in Schulen oder deren Verwaltungsbezirken festgestellt. Dabei schien zunächst nicht klar zu sein, ob diese Schwierigkeiten auf die mangelnde Qualität der Programmkonzepte oder die unzureichende Umsetzung auf lokaler Ebene zurückzuführen sind (Supovitz & Weinbaum, 2008). In der frühen Phase der schulischen Implementationsforschung, die sich noch stärker an den Ansätzen der Politik- und Wirtschaftswissenschaften orientierte, herrschte ein eher rational-empirisches, technologisches Verständnis des Prozesses (Goldenbaum, 2012; Odden, 1991; Supovitz & Weinbaum, 2008). Reformbestrebungen wurden nach diesem Verständnis in der Regel von oben, also durch die Bildungspolitik angeordnet oder von Wissenschaftler*innen erarbeitet und initiiert (Gräsel & Parchmann, 2004), wobei der Prozess dann als erfolgreich eingestuft wurde, wenn z. B. die Planung für den Unterricht gemäß den Vorgaben realisiert und nur wenig verändert wurde (ebd.). Weiterhin wurde angenommen, dass fähige Lehrkräfte in der Lage sind, eine Innovation mithilfe von Materialien und Handbüchern zu erfassen, ihr Verhalten den neuen Anforderungen anzupassen, auch wenn es zum Teil radikale Veränderung erforderte, und die Innovation gemäß den Vorgaben exakt umzusetzen (Gräsel & Parchmann, 2004, S. 200).

Bei nicht erfolgreicher Implementation einer Reform wurde daher häufig auf Nicht-Wollen oder gezielte Sabotage seitens der Zielgruppen, z. B. Lehrkräfte, verwiesen (Supovitz & Weinbaum, 2008). Im Laufe der folgenden Jahrzehnte verbreitete sich mehr und mehr die Ansicht von Implementation als einem Interaktionsprozess zwischen initiiender und umsetzender Instanz

(ebd.), bei dem sich die Anforderungen und Erwartungen der jeweiligen Rollen in verschiedenen Bereichen überschneiden.

Nach heutigem empirischen und theoretischen Erkenntnisstand wird in der Bildungsforschung mehrheitlich von einem Vier-Phasen-Modell des Implementationsprozesses ausgegangen (Goldenbaum, 2012; Hunneshagen, 2005): (1) Initiation – diese Phase dient der Vorbereitung und es stehen u. a. Aktivitäten wie die Feststellung der Ausgangslage, Zieldefinition und die Erstellung eines Umsetzungsplans im Fokus; (2) (erste) Umsetzung – hier steht die Durchführung im Vordergrund und es zählen Tätigkeiten wie die Entwicklung, Aneignung und Vermittlung der Neuerung dazu; (3) Evaluation – in dieser Phase werden Informationen rund um die Neuerung sowie die Neuerung selbst bewertet und daran anschließend Entscheidungen bezüglich der Neuerung getroffen; (4) Institutionalisierung – diese abschließende Phase beinhaltet verankernde Aktivitäten, wie die Stabilisierung und Verbreitung der Neuerung. Dabei gilt eine Implementation als abgeschlossen, wenn sie die Phase der Institutionalisierung erreicht hat. Bis zu diesem Punkt können nach Desimone (2002, S. 455) fünf bis zehn Jahre, nach Hameyer (1978) im Falle von Bildungsinnovationen 15 bis 20 Jahre und bei der Implementation einer Innovation in ein bestehendes System – was im Bereich der Schule die Regel sein sollte – sollten für den Prozess fünf bis sieben Jahre eingeplant werden (Fullan, 2007)). Tabelle 1 fasst die vier Phasen mit ihren jeweiligen beispielhaften Tätigkeiten zusammen.

Tabelle 1: Tätigkeiten im Implementationsprozess nach Goldenbaum (2012, S. 91)

Phase	Tätigkeiten
Initiation	Vorbereitende Tätigkeiten: u. a. Analyse der Ausgangssituation, Festlegung von Zielen, Konkretisierung von Startidee und Auftrag, Erstellung eines Umsetzungsplans, Teambildung, Kontaktaufbau zu Kooperationspartner*innen
Umsetzung	Durchführende Tätigkeiten: u. a. Entwicklung, Aneignung und Vermittlung der Innovation, eventuell auch Nutzung von Unterstützungsmaßnahmen und Durchführung von Zwischenevaluationen
Evaluation ¹	Auswertende Tätigkeiten: i. d. R. das Sammeln von Informationen, die Bewertung der Innovation anhand der Informationen und das Treffen von Entscheidungen in Bezug auf die Innovationen
Institutionalisierung	Verankernde Tätigkeiten: u. a. Stabilisierung, Verbreitung (z. B. durch Anreizsysteme, Weiterbildungen); Innovation wird als Standardpraktik angewendet

¹ Evaluation wird nicht in allen Modellen als einzelne Phase beschrieben (Goldenbaum, 2012) und auch in einer Zusammenfassung von Reinmann & Vohle (2004) nicht inhaltlich spezifiziert.

Kritik an den Phasenmodellen wird bezüglich der postulierten Linearität und den somit nicht berücksichtigten rekursiven Schleifen, Umwegen und Rückschlägen, der exakten Planbarkeit des Implementationsprozesses bezogen auf diese vier Phasen, sowie an den in den Modellen wenig thematisierten Übergängen zwischen den einzelnen Phasen geübt (Goldenbaum, 2012). Zudem bleibt in diesem Prozess zunächst ungeklärt, in welcher Qualität und Intensität sowie mit welcher Dauerhaftigkeit und Treue, kurz gesagt mit welchem Erfolg eine Reform oder Neuerung implementiert wurde (vgl. Abschnitt 1.4). Bezogen auf die Umsetzung einer Innovation wird zwischen ungeplanter (Diffusion) und gesteuerter (Dissemination, Transfer) Verbreitung unterschieden (Jäger, 2004). Bei der Diffusion von Innovationen und Reformen liegt der Blick auf der Makroebene, also das Gesamtsystem betreffend, und die ungeplante Verbreitung erfolgt vordergründig über einzelnen Systemelemente, wie Lehrkräfte, Schulleitungen oder Einzelschulen. Bei der Dissemination handelt es sich ebenfalls um einen auf der Makroebene stattfindenden Prozess, jedoch erfolgt hier die Verbreitung zielgerichtet und geplant. Ebenfalls gesteuert wird die Verbreitung bei einem Transfer von Neuerungen, wobei hier der Fokus auf den Einzelschulen, ihren Schulleitungen und Lehrkräften liegt. Bei Transferprozessen steht die Übertragung von Neuerungen von einem Kontext in einen anderen im Vordergrund, wobei sich die Kontexte in unterschiedlichen Merkmalen unterscheiden (z. B. Inhalt, Struktur und Person; ebd.).

1.3 Implementationsstrategien

Der gesamte Implementationsprozess ist geprägt von den Strategien, mit welcher Reformen und Innovationen implementiert werden sollen. Aus der Forschungsliteratur zur Implementation im Bildungswesen lassen sich vier wesentliche Strategien zusammenfassen: (1) *Top-down-Strategien*, (2) *symbiotische Strategien* und (3) *Bottom-up-Strategien* (vgl. u.a. Goldenbaum, 2012; Gräsel & Parchmann, 2004; Hasselhorn et al., 2014) und in den letzten zwei Jahrzehnten auch (4) *Outputorientierte Strategien* (Gräsel, 2014).

Bei der Anwendung von *Top-down-Strategien* erfolgen sowohl die Initiation als auch die Festlegung der Inhalte und Ziele von einer externen, den Schulen meist übergeordneten Instanz (Goldenbaum, 2012; vgl. Gräsel & Parchmann, 2004) und entsprechen somit einem eher rational-empirischen oder technologischen Verständnis von Implementation. Top-down-Strategien werden danach unterscheiden, ob eine Reform als erfolgreich implementiert gilt, wenn sie wie vorgegeben umgesetzt wurde (*fidelity perspective*) oder es Raum für Anpassungen

und Veränderungen gab (*mutual adaption*). Die erstere der beiden Umsetzungsmöglichkeiten der Top-down-Strategie – auch programmierte Strategie (Altrichter & Wiesinger, 2005) oder Replikation (Coburn, 2003) – gilt als angemessen, wenn das Ausmaß der angestrebten Neuerung eher gering ist oder ihre Umsetzung auch schrittweise verwirklicht werden kann. Zudem sollte die Innovation bei den Adressat*innen weitgehend begrüßt werden und die umsetzende Institution in eine weitgehend stabile Umwelt eingebettet sein (Altrichter & Wiesinger, 2005). Die zweite Form – auch adaptiv-evolutionäre Strategie (ebd.) oder normativ-reduktive Strategie (Goldenbaum, 2012) – einer Implementation von oben sollte dann in Betracht gezogen werden, wenn die eben genannten Bedingungen (Akzeptanz, Stabilität, usw.) nicht gegeben sind und bei der Umsetzung institutionelle und organisatorische Rahmenbedingungen berücksichtigt werden sollten. Bezogen auf den Implementationsprozess ist beiden Perspektiven jedoch gemeinsam, dass vor allem die Phasen der Initiation und Umsetzung personell und zeitlich unabhängig voneinander zu sehen sind (Gräsel & Parchmann, 2004).

Von einer starken Überschneidung und Wiederholung einzelner Phasen des Implementationsprozesses ist bei der Anwendung *symbiotischer*³ Strategien auszugehen. Hier arbeiten Akteur*innen „mit unterschiedlicher Expertise“⁴ gemeinsam an der Umsetzung pädagogischer Innovationen“ (Gräsel & Parchmann, 2004, S. 205). Dabei kommt der Impuls zur Neuerung meist aus der pädagogischen Praxis und wird von „allen Beteiligten als bedeutsam und veränderenswert beurteilt“ (ebd., S. 206). Neben den vier bereits beschriebenen Phasen im Implementationsprozess (siehe Abschnitt 1.2) stellen Revision, Selbst- und Fremdevaluation sowie die kontinuierliche Prüfung der konkreten Maßnahmen zentrale Merkmale dar (Goldenbaum, 2012, S. 97). Auch die Umsetzung als solche erfolgt hier im ständigen Austausch zwischen den Beteiligten, sodass die Etablierung von Kooperationsstrukturen grundlegend ist für die erfolgreiche Anwendung von symbiotischen Implementationsstrategien.

Ebenfalls häufig durch die pädagogische Praxis initiiert sind Innovationen, die mithilfe der Bottom-up-Strategie implementiert werden. Beispielhaft für diese Neuerungen werden von Gräsel und Parchmann (2004) Schulentwicklungsprozesse genannt, bei denen zum Teil auch Erziehungswissenschaftler*innen zu Rate gezogen werden. Im Gegensatz zu den bereits

³ Der Begriff wurde von Gräsel und Parchmann geprägt. Er stammt aus der Biologie und „bezeichnet das Zusammenleben verschiedener Arten zum gegenseitigen Vorteil“ (2004, S. 205).

⁴ Z. B. aus pädagogischer Praxis, Wissenschaft und Bildungsadministration

beschriebenen Strategien liegt hier die Entwicklung der Implementationsziele in der Hand der Schulen. Sie sind es auch, die die konkreten Maßnahmen der Innovation entwickeln, erproben und evaluieren (Goldenbaum, 2012).

Historisch betrachtet, können symbiotische Strategien als Antwort auf den auf schulsystemischer Ebene mangelnden Erfolg in der Anwendung von zunächst Top-down- und anschließend Bottom-up-Strategien verstanden werden (Goldenbaum, 2012). Zusammenfassend kann für die im Rahmen dieser Arbeit untersuchten Reformen in der Berliner Schulanfangsphase festgehalten werden:

Für die Implementation von staatlich administrierten Neuerungen ist eine steigende Dezentralisierung deutlich geworden, die verbunden ist mit einer Erweiterung der Eigenverantwortlichkeit der Einzelschule und der Etablierung von Schulautonomie. Zentral werden dabei klare Rahmenrichtlinien und Rechenschaftsverpflichtungen auf der Makroebene vorgegeben; die eigentlichen Umsetzungs- und Entscheidungsprozesse liegen dezentral in der Verantwortung der Einzelschule, also auf der Mesoebene. (ebd.)

Die Berücksichtigung von Outputs stellt im deutschen Bildungssystem selbst und in der Bildungsforschung seit etwa 20 Jahren einen Schwerpunkt dar. Auch im Bereich der Implementation von Reformen im Bildungsbereich werden aktuell die Outputs der Schüler*innen, Lehrkräfte, Schulleitungen und Schulen verstärkt in den Blick genommen (vgl. Gräsel, 2014; O'Day, 2002). Bei einer *outputorientierten Implementationsstrategie* stehen die Ziele einer Reform, also z. B. das Halten oder Verbessern der durchschnittlichen Schüler*innenleistungen einer Schule, im Mittelpunkt. Die einzelnen Phasen des Implementationsprozesses sind ebenfalls auf diese Ziele ausgerichtet. Zentral ist dabei die Phase der Evaluation, bei der die Zielerreichung, z. B. in Form von regelmäßigen standardisierten Lernstandserhebungen, überprüft wird. O'Day (2002) formuliert in ihren Ausführungen drei Herausforderungen outputorientierter Schulentwicklung: (1) Die Schule stellt einerseits die Einheit der Innovation dar, aktiv werden müssen jedoch die Akteure innerhalb dieser Einheit. (2) Externe Einflüsse wollen interne Handlungen und Interaktionen beeinflussen. (3) Informationen verbreiten und verarbeiten ist ein anspruchsvoller, aber notwendiger Prozess für Qualitätsentwicklung in Schulen (ebd., S. 295 f.).

1.4 Implementationserfolg

Am Ende eines jeden Reformprozesses stellt sich die Frage einer Bewertung seiner Umsetzung, also die des Implementationserfolgs. Neben dem Implementationsprozess selbst und der angewandten Implementationsstrategie werden in der Literatur weitere Einflussfaktoren und Gelingensbedingungen für den Implementationserfolg beschrieben.

1.4.1 Einflüsse auf den Implementationserfolg

Mit der Anwendung unterschiedlicher Strategien zur Implementation werden auch verschiedene Ziele verfolgt (siehe Abschnitt 1.3). So wird in top-down organisierten Prozessen meist die Umsetzung einer Neuerung nach genauen Vorgaben als Ziel definiert, während eine outputorientierte Implementation die (Verbesserung der) Ergebnisse einzelner Akteure (z. B. Schüler*innen, Lehrkräfte) verfolgt. Folglich muss auch der Erfolg eines Implementationsprozesses an seinen Zielen gemessen werden⁵. Darüber hinaus werden national und international verschiedene Einflussfaktoren für den Erfolg von Implementationsprozessen zusammengefasst.

Aufgrund ihrer umfassenden vor allem praktischen Erfahrungen mit der Implementation von Innovationen im US-amerikanischen Bildungssystem formulieren Hall und Hord zwölf „Principles of Change“ (2015), wobei u. a. das Lernen als prägendes Merkmal, die Prozesshaftigkeit, die Schule als Einheit der Veränderung sowie der Einsatz gezielter Interventionen als einflussreiche Faktoren benannt werden. Für den deutschsprachigen Raum unterscheidet Jäger (2004) auf Grundlage theoretischer Überlegungen und der Durchführung eines Workshops mit Schulentwicklungsexpert*innen aus der Praxis zwischen förderlichen (Beteiligung von Schulleitung, Kollegium, Umfeld; inhaltliche Voraussetzungen: methodische Unterstützung, Nutzen; Maßnahmen: Planung, Umsetzung, Nacharbeitung) und hinderlichen Faktoren (geringe Anzahl beteiligter Kolleg*innen, formale Strukturen, Absichtserklärungen ohne Umsetzung). Als wesentliche Einflussfaktoren auf den Erfolg einer Innovation werden (Gräsel & Parchmann, 2004; Gräsel, 2010; Hasselhorn et al., 2014) zusammenfassend Merkmale der Innovation, Merkmale der Lehrkräfte, Merkmale der Einzelschule und Merkmale des Schulsystems unterschieden. Diese Differenzierung geht auf Fullans Faktoren erfolgreicher Implementation zurück (Fullan, 2000) zurück und wird vor allem in der

⁵ Der Vergleich von Reformzielen mit der tatsächlichen Umsetzung wird auch Implementationstreue oder *fidelity* Supovitz und Weinbaum (2008) genannt.

deutschsprachigen Implementationsliteratur immer wieder aufgegriffen. Eine anschauliche Zusammenfassung der vier Einflussfaktoren von Goldenbaum (2012, 2013) ist in Tabelle 2 dargestellt.

Tabelle 2: Einflussfaktoren auf die Implementation schulischer Innovationen (Goldenbaum, 2012, S. 115, 2013, S. 50)

Ebene	Merkmal
Innovation	Lösungsangebot für schulischen Bedarf, Klarheit über Zielstellungen und Vorgehensweisen, Komplexität, Qualität und Praktikabilität
Lehrkräfte	Qualifikationen/Kompetenzen, Einstellungen, Subjektive Theorien, Selbstwirksamkeitserwartungen, Motivation, (schulübergreifende) Kooperation, Anzahl beteiligter Lehrkräfte
Einzelstufe	Schulkultur, Schulleitung, (schulübergreifende) Kooperation, Steuergruppe
Schulsystem	Einbindung von Innovationen in Bildungsreformen, Stabilität schulischer Kontexte, Langfristige Unterstützung und Begleitung, Realistische Zeit- und Ressourcenplanung, Ausreichende Zeit, Informationssystem

Auf der Ebene der Innovation steht der wahrgenommene Wert oder Nutzen einer Innovation im Vordergrund (Hasselhorn et al., 2014). Dabei ist ebenfalls der (empfundene) Neuheitsgehalt von Bedeutung. Einerseits unterscheidet Gillwald (2000) in diesem Zusammenhang absolute von relativen Neuheiten. Andererseits ist nach Rogers (2003) vor allem die Unterscheidung von realer und empfundener Neuheit von Bedeutung. Es ist anzunehmen, dass eine Weiterentwicklung (relative Neuheit) eines im Großen und Ganzen funktionierenden Bereichs in der Schule mehr Zustimmung erfahren wird, als wenn für diesen Bereich eine absolute Neuheit implementiert werden sollte (Reinmann, 2005). Weishardt (2008) folgend, liegt eine Innovation bereits dann vor, wenn eine entsprechende Maßnahme in der Schule unbekannt ist.

Einen wesentlichen Einfluss auf den Implementationsprozess und dessen Erfolg stellen die Lehrkräfte dar, da in der Regel sie es sind, die ein innovatives Vorhaben in die pädagogische Praxis umsetzen müssen (O'Day, 2002). Dabei zeigt sich vor allem die Akzeptanz (Odden, 1991; Pant, Vock, Pöhlmann & Köller, 2008b), aber auch die Beurteilung einer Maßnahme als „nützlich, sinnvoll, realisierbar, wichtig usw.“ (Gräsel & Parchmann, 2004, S. 203) als wesentliche Merkmale. Hall und Hord (2015) definieren in diesem Zusammenhang sieben Stufen der Akzeptanz (*Stages of Concern*) einer Bildungsinnovation bei den beteiligten Akteur*innen (Lehrkräfte, aber auch Schulleitungen): beginnend mit keinem oder geringem Bewusstsein für eine Innovation (1), verändert sich der Fokus der beteiligten Akteur*innen von der Information (2) zur persönlichen Betroffenheit (3). Anschließend zum

Aufgabenmanagement (4), zu den Auswirkungen auf Lernende (5) und den Kooperationsmöglichkeiten (6), um mit dem Fokus auf Revision und Optimierung (7) abzuschließen. Auf der Basis dieser sieben Stufen entwickelten Hall und Hord (2015) das sogenannte *Concerns-Based Adoption Model*, welches die Auseinandersetzung beteiligter Akteur*innen mit einer Reform oder Neuerung zu erfassen versucht (für eine Übersicht vgl. Böse, 2015). Empirisch konnte das Modell für den deutschsprachigen Raum zumindest teilweise belegt werden (Böse, 2015; Pant, Vock, Pöhlmann & Köller, 2008a, 2008b). Weiterhin scheint für die Verbreitung einer Neuerung „die Vereinbarkeit [...] mit tief sitzenden Überzeugungen und sozial ausgehandelten Praktiken eine Schlüsselrolle [...] einzunehmen“ (Gräsel & Parchmann, 2004, S. 201). Empirisch können Trempler, Schellenbach-Zell und Gräsel (2013) aufzeigen, dass die (wahrgenommene) inhaltliche Relevanz des Innovationsthemas sich als stärkster Prädiktor vor allem für intrinsische, aber auch für autonome und external-motivierte Motivation erweist. Schließlich wird auf die Notwendigkeit der Professionalisierung und stetige Fort- und Weiterbildung (Odden, 1991) sowie Stärkung individueller Kompetenz hingewiesen (Gräsel & Parchmann, 2004). Als weitere Bedingung für einen gelingenden Implementationsprozess beschreibt Goldenbaum (2013) die Kooperation von Lehrkräften, die auch im aktuellen bildungswissenschaftlichen Diskurs stetig gefordert wird (vgl. Abschnitt 2.4; Terhart & Klieme, 2006). Hier scheinen vor allem die Organisation und Struktur von Schule (als lose gekoppeltes System; Weick, 1976) und Unterricht (Organisation in Stundenplänen, Breidenstein & Rademacher, 2013) große Herausforderungen für die Etablierung von Kooperation darzustellen (O'Day, 2002).

Auch haben Merkmale der Einzelschule Einfluss auf schulische Innovationen (Goldenbaum, 2013). So wird der Schulleitung eine wichtige Rolle zugeschrieben, da ihre Unterstützung die Einstellungen und Akzeptanz bei Lehrkräften und Eltern beeinflusst (Gräsel & Parchmann, 2004). Holtappels (1995) verweist in seinen Überlegungen zum Implementationsprozess auf die Bedeutsamkeit der Schulkultur mit ihren Komponenten der Erziehungs-, Lern- und Organisationskultur. Unter dem Begriff Schulkultur wird dabei das kulturelle Eigenleben der Schule mit spezifischen Wert- und Normgefühlen, inhaltlich methodischen Konfigurationen sowie Umgangs- und Verkehrsformen verstanden (ebd.). Beeinflusst wird dieses Gebilde einerseits von innen heraus in Form der drei erwähnten Komponenten sowie Selbstentwicklung und Selbstevaluation und andererseits von außen durch den Schulträger, das Umfeld, die Schulaufsicht und vorhandene Unterstützungsstrukturen. Die Sicht der Einzelschule als bedeutsame pädagogische Gestaltungseinheit und damit Ort relevanter Entscheidungen in Bezug auf Bildung hat sich vor allem in aktuellen (Reform-) Diskursen (Stichwort: „Neue

Steuerung“, für eine Übersicht: Altrichter & Maag Merki, 2010) herausgestellt (Goldenbaum, 2012; Koch & Gräsel, 2004).

Das vierte einflussreiche Merkmal schulischer Innovationen bildet das Schulsystem. Gerade in Top-down-Prozessen ist eine systemische Begleitung und Unterstützung auch über die Einführungsphase hinaus von großer Bedeutung. Hunneshagen (2005) ergänzt die Notwendigkeit von parallel laufenden und ineinandergreifenden Entwicklungen für einen erfolgreichen Implementationsprozess. Außerdem sind bei der sozialen Organisation Schule stets ihre besonderen Bedingungen (Bildungsauftrag, der pädagogische Zusammenhang, Schüler*innen im Mittelpunkt, Arbeitsteilung in gefügeartigen Kooperationen etc.) und der systemische Zusammenhang (jede Veränderung zeigt Auswirkungen an verschiedenen Stellen) zu berücksichtigen (ebd.).

1.4.2 Operationalisierung des Implementationserfolgs

Zur theoretischen Beschreibung einerseits, aber auch als Möglichkeit der Operationalisierung beschreibt Coburn (2003) in ihrem Aufsatz zum Transfer (engl. *scaling up*) von Schulreformen vier allgemeine Indikatoren erfolgreicher Reformimplementationen: Identifikation, Nachhaltigkeit, Tiefe und Verbreitung (Tabelle 3). (1) Da Individuen und Organisationen im Allgemeinen und auch schulische Akteur*innen und ihre Schulen im Speziellen in ihren Handlungen meist auf bisher erworbenes Wissen, gemachte Erfahrungen und entwickelte Vorstellungen und Werte aufbauen, neigen sie häufig dazu, eher solche Interventionen, Konzepte oder Programme zu bevorzugen, welche mit ihren bisherigen Praktiken übereinstimmen (Spillane, 2000). Coburn schlussfolgert daraus, dass Schulreformen in der Tiefe (*depth*) implementiert werden müssen, also tiefgreifende und konsequente Veränderungen des Unterrichtsgeschehens mit sich bringen sollten. Dieses Merkmal ist vor allem deshalb als hochrelevant einzuschätzen, weil auch Reformen, einschließlich solcher, die rein organisatorischer Natur scheinen, auf einem bestimmten pädagogischen Verständnis basieren, welches mit dem Implementationsprozess in die Unterrichtspraxis transportiert wird bzw. werden soll. Fördernd für eine Verbreitung in der Tiefe zeigten sich einer Studie von Goldenbaum (2013) zufolge innerschulische Kooperationen. (2) Als zweites Merkmal beschreibt Coburn (2003) die Nachhaltigkeit (*sustainability*) einer Reform. Da Reformen, vor allem, wenn sie nicht in der Tiefe implementiert wurden, nach einiger Zeit häufig aufgrund sich verschiebender Prioritäten, mannigfaltiger Ansprüche und Personalfuktuation wieder aufgegeben werden, stellt die Sicherung der Nachhaltigkeit einer Reform ein zentrales

Kriterium erfolgreicher Implementation dar. Dies gilt insbesondere für extern (z. B. durch die Bildungspolitik) initiierte Reformen, da diese häufig für einen kurzen Zugang zu zusätzlichen Ressourcen und Unterstützungssystemen sorgen, jedoch mit der Abkehr der Initiator*innen dieser Zugang entfällt.

Tabelle 3: Erfolgsindikatoren von Implementationsprozessen nach Coburn (2003)

Indikator	Beschreibung
Identifikation	Die Identifikation ist besonders bei top-down initiierten Reformen relevant, da hier eine external verstandene und unterstützte Theorie in eine internal verstandene und unterstützte theoriebasierte Praxis übergehen soll. Beinhaltet auch die Übertragung substanzieller und strategischer Entscheidungen von der initiiierenden Organisation an den Schulbezirk und die Schulleitungen.
Nachhaltigkeit	Die Verbreitung und Übernahme einer Innovation ist nur dann von Bedeutung, wenn sie in den Schulen dauerhaft umgesetzt wird. Besonders bei top-down initiierten Reformen, ist das eine Herausforderung, da es bei diesen Reformen meist eine kurze Phase der verstärkten Unterstützung durch materielle und personelle Ressourcen, Fort- und Weiterbildungen u.a. gibt und der Implementationsprozess auch nach Kürzung der Unterstützungssysteme voranschreiten soll.
Tiefe	Beschreibt tiefgreifende und konsequente Veränderungen in der Unterrichtspraxis, die über das bloße Ändern von Unterrichtsmaterialien, der Organisation im Klassenzimmer und kleine Aktivitäten hinausgehen. Ist von besonderer Relevanz für den Implementationsprozess, da Lehrkräfte und andere angesprochene Akteur*innen Reformen häufig im Sinne ihres bisherigen Wissens, ihren Erfahrungen und Einstellungen entsprechend anpassen und zu Praktiken tendieren, die am ehesten ihren aktuellen Praktiken entsprechen.
Verbreitung	Beschreibt einerseits die Verbreitung einer Reform in eine größere Anzahl an Klassen und Schulen. Andererseits ist hier auch die Verbreitung innerhalb von Klassen, Schulen und Schulbezirken gemeint, die die Anwendung reformspezifischer Normen und Prinzipien unterschiedlicher Akteur*innen einschließt. Die Annahme ist hierbei, dass eine Reform seitens der Schulen und Lehrkräfte eher dann nachhaltig und tiefgreifend implementiert wird, wenn Schulen und Schulbezirke ihre Aktivitäten auf die Reform ausgerichtet bzw. diese mit ihr vereinbar sind.

(3) Weiterhin lässt sich der Erfolg einer Reform an ihrer Verbreitung (*spread*) bemessen. Dabei trägt laut Coburn (ebd.) eine größere Verbreitung sowohl über Schulen hinweg als auch innerhalb von Schulen dazu bei, dass Lehrkräfte und Schulen eine Reform nachhaltig und vertiefend implementieren können. Dieses Merkmal kann vor allem für solche Reformen von Bedeutung sein, die bisherige Konventionen und institutionalisierte Ansätze in Schulen, aber vor allem bei ihren Akteur*innen herausfordern. Bezogen auf die verschiedenen Ebenen von Schule bedeutet eine Verbreitung auf der Mikroebene, dass Lehrkräfte Normen und Werte einer

Reform nicht nur auf spezifische reformverwandte Aktivitäten anwenden, sondern darüber hinaus auch in anderen Bereichen ihrer pädagogischen Praxis. Auf der Meso- bzw. Makroebene, also der Ebene der Schule und des Schulverwaltungsbezirks, bedeutet die Verbreitung einer Reform neben der Umsetzung in einer steigenden Anzahl von Klassen bzw. Schulen ebenfalls die Anwendung reformspezifischer Normen und Werte auf schulische bzw. verwaltungsbezirkliche Grundsätze und Routinen. (4) Schließlich kann eine Schulreform dann als erfolgreich gewertet werden, wenn sich schulische Akteur*innen mit ihr identifizieren (*shift in reform ownership*; übersetzt nach Gräsel, 2010), wenn also aus einer extern verstandenen und unterstützten eine intern verstandene und unterstützte Reform wird. Dies schließt auch den Transfer von Entscheidungsbefugnissen an die Schulverwaltungsbezirke und Schulleitungen ein und ist nach Auffassung von Coburn (2003) ein zentrales Element für die Nachhaltigkeit und Verbreitung bildungspolitischer Reformen.

Wie bereits dargestellt (Abschnitt 1.4.1), ist auf der Ebene der Lehrkräfte die Akzeptanz einer Neuerung von zentraler Bedeutung (Fullan, 2000; Pant et al., 2008b). Neben der Erfassung der Akzeptanz einer Neuerung bei Lehrkräften mithilfe des *Concerns-Based Adoption Models* (Hall & Hord, 2015) kann eine weitere Operationalisierung der Entwicklung der Akzeptanz einer Reform, nicht nur auf Ebene der Lehrkräfte, unter Anwendung der Diffusionstheorie (Rogers, 2003) erfolgen. Rogers geht dabei davon aus, dass Neuerungen in der Regel nicht sofort von allen Akteuren eines Systems, wie z. B. des Bildungssystems oder einer einzelnen Schule, akzeptiert und umgesetzt werden, sondern zeitversetzt entsprechend ihrer Innovationsbereitschaft (*Innovativeness*; ebd. S. 267 ff.). Er unterscheidet dabei fünf idealtypische Kategorien: *Innovative*, *Vorreiter*innen*, *frühes Hauptfeld*, *spätes Hauptfeld*, *Nachzügler*innen* (S. 282ff.; Übersetzung nach Jäger, 2004, 95f.). Diese Typen sollen dazu dienen, Implementation zu erfassen, und eine Charakterisierung oder Gegenüberstellung von Akteur*innen entsprechend deren Zugehörigkeit zu einer Kategorie ermöglichen (Rogers, 2003). Die Verteilung dieser Anwendungskategorien lässt sich mit einer Normalverteilung darstellen, die bei einem Anwenderanteil von 50 % ihren Mittelwert hat (Abbildung 1). Die Innovativen machen idealerweise 2,5 % der Anwender*innen aus und werden von Rogers als risikofreudig oder -bereit beschrieben. Sie müssen in der Lage sein, mit einem hohen Maß an Unsicherheit im Umgang mit der Neuerung zum Zeitpunkt der Umsetzung zurechtzukommen, und eventuelle Rückschläge bei missglückter Neuerung oder Umsetzung in Kauf nehmen. Während diese Innovativen, seien es Schulen oder einzelne Lehrkräfte innerhalb von Schulen, meist eher eine Außenseiterrolle einnehmen, sind sie doch zentral für den

Implementationsprozess, indem sie Neuerungen in der (über-)regionalen Bildungslandschaft oder ihrer Schule bekannt machen. Sie übernehmen daher eine Schlüsselrolle (*gatekeeper*) im Prozess (ebd.). Nach den Innovativen beginnen die Vorreiter*innen mit der Implementation der Neuerung. Sie machen einen Anteil von 13,5 % der Anwender*innen aus und werden von Rogers als respektierte Personen oder Einrichtungen beschrieben. Im Vergleich zu den Innovativen sind die Vorreiter*innen besser in das Sozialsystem der Akteur*innen (hier: der regionalen Schullandschaft oder Einzelschule) integriert und gelten innerhalb dieses Systems als Meinungsführende. Sie übernehmen die Rolle der *change agents* (Fullan, 1993) und haben somit die Aufgabe, die kritische Masse von der Neuerung und ihrer Implementation zu überzeugen. Darüber hinaus haben sie die Aufgabe, die Neuerung im Sinne des Sozialsystems ggf. zu adaptieren und somit auch die Unsicherheiten über die Umsetzung zu reduzieren. Im

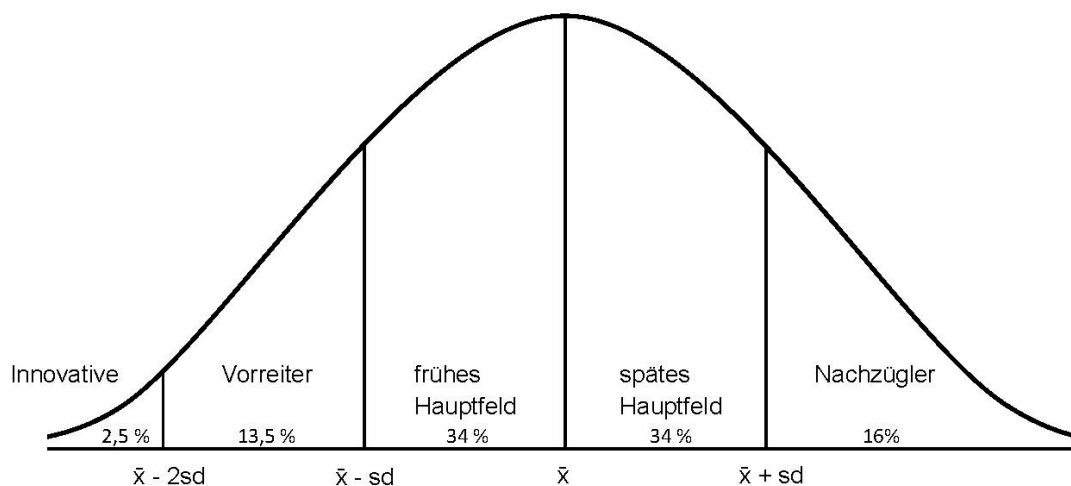


Abbildung 1: Anwenderkategorien entsprechend ihrer Innovationsfreude (Rogers, 2003, S. 281)

Anschluss an die Vorreiter*innen beginnt das frühe Hauptfeld mit der Umsetzung der Reform. Mitglieder dieser Gruppe stehen in regelmäßigem Austausch mit dem Sozialsystem, zählen aber selten zu den Meinungsführenden. Da sie die Reform den Vorreiter*innen folgend umsetzen, stellen sie eine wichtige Verbindung zu den übrigen Akteur*innen des Sozialsystems dar. Nach dem frühen folgt das späte Hauptfeld. Die Reform wird hier häufig skeptisch und mit Vorsicht betrachtet, sodass die Umsetzung erst erfolgt, wenn der ökonomische oder soziale Druck groß ist bzw. sich die Normvorstellungen des Sozialsystems im Sinne der Neuerung geändert haben. Die letzte Anwendungskategorie nennt Rogers die Nachzügler*innen. Die Mitglieder dieser Gruppe sind innerhalb des Sozialsystems nicht sehr gut vernetzt und können auch als Traditionalist*innen bezeichnet werden, die sich vor allem an Vergangem orientieren. Eine Neuerung wird hier meist erst dann implementiert, wenn eine Nicht-Implementation erhebliche Nachteile mit sich bringen würde (ebd.).

Die beschriebenen Systematisierungen von Coburn (2003) und Rogers (2003) finden in der (Bildungs-) Forschung bei der Untersuchung des Erfolgs von Implementationsprozessen Anwendung (Goldenbaum, 2012). Vor allem die Erfolgsindikatoren von Coburn (2003) werden aufgrund ihrer Fokussierung auf die Qualität von Implementation als bedeutsam angesehen. So untersuchen Trempler, Schellenbach-Zell und Gräsel (2012) den Transfer der UN-Dekade „Bildung für eine nachhaltige Entwicklung (BNE)“ mithilfe eines Fragebogens, der, basierend auf Coburns Indikatoren, wesentliche Veränderungen bzw. Aktivitäten auf Schulebene erfasst. Auch Goldenbaum, Schmittat, Stockmann und Kuhn (2012) greifen für die Untersuchung eines Transferprojekts auf die vier Indikatoren zurück. Die in diesem Zusammenhang von Schulentwicklungsberater*innen herausgestellten Gelingensbedingungen für den Transfererfolg lassen sich nach Ansicht der Wissenschaftler*innen gut in das „Coburn-Modell einpassen und stellen eine projektbezogene Konkretisierung des Modells dar“ (S. 47).

Die Anwendungskategorien nach Rogers (2003) finden ebenfalls in vielen Studien zur Einschätzung des Implementations- oder Transfererfolgs als theoretische Grundlage Anwendung. Vor allem für die Beschreibung der Verbreitung einer Reform scheinen sie geeignet (Gräsel, Jäger & Willke, 2006; Jäger, 2004) und wurden praktisch bisher vor allem für Studien zur Implementation von technologischen Neuerungen und E-Learning im Bildungsbereich eingesetzt (Chan & Ngai, 2012; Gallaher & Wentling, 2004; Yildirim, Zayim & Saka, 2006). Die Anwendungskategorien eignen sich nicht für die Erklärung von Implementationsprozessen (Rekkor, Ümarik & Loogma, 2013), stellen jedoch eine gute Möglichkeit dar, sie zu beschreiben. Weiterhin dienen die Kategorien als Grundlage, um Unterschiede in Implementationsprozessen bzw. ihrem Erfolg zu beschreiben (Chan & Ngai, 2012; Yildirim et al., 2006). Individuen werden demnach zunächst also kategorisiert und diese Kategorien anschließend charakterisiert. Die Kategorien als solche, wie sie von Rogers beschrieben wurden, konnten in Studien in (sehr) ähnlicher Form repliziert werden (Loogma, Kruusvall & Ümarik, 2012; Singhal & Dearing, 2006). Dabei wurde jedoch auch eine Implementationslücke (engl. *implementation gap*, vgl. Supovitz & Weinbaum, 2008) zwischen dem frühen und dem späten Hauptfeld sichtbar, die darauf hinweist, dass das Erreichen des späten Hauptfelds in Implementationsprozessen eine weitere Herausforderung darstellt (Loogma et al., 2012; Singhal & Dearing, 2006).

1.5 Fazit

Im Rahmen dieser Arbeit werden Reformen im Anschluss an Bormann (2011) als „formale Absichten zur Veränderung respektive Verbesserung“ verstanden, „die vom politisch-administrativen System ausgehen und durch Handlungen konkretisiert werden müssen“ (S. 13). Diese Konkretisierung von Reformen, also deren Übertragung in die Praxis, wird als Implementation bezeichnet (vgl. Goldenbaum, 2013). Für die Beschreibung des Implementationsprozesses hat sich in der Bildungsforschung weitestgehend ein Vier-Phasen-Modell durchgesetzt, das die Phasen Initiation, Umsetzung, Evaluation und Institutionalisierung unterscheidet (Goldenbaum, 2012, Hunneshagen, 2005). Die vier Phasen stehen dabei für vorbereitende, durchführende, auswertende und verankernde Tätigkeiten einer Implementation und laufen nicht zwingend linear ab bzw. können sich überschneiden (Goldenbaum, 2012). Für die Beschreibung der Umsetzung von Reformen wird zwischen ungeplanter Verbreitung auf der Makroebene (Diffusion) und gesteuerter Verbreitung auf Makro- (Dissemination) und Meso-/Mikroebene (Transfer) unterschieden (Jäger, 2004). Zudem werden die einzelnen Phasen des Implementationsprozesses geprägt von der Strategie, mit welcher dieser initiiert und begleitet wird. Dabei sind Top-down-, symbiotische, Bottom-up- und outputorientierte Strategien zu unterscheiden (Gräsel, 2014; Hasselhorn et al., 2014). Schulreformen, die auf der Makroebene durchgeführt werden und alle Schulen (zumindest einer Schulart) betreffen, werden in Deutschland meist top-down initiiert (z. B. die Schulstrukturreform in Berlin). In den vergangenen zwanzig Jahren kamen zudem verstärkt outputorientierte Strategien, ebenfalls häufig auf der Makroebene (z. B. die Implementation der Bildungsstandards), zum Einsatz (Gräsel, 2014).

Die Bewertung des Erfolgs eines Implementationsprozesses muss zuvorderst entlang der zu Beginn definierten Ziele der Reform erfolgen. Darüber hinaus werden in der Literatur verschiedene Gelingensbedingungen zusammengefasst, die sich jeweils der Ebene der Innovation oder Reform, der Lehrkräfte, der Einzelschule und des Schulsystems zuordnen lassen (Goldenbaum, 2012, 2013). Für die Neuerung selbst stellen dabei der wahrgenommene Wert und Nutzen (Hasselhorn et al., 2014) sowie der (empfundene) Neuheitsgehalt (Gillwald, 2000; Rogers, 2003) relevante Merkmale dar. Für die Ebene der Lehrkräfte steht die Akzeptanz einer Reform im Zentrum (Gräsel & Parchmann, 2004; Hall & Hord, 2015; Odden, 1991). Zudem werden die (wahrgenommene) inhaltliche Relevanz (Trempler et al., 2013), die kontinuierliche Professionalisierung (Odden, 1991) sowie die Kooperation der Lehrkräfte untereinander (Goldenbaum, 2013) als förderliche Faktoren identifiziert. Auf Einzelschulebene

wird der Schulleitung eine zentrale Rolle im Implementationsprozess zugeschrieben (Gräsel & Parchmann, 2004), aber auch die Schulkultur mit ihren spezifischen Werten und Normen entscheidet darüber, ob eine Reformimplementation gelingen kann (Holtappels, 1995). Für das Schulsystem insgesamt wird als förderlich beschrieben, wenn Implementationsprozesse systemisch begleitet und unterstützt werden sowie die einzelnen Entwicklungen innerhalb des Prozesses parallel laufen und ineinandergreifen (Hunneshagen, 2005). Darüber hinaus müssen in Schulreformprozessen besondere Bedingungen, wie der Bildungsauftrag, der pädagogische Zusammenhang (ebd.) oder die Arbeitsteilung in gefügearartigen Kooperationen (vgl. u. a. Weick, 1976), berücksichtigt werden.

Für die Erfassung des Implementationserfolgs bieten sich im Rahmen dieser Arbeit die Arbeiten von Coburn (2003) und Rogers (2003) an. Coburn unterscheidet in ihren Überlegungen zur Verbreitung von Schulreformen zwischen (1) der Tiefe, also tiefgreifenden und konsequenten Veränderungen des Unterrichtsgeschehens im Sinne der Reform; (2) der Nachhaltigkeit, im Sinne einer langfristigen Umsetzung und Institutionalisierung der mit der Reform vermittelten Inhalte; (3) der Verbreitung, die neben der quantitativen Verbreitung beteiligter Akteur*innen innerhalb und außerhalb einer Schule vor allem die Übertragung der vermittelten Werte und Normen oder auch pädagogischen Absichten auf andere Bereiche ihrer professionellen Praxis beschreibt; sowie (4) die Identifikation mit der Reform dergestalt, dass aus einer extern verstandenen und unterstützten eine intern verstandene und unterstützte Reform wird. Für die Beschreibung der Verbreitung eines Implementationsprozesses in einem eher quantitativen Sinne beschreibt Rogers (2003) fünf idealtypische Anwendungskategorien: (1) Die Innovativen machen dabei eine Neuerung in der (regionalen) Schullandschaft bekannt und übernehmen eine Schlüsselrolle in der Verbreitung (*gatekeeper*), (2) die Vorreiter*innen greifen die Erfahrungen der Innovativen auf und übernehmen als Meinungsführende die Rolle der *change agents* mit dem Ziel, die kritische Masse von der Reform zu überzeugen. (3) Anschließend beginnt das frühe Hauptfeld mit der Reformumsetzung. Es bildet ein wichtiges Bindeglied zwischen den Vorreiter*innen und den Akteur*innen, die die Reform noch nicht umsetzen. (4) Dem frühen folgt idealerweise das späte Hauptfeld. Diese Gruppe ist eher skeptisch und wendet Reformen und Neuerungen erst an, wenn der ökonomische oder soziale Druck groß ist. (5) Als letzte Anwendungskategorie beschreibt Rogers die Nachzügler*innen, die eine Reform erst umsetzen, wenn eine Nicht-Umsetzung erhebliche Nachteile mit sich bringen würde.

2. Schulbeginn in Berlin und Deutschland

Die Neukonzeption des Schulanfangs war eine der wesentlichen Reformanstrengungen im deutschen Bildungswesen der letzten Jahrzehnte (Berthold, 2008b; Carle, 2008; Liebers, Prenzel & Bieber, 2008). Zum einen von einem Umdenken in den Erwartungen an Kinder und Schule am Schulanfang geprägt und zum anderen durch die Reaktionen auf die erwartungswidrig mittelmäßig bis schlechten Ergebnisse deutscher Schüler*innen in internationalen Vergleichsstudien (u.a. Mullis et al., 2004; OECD, 2002) ausgelöst, wurde die Reform des Schulanfangs in den Schulgesetzen der Bundesländer ab Mitte der 2000er Jahre verankert.

Das folgende Kapitel stellt diese Entwicklungen für Deutschland im Allgemeinen und Berlin im Speziellen dar. Dabei werden zunächst der Wandel bezüglich der Erwartungen an Schulanfänger*innen und die entsprechenden Begrifflichkeiten beschrieben (Abschnitt 2.1). Anschließend werden Entwicklung und Merkmale der heutigen Schulanfangsphase (SAPH) für Deutschland (Abschnitt 2.2) und Berlin (Abschnitt 2.3) im Detail dargestellt. Darauf aufbauend, werden die beiden umfangreichsten Reformen an Grundschulen dieser Zeit, die auch den Schwerpunkt dieser Dissertation darstellen, in den Blick genommen: das Jahrgangsübergreifende Lernen (Abschnitt 2.4) und die Senkung des Schuleintrittsalters (Abschnitt 2.5). Abschließend werden die wesentlichen Erkenntnisse des Kapitels zusammengefasst (Abschnitt 2.6)

2.1 Schulbeginn in Deutschland

Die Debatte um die Gestaltung des Schulanfangs ist sowohl in Deutschland als auch in anderen Ländern der OECD geprägt von Bemühungen zur Definition der Erwartungen an die beteiligten Akteur*innen und Institutionen. Konkret stehen hier die Kompetenzen und Fähigkeiten der Schulanfänger*innen einerseits und die Anforderungen an Schule und Unterricht andererseits im Mittelpunkt. Begrifflich spiegelt sich die Entwicklung dieser Debatte in den Bezeichnungen Schulreife, Schulfähigkeit und Schulbereitschaft wider (Carle & Metzen, 2008a; Hasselhorn & Lohaus, 2008; Stamm, 2013; Tervooren, 2010).

Kern (1951) ging zunächst davon aus, dass mit einem bestimmten Alter eines Kindes auch ein bestimmter Entwicklungsstand einhergeht, der dieses Kind „reif“ für die Schule macht. In

seinem grundlegenden Werk „Sitzenbleiberelend und Schulreife“ fasst er seine Sicht wie folgt zusammen:

Jedes Kind, extrem schwache Begabung (Idiotie und Imbezillität) ausgenommen, erreicht im Laufe seiner Entwicklung einmal die Entwicklungsphase, der jenes Leistungsgefüge zugeordnet ist, das als Voraussetzung für ein erfolgreiches Durchlaufen der Schule angesetzt werden muss. Das eine Kind kommt lediglich früher, das andere später zu diesem Entwicklungspunkt. Insofern ist die schulische Leistung primär von der Entwicklung abhängig. (ebd. S. 67)

Nach dieser Annahme stellt das Alter den besten Indikator für die Feststellung der Schulreife dar. Mithilfe von Schulreifetests (Kern, 1970a) sollte darüber hinaus festgestellt werden, ob ein Kind den notwendigen Entwicklungsstand erreicht hat. Ist dies zum Zeitpunkt der Einschulung nicht der Fall, so ist das Kind, nach Kern, von der Einschulung zurückzustellen, da es sonst die Anforderungen des schulischen Anfangsunterrichts nicht erfolgreich bewältigen kann (Hasselhorn & Lohaus, 2008; Kern, 1970b). Eine in seinem Sinne verfrühte Einschulung sah Kern als ursächlich für die hohen Sitzenbleiberquoten jener Zeit an. Und auch eine Beschleunigung der Entwicklung eines Kindes wurde in der sogenannten Reifungstheorie ausgeschlossen (Ehm & Hasselhorn, 2017). Obwohl bereits damals Kritik an der Unzuverlässigkeit der Schulreifetests zu vernehmen war (Hasselhorn & Lohaus, 2008; Kammermeyer, 2000), galt das Konzept der Schulreife zu dieser Zeit als anerkannt und war weit verbreitet. In der heutigen Diskussion scheint es einerseits als leitendes Konzept in der Einschulungsdiskussion überwunden zu sein. Andererseits stellt das Alter nach wie vor das zentrale Schuleintrittskriterium dar (Faust, 2006; Kerstan & Spiewak, 2008; Stamm, 2013).

Bereits zu Beginn der 1960er Jahre wurde die Annahme, dass die Entwicklung eines Kindes nicht beeinflusst werden kann, von Kemmler und Heckhausen (1962) widerlegt. Schenk-Danzinger (1969) plädierte für die Verwendung des Begriffs der Schulfähigkeit (Hasselhorn & Lohaus, 2008; Tervooren, 2010) sowie für die Erweiterung des Aspekts der Schulbereitschaft, also die Fokussierung auf motivationale und sozial-emotionale Faktoren (Schenk-Danzinger, 1969). Zudem lieferten die begleitenden Untersuchungen zu den Förderprogrammen der 1970er in den USA (z. B. Head Start) ebenfalls Belege einer „prinzipiellen Trainierbarkeit der kognitiven Entwicklung“ (Hasselhorn & Lohaus, 2008, S. 412), sodass sich nach dem reifungstheoretischen das lerntheoretische Paradigma und somit der Begriff der Schulfähigkeit verbreitete. Dabei grenzte sich die Schulfähigkeit zunächst von der Schulreife dahingehend ab, dass zwar immer noch die Eigenschaften des Kindes im Vordergrund standen, diese aber in

Bezug zu den Anforderungen der jeweiligen Grundschule gesetzt wurden (Kammermeyer, 2000; Nickel & Schmidt, 1975). Nickel (1990) erweiterte anschließend das Konzept der Schulfähigkeit als abhängig vom individuellen Entwicklungs- und Förderstand, ökosystemischer Umweltfaktoren (Merkmal der schulischen, vorschulischen und häuslichen Lernumwelt) sowie von der Schule mit ihren Anforderung und Lernbedingungen (Abbildung 2).

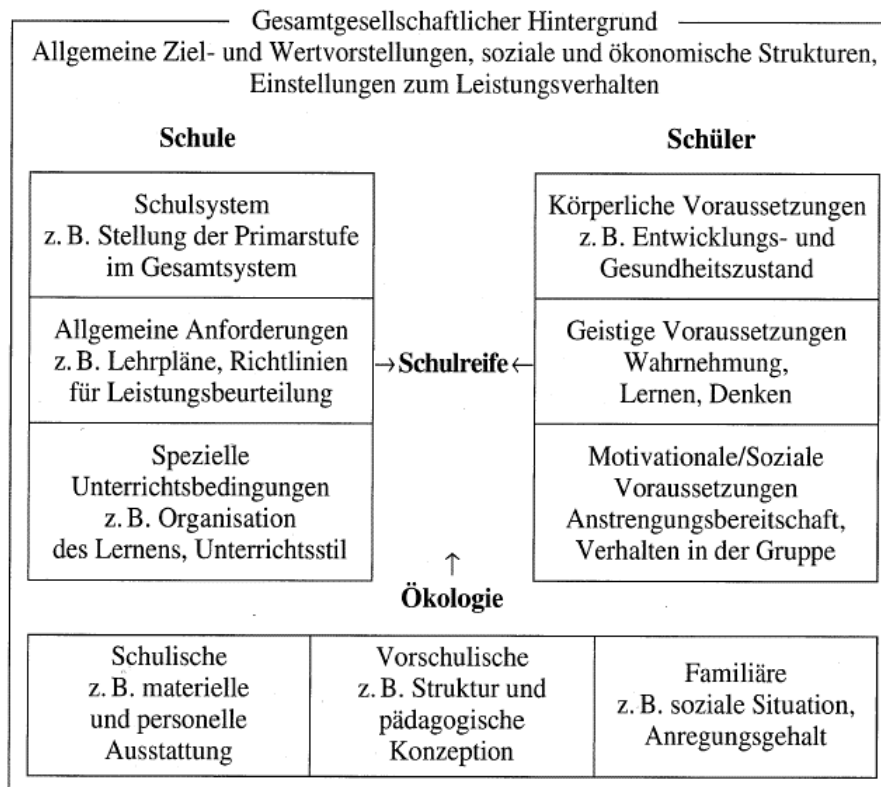


Abbildung 2: Ökosystemisches Modell der Schulfähigkeit (Nickel & Schmidt-Denter, 1995, S. 227)

Mit Berücksichtigung der Perspektive der Lernumwelt, wie es im ökosystemischen Modell von Nickel der Fall ist, hat sich international der Begriff der „Schulbereitschaft“, im Sinne einer Bereitschaft für die Schule und der Bereitschaft der Schule selbst, verbreitet (Ehm & Hasselhorn, 2017). Im deutschsprachigen Raum wird weiterhin häufig von Schulfähigkeit gesprochen und damit für ein Bildungsverständnis geworben, das „sich im Hinblick auf die Schulfähigkeit nicht ausschließlich auf konstruktivistische Elemente beschränkt“, wie sie der „Schulbereitschaft“ zugrunde liegen, „sondern die instruktiv-schulvorbereitende Perspektive einbezieht und damit der pädagogischen Fachperson eine neue ‚befähigende‘ Rolle überantwortet“ (Stamm, 2013, S. 278). Die KMK folgt dieser Ansicht, indem sie Schulfähigkeit als „Schnittpunkt der Lernvoraussetzungen des Kindes, des sachlichen Anspruchs der Inhalte und des pädagogischen Konzepts der Schule“ (KMK, 1997, S. 1) beschreibt.

Was sich hinter dem Begriff der Schulfähigkeit aus Sicht der Eltern, Lehrkräfte und Forschung genau verbirgt, wurde u. a. in der BiKS-Studie ("Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter"; Faust, 2013) untersucht. Pohlmann-Rother und Kolleg*innen (2013) zeigten dabei, dass sowohl für Eltern als auch für Lehrkräfte und Erzieher*innen Ausdauer und Konzentrationsfähigkeit der Kinder die wichtigsten Indikatoren für Schulfähigkeit darstellen. Von den Eltern wurden anschließend geistiger Entwicklungsstand, Selbstständigkeit und soziale Fähigkeiten genannt. Bei den Lehrkräften folgten Interesse an Schule und/oder am Lernen, ruhig sitzen und zuhören können sowie der geistige Entwicklungsstand und für die Erzieher*innen stellten Selbstständigkeit, geistiger Entwicklungsstand und soziale Fähigkeiten die anschließend wichtigsten Kompetenzen dar (ebd.). Das Alter erwies sich aus Sicht aller drei Akteursgruppen als eines der unwichtigeren Kriterien, wurde aber von Eltern und Lehrkräften immerhin noch von 13 % bzw. 11 % der Befragten genannt. Bei den Erzieher*innen verwiesen lediglich 2 % der Befragten auf das Alter der Kinder. Weiterhin, als Hauptargumente für eine vorzeitige Einschulung werden vonseiten der Eltern die Vermeidung von Langeweile und Unterforderung im Kindergarten (Faust, Klucznik & Pohlmann-Rother, 2013) und für eine verzögerte Einschulung das Alter, mangelndes Interesse des Kindes an schulischen Inhalten und Verspieltheit (Wehner, 2013) angegeben.

Aus Sicht der Forschung zu (vor-) schulischen Fähigkeiten und Kompetenzen und ihren Auswirkungen auf den späteren Schulerfolg zeigt sich, dass es vor allem bereichsspezifische Vorläuferfähigkeiten, wie die phonologische Informationsverarbeitung und räumlich-konstruktive und pränumerische Fertigkeiten (Hasselhorn & Lohaus, 2008) sind, die vergleichsweise hohe Anteile schulischer Leistungen erklären können (Duncan et al., 2007). Aber auch sozial-emotionale Fähigkeiten sind prädiktiv für spätere Einstellungen zur Schule und die Anpassung an Schulerfordernisse (Hasselhorn & Lohaus, 2008). Als gut bestätigt gilt die Annahme, dass die Fähigkeit zur Kontrolle des eigenen Sozialverhaltens als Prädiktor für die Integration in eine Lerngruppe sowie die Anerkennung bei den Mitschüler*innen für das Engagement im Unterricht wirkt (Faust, 2013). Hasselhorn, von Goldammer und Weber (2008) zeigen den Zusammenhang zwischen der Fähigkeit des Belohnungsaufschubs mit vier bis fünf Jahren und dem Sozialverhalten und Schulerfolg im Jugendalter. Weiterhin ist der positive Einfluss des Besuchs einer vorschulischen Einrichtung, wie dem Kindergarten, aber auch der häuslichen Lernumgebung auf die kindlichen Kompetenzen sowohl international (OECD, 2011) als auch national gut dokumentiert (Schober & Spieß, 2012).

2.2 Historische und aktuelle Entwicklungen der Schulanfangsphase in Deutschland

Nach den großen Reformen in den 1960er und 1970er Jahren (Cortina, Baumert, Leschinsky, Mayer & Trommer, 2005; Gudjons, 1999) geriet zu Beginn der 1990er Jahre das deutsche Bildungssystem erneut in die Kritik (Ramseger, 2004). Vor allem die vergleichsweise lange Verweildauer deutscher Schüler*innen und der späte Eintritt in das Berufsleben (OECD, 2001a) wurden hierbei hinterfragt (Berthold, 2008a; Ramseger, 2004). Ein Teil dieser Probleme zeigte sich bereits in den Grundschulen: die Rückstellungsquoten bei der Einschulung waren seit den 1990er Jahren kontinuierlich gestiegen, was das ohnehin schon im Vergleich zu anderen europäischen Ländern höher angesetzte Einschulungsalter weiter erhöhte. Die tägliche Lern- und Unterrichtszeit deutscher Grundschüler*innen war gering⁶ und die Kinder verbrachten relativ wenig Zeit ihres Tages in der Schule. Auch waren die Schulen, was Personal und Material betraf, vergleichsweise schlecht ausgestattet (Ramseger, 2004). Als Antwort auf diese Mängel einerseits, aber auch als Folge der mangelnden Aussagekraft und Objektivität der weit verbreiteten Schulreifetests (Hasselhorn & Lohaus, 2008; Kammermeyer, 2000) und der Erkenntnis, dass Kinder mit sehr heterogenen Lernausgangslagen, wie sie in den sich verbreitenden Integrationsklassen häufig gegeben sind, erfolgreich gemeinsam lernen können (Berthold, 2008a, 2008b; Kastirke & Jennessen, 2006) wurde eine Reform der Schulanfangsphase⁷ diskutiert. Seit 1992 wurde die veränderte SAPH teils in Schulversuchen, teils als offiziell eingeführtes Einschulungsmodell⁸ oder auch als nicht überall umgesetzte Regeleinrichtung⁹ oder Projekt der Schulentwicklung¹⁰ eingeführt (Bos, Hornberg & Arnold, 2008; Bos, Schwippert & Arnold, 2010). Den Empfehlungen zum Schulanfang der KMK (1997) folgend, hatte die neue SAPH in allen Bundesländern eine Reduktion der Rückstellungsquoten, die Ermutigung der Eltern zur vorzeitigen Einschulung von sogenannten „Kann-Kindern“¹¹ und eine mögliche Flexibilisierung der Stichtage zur Einschulung bzw. die Einführung mehrerer Stichtage zum Ziel. Zudem beschlossen sechs der 16 Bundesländer¹² eine Verschiebung des Stichtags zur Einschulung und somit eine weitere Senkung des

⁶ Hier wurde vor allem die in den alten Bundesländern und West-Berlins übliche „Vierteltagsgrundschule“ den Bedürfnissen der Eltern und Kinder nicht mehr gerecht (Ramseger, et al., 2004)

⁷ In der Literatur verwendete Synonyme: Schuleingangsstufe bzw. -phase

⁸ Seit 2004/05 in Sachsen, seit 2005/06 in Berlin, Nordrhein-Westfalen, Sachsen-Anhalt

⁹ Gesetzlich verpflichtend in Brandenburg, Bremen, Niedersachsen, Rheinland-Pfalz

¹⁰ Thüringen, Schleswig-Holstein

¹¹ „Kann-Kinder“ oder seitens der Bildungsadministration auch Antragskinder genannt, sind Kinder, die zum Stichtag das sechste Lebensjahr noch nicht vollendet haben, jedoch vom Stand ihrer Entwicklung her als schulfähig eingeschätzt werden.

¹² Baden-Württemberg, Bayern, Berlin, Brandenburg, Nordrhein-Westfalen, Rheinland-Pfalz

Schuleintrittsalters (für eine Übersicht: Berthold, 2008a). Zum Erreichen dieser Ziele sollte bundesweit bei der Einschulung auf die Selektion durch Feststellung der Schulfähigkeit verzichtet und alle Kinder, die ein festgelegtes Alter erreicht haben, eingeschult werden. Kinder, die ansonsten zurückgestellt worden wären, sollten durch multiprofessionelle Teams (Lehrkräfte, sozialpädagogisches und ggf. auch sonderpädagogisches Fachpersonal) integrativ gefördert werden. Bei jahrgangsübergreifender Führung der SAPH sollten Schüler*innen diese in einem bis drei Jahren durchlaufen können (flexible Verweildauer), um somit den großen Entwicklungsunterschieden der Kinder zu Beginn ihrer Schullaufbahn gerecht zu werden (Bos et al., 2008; Bos et al., 2010). Diese strukturellen Veränderungen sollten, pädagogisch begleitet, die Voraussetzung bilden, um der Heterogenität der Schülerschaft zu begegnen und, bei gleichzeitiger Anerkennung der Beschränktheit schulischen Einflusses, soziokulturelle Benachteiligung zu kompensieren (Berthold, 2008b). Die Reform des Schulanfangs war demnach geprägt durch die Annahme einer kindgerechten Schule anstelle des schulreifen oder -fähigen Kindes und folgte dem von Prenzel (2006) maßgeblich geprägten Ansatz einer Pädagogik der Vielfalt (ebd.). Kennzeichnend für diese Pädagogik sind die Annahme vielfältiger kindlicher Kompetenzen und die Betrachtung von Verschiedenheit als Bereicherung.

Mit Blick auf die Umsetzung der einzelnen Elemente der SAPH zeigt sich im Schuljahr 2016/17 in der BRD folgendes Bild¹³: sieben der 16 Bundesländer haben den Stichtag zur Einschulung auf den 30. Juni (HB, HE, MV, SH, SL, SN, ST) gelegt, sechs weitere auf den 30. September (BW, BY, BB, NI, NW) und Berlin auf dem 31. Dezember. Die übrigen drei Bundesländer (HH, RP, TH) haben ihre Stichtage zwischen diese beiden Daten gelegt. In allen sechzehn Bundesländern können Kinder von der Einschulung prinzipiell um ein Jahr zurückgestellt werden. Jedoch legen die Bundesländer dabei verschiedene (mehr oder weniger strenge) Voraussetzungen fest. Ebenfalls in allen Bundesländern möglich ist es, ein Kind bis zu ein Jahr früher einzuschulen, sofern die entsprechenden Voraussetzungen, i. d. R. durch einen Antrag der Eltern und ein schulärztliches Gutachten, erfüllt werden. Schulreife-tests werden in Deutschland auch weiterhin nicht mehr standardmäßig eingesetzt. In den meisten Bundesländern sind schulpflichtige Kinder und die, die vorzeitig eingeschult werden sollen, zur Teilnahme an einer schulärztlichen Untersuchung (Schuleingangsuntersuchung) verpflichtet. In Bayern kann die Schule anschließend eine Schulreifeuntersuchung anordnen. In einigen

¹³ Angaben wurden den länderspezifischen Schulgesetzen bzw. Verordnungen entnommen. Eine Übersicht zu den rechtlichen Grundlagen ist Anhang A.1 zu entnehmen

Bundesländern ist das Einverständnis der Schulleitungen zur Schuleingangsuntersuchung der bei ihnen angemeldeten Kinder nötig (MV, RP, SL, SH, TH). Zur Umsetzung des Jahrgangsübergreifenden Lernens (JÜL) und der damit einhergehenden Möglichkeit einer flexiblen Verweildauer in der Schulanfangsphase lässt sich festhalten, dass in zwölf Bundesländern die Möglichkeit besteht, die ersten beiden (z. T. auch die ersten drei) Jahrgangsstufen übergreifend zu organisieren. Berlin, Bremen und Sachsen-Anhalt haben dabei JÜL entweder als Regeleinrichtung oder als bevorzugte Organisationsform in ihren Schulgesetzen verankert. In sieben Bundesländern sind die ersten beiden Jahrgangsstufen als pädagogisch-curriculare Einheit im Gesetz festgehalten (BE, HE, NW, RP, SH, SN, TH) und in zwei weiteren werden sie zumindest gemeinsam als Schulanfangsphase o. Ä. bezeichnet (NI, SL). Zusammenfassend ist festzustellen, dass einzelne Aspekte der ursprünglichen Überlegungen zur Reform des Schulanfangs inzwischen Eingang in die Schulgesetzgebung der Länder der Bundesrepublik Deutschland gefunden haben. In einem Teil der Bundesländer werden Kinder in einem jüngeren Alter eingeschult und die Mehrzahl der Bundesländer ermöglicht die übergreifende Organisation der ersten beiden Jahrgangsstufen und somit die flexible Verweildauer. Andere Aspekte, wie das Einschulen aller Kinder ohne Ausnahme (keine Rückstellungen), wurden nicht nachhaltig in den einzelnen Bundesländern verankert. Jedoch ist es neun Bundesländern gelungen, ihre Rückstellungsquoten zwischen den Schuljahren 2000/01 und 2014/15 zu senken (HB, HH, MV, SH; SN, ST, NI, SW, RP). Die Hansestadt Bremen kann dabei sogar eine Reduktion um 10,7 % auf 1,5 % aller schulpflichtigen Kinder verzeichnen.

Wissenschaftliche Untersuchungen zur Schulanfangsphase liegen für die deutschen Bundesländer in Form von Begleitstudien und Evaluationen vor. Als prominente Beispiele sind hier die Evaluationen der Schulversuche in Baden-Württemberg (Stuchlik, 2002) und Brandenburg (Liebers et al., 2008) sowie die wissenschaftliche Begleitung zum Transfer der Schulanfangsphase in Thüringen (BeSTe; Carle & Metzen, 2009) zu nennen. Der Stand der Umsetzung der SAPH in Deutschland im Schuljahr 2005/06 wurde im Rahmen der Erweiterungsstudien zu IGLU (IGLU-E) von Bos und Kolleg*innen (2010) untersucht. Die Wissenschaftler*innen kamen damals zu dem Schluss, dass durch die veränderten Regelungen zur Einschulung (z. B. keine Rückstellungen) in der Schulanfangsphase¹⁴ die Klassen heterogener geworden sind. Weiterhin ist vor allem der Unterricht in jahrgangsübergreifenden Klassen anspruchsvoller und erfordert zusätzliche Ressourcen (ebd.). Es wird auch festgestellt,

¹⁴ In der Studie wurde der Begriff neue Schuleingangsstufe verwendet.

dass es in den meisten Bundesländern nach wie vor Rückstellungen gibt und auch nicht völlig auf die Feststellung der Schulfähigkeit verzichtet wird. Der überwiegende Anteil der Klassen wird jahrgangsbezogen organisiert und als elementar für die Einführung der SAPH definierte Elemente, wie die Kooperation von vorschulischen Einrichtungen, Hort und Schule oder die Binnendifferenzierung, sind nicht überall umgesetzt. Faust (2006) zieht auf Grundlage unterschiedliche Evaluationsergebnisse der einzelnen Länder, aber auch einzelner Schulen, den Schluss, „dass wohl auch bei der neuen Schuleingangsstufe die *Konzeption allein* nicht für bessere oder schlechtere Schülerleistungen verantwortlich zu machen ist, sondern dass bezogen auf die einzelne Klasse die *Schul- und Unterrichtsqualität* entscheidet“ (ebd., S. 338, Hervorhebung im Original, d. Verf.). Viele Studien beziehen sich daher häufig auf einzelnen Aspekte der SAPH (z. B. Berthold, 2008a; Carle & Metzen, 2008a; Liebers et al., 2008).

2.3 Die Schulanfangsphase in Berlin

In Berlin reagierte die ab 1995 für Schulen zuständige Senatorin, Ingrid Stahmer (SPD), auf die sich ändernden Anforderungen an den Schulanfang mit der sogenannten Grundschulreform 2000. Mit Beginn des Schuljahres 1998/99 sollte die Reform an Berliner Grundschulen implementiert werden. Das Programm beinhaltete, gemäß den KMK-Empfehlungen (KMK, 1997), sowohl die Flexibilisierung des Verfahrens zum Schuleintritt als auch die Individualisierung der SAPH sowie im Weiteren die Entwicklung einer verlässlichen Halbtagsgrundschule, die Förderung benachteiligter Kinder in sozialen Brennpunkten, Begegnung mit einer Fremdsprache ab Klasse 3, Profilausbildung und Differenzierung in den Klassen 5 und 6 sowie CidS – Computer in die (Grund-)Schule (SenSJS, 2004). Zunächst wurde die Grundschulreform 2000 durch die Senatorin als offen angelegt, was bedeutete, dass Schulen eigene Schwerpunkte setzen sowie Ziele und Wege der eigenen Schulreform selbst bestimmen konnten. Mit Beginn der neuen Legislaturperiode und der Übernahme der Senatsschulverwaltung durch Klaus Böger (SPD) im Herbst 1999, aber vor allem auch als Antwort auf die Ergebnisse der internationalen Leistungsstudien (Deutsches PISA-Konsortium, 2002; 2001b; OECD, 2002), wo sich weite Teile des deutschen, aber auch speziell des Berliner Schulsystems als reformbedürftig darstellten, wurde eine weitreichende Überarbeitung des Berliner Schulgesetzes (gültig ab 2004) und damit aller Bildungsgänge angestrebt (Anhang A.2). Eine Änderung des Schulgesetzes hatte, im Gegensatz zur ursprünglich offen angelegten Grundschulreform 2000, einen verpflichtenden Charakter zur Folge. Die Überarbeitung beinhaltete im Sekundarschulbereich u. a. die Umstellung von der Viergliedrigkeit (Gesamt-,

Haupt-, Realschule und Gymnasium) zur Zweigliedrigkeit (Integrierte Sekundarschule und Gymnasium), wobei beide Schulformen gleichwertig nebeneinanderstehen sollen¹⁵. Weiterhin wurde das Abitur an Gymnasien um ein Jahr auf zwölf Schuljahre verkürzt. Für die Grundschule wurden große Teile der Grundschulreform 2000 aufgegriffen, erweitert und konkretisiert. Die Reform der SAPH zum Schuljahr 2005/06 als Neuausrichtung der Einschulung und der ersten Schuljahre nahm dabei eine zentrale Rolle ein: die im Rahmen eines Schulversuchs teilweise eingeführte verlässliche Halbtagsgrundschule (Ramseger, 2004) sollte bis zum Schuljahr 2006/07 flächendeckend an Berliner Grundschulen implementiert werden und auch die in der Grundschulreform 2000 bereits formulierte Förderung benachteiligter Kinder in sozialen Brennpunkten, der Kontakt mit einer Fremdsprache ab Klasse 3 sowie die Profilbildung in den Klassen 5 und 6 fanden Eingang in das neue Schulgesetz (Abgeordnetenhaus Berlin, 2004, Abgeordnetenhaus Berlin, 2008). Wie auch bei der Grundschulreform 2000 erhielt die Individualisierung des Lehren und Lernens in der neu ausgerichteten Schulanfangsphase in den Konzeptionen zum neuen Schulgesetz besondere Aufmerksamkeit (Abgeordnetenhaus Berlin, 2000, Abgeordnetenhaus Berlin, 2004; SenSJS, 2004). Allgemeine Ziele waren dabei, die Qualität von Unterricht zu verbessern und Kinder in der Schule sowie in ihrer Motivation und Lernfreude optimal zu fördern (Abgeordnetenhaus Berlin, 2004). Gelingen sollte dies durch die Etablierung einer neuen Lernkultur, die mit den Schwerpunkten besetzt wurde, die Interessen der Kinder aufzugreifen, Lernwege zu individualisieren, selbsttätige Lernprozesse zu initiieren, Lernen in Zusammenhängen und an bedeutsamen Inhalten sowie Lernen voneinander und miteinander zu ermöglichen, Lernkompetenz zu entwickeln, mit Fehlern konstruktiv umzugehen, Elternarbeit zu intensivieren und schließlich Räume in der Schule zu schaffen (ebd.). Kennzeichen des neuen Schuleingangs in Berlin waren zunächst, wie in den übrigen Bundesländern auch, die Senkung des Schuleintrittsalters durch die Verschiebung des Einschulungstichtags, das Unterbinden von Einschulungsrückstellungen und die Motivation zur vorzeitigen Einschulung. Ebenfalls mit dem Ziel der Individualisierung der Lehr-Lern-Prozesse und der somit optimalen Förderungen der Schüler*innen, wurde in der Reform des Schulgesetzes das Jahrgangsübergreifende Lernen als verpflichtender Bestandteil der Schulanfangsphase aufgenommen. Mithilfe der Umsetzung von JÜL sollte zudem ein weiterer wesentlicher Bestandteil der Schulanfangsphase gesichert werden: das flexible Verweilen. Flexibles Verweilen ermöglicht den Schüler*innen „ein langsames oder schnelleres Durchlaufen [der Schulanfangsphase] ohne [größere] Brüche und

15 Z. B. bieten beide Schulformen alle allgemeinbildenden Schulabschlüsse (inkl. Abitur) an und es gelten die gleichen Bildungsstandards (für eine Übersicht siehe Baumert, Maaz, Neumann, Becker und Dumont (2013)).

Wechsel der sozialen Gruppen“ (Abgeordnetenhaus Berlin, 2004). Sowohl in Berlin als auch in sechs weiteren Bundesländern können Schüler*innen die SAPH in ein bis drei Jahren durchlaufen und anschließend in die dritte Jahrgangsstufe übertreten, ohne dass sich dies auf die Anzahl ihrer Gesamtschuljahre auswirkt. Dadurch sollte in besonderem Maße der Heterogenität der Schülerschaft Rechnung getragen und sichergestellt werden, dass jedes Kind am Ende der Schulanfangsphase über grundlegende Kompetenzen verfügt. Den KMK-Empfehlungen zum Schulanfang (1997) folgend, war die Überarbeitung der Einschulung selbst wesentlicher Bestandteil der Neukonzeption der Schulanfangsphase sowohl in der Grundschulreform 2000 als auch in der Reform des Berliner Schulgesetzes. Neben der Senkung des Einschulungsalters von durchschnittlich sechs Jahren und sieben Monate auf sechs Jahre und einen Monat sollte zukünftig auf die Schulfähigkeitsprüfung verzichtet und alle Kinder, die ihrem Alter nach schulpflichtig waren, eingeschult werden. Vorklassen wurden abgeschafft und von der Feststellung der Förderschwerpunkte „Lernen“ und „emotionale und soziale Entwicklung“ in Jahrgangsstufe 1 abgesehen (Abgeordnetenhaus Berlin, 2004). Weiterhin sollten Kinder mit mangelnden Deutschkenntnissen nicht wie bisher separat in Kleinklassen, sondern gemeinsam mit allen Kindern beschult und sprachlich gefördert werden. Die ohnehin große Heterogenität der Berliner Schülerschaft beim Schuleintritt (Abgeordnetenhaus Berlin, 2004; SenBJW, 2013b) wurde infolge dieser Änderung weiter erhöht und die Notwendigkeit individualisierter Lehr-Lern-Prozesse verdeutlicht.

Zu den bisher genannten Veränderungen stand die Schulgesetzreform in 2004 auch im Zeichen des Bildungsmonitorings (SenBWF, 2004). Mit dem Ziel der ganzheitlichen Analyse und Steuerung des Bildungswesens sollten die in den Jahren 2003 und 2004 entwickelten, bundesweit einheitlichen Bildungsstandards¹⁶ für die Fächer Deutsch, Fremdsprachen und Mathematik mithilfe verschiedener standardisierter Leistungsfeststellungsverfahren einerseits in die Schule gebracht, andererseits ihr Erreichen bei den Schüler*innen überprüft werden. So sollten, nach Ansicht der Senatsbildungsverwaltung, in der Stärkung der Eigenverantwortung der Schule diese Verfahren dazu beitragen, eine einheitliche Schulqualität zu sichern. Mithilfe der Orientierungsarbeiten in Jahrgangsstufe 2 (in den Schuljahren 2003/04 bis 2006/07) und der Vergleichsarbeiten in Jahrgangsstufe 4 (in den Schuljahren 2004/05 bis 2006/07) bzw. den Vergleichsarbeiten in Jahrgangsstufe 3 (ab dem Schuljahr 2007/08) sollten die Kompetenzen

¹⁶ Bildungsstandards geben Ziele der pädagogischen Arbeit in Form von anzustrebenden Lernergebnissen vor. Sie legen fächerspezifisch fest, welche Kompetenzen Kinder und Jugendliche bis zu einem bestimmten Abschnitt in ihrer Schullaufbahn entwickeln sollen. Unter einer Kompetenz wird dabei die Fähigkeit verstanden, Wissen und Können in den jeweiligen Fächern zur Lösung von Problemen anzuwenden. (IQB (2017a))

der Schüler*innen erfasst und der Erfolg des Unterrichts dokumentiert werden (Abgeordnetenhaus Berlin, 2007b; SenBJS). Die Vergleichsarbeiten dienten zusätzlich als „Unterstützung der Unterrichts- und Schulentwicklung jeder einzelnen Schule durch eine an den Bildungsstandards orientierte Rückmeldung als Standortbestimmung mit Bezug zu den Landesergebnissen“ (KMK, 2015).

Die Reform des Berliner Schulgesetzes im Allgemeinen, vor allem aber die Neuausrichtung des Schulanfangs im Speziellen wurde von den lokalen, zum Teil aber auch überregionalen Medien und der breiten Öffentlichkeit mit großem Interesse verfolgt (für eine Übersicht siehe Anhang A.3). Auch die politischen Parteien im Berliner Senat, allen voran die Fraktionen der damaligen Opposition (Bündnis 90/Die Grünen, CDU, FDP), und die zuständigen Gewerkschaften beobachteten einzelne Schritte der Reform und setzten sich (kritisch) mit den Neuerungen auseinander (GEW Berlin, 2006; 2008; Klesmann, 2013). Hauptkritikpunkte seitens der Politik waren dabei die zu kurzfristige, die Schulen überfordernde Umsetzung (Abgeordnetenhaus Berlin, 2005b, Abgeordnetenhaus Berlin, 2006c, Abgeordnetenhaus Berlin, 2006d) und fehlende wissenschaftliche Begleitung (Abgeordnetenhaus Berlin, 2005a, Abgeordnetenhaus Berlin, 2007a, Abgeordnetenhaus Berlin, 2007d) dieser umfassenden Reform und zu wenig Unterstützung für die Schulen und Akteure (Abgeordnetenhaus Berlin, 2006a, Abgeordnetenhaus Berlin, 2007e). Zahlreiche kleine Anfragen der Abgeordneten zeigen weiterhin, dass sie sich vor allem um die (ausreichende) Qualifizierung der Beteiligten in den Schulen sorgten (ebd.).

Zur aktuellen Lage der SAPH lässt sich festhalten, dass einige wesentliche Merkmale Eingang in die Schulgesetzgebung des Landes gefunden haben: die ersten beiden Jahrgangsstufen der Grundschule sind als pädagogisch-curriculare Einheit definiert und sollten jahrgangsübergreifend organisiert sein (SenBWF, 2005 i.d.F.v. 2012). Seit einer Gesetzesänderung im Jahr 2012 kann die Schulkonferenz mit einer Zweidrittelmehrheit beschließen, die Schulanfangsphase jahrgangsbezogen zu organisieren (ebd.). Auch das flexible Verweilen ist gesetzlich gesichert (SenBWF, 2005 i.d.F.v. 2012), wobei ungeklärt bleibt, wie ein Kind im jahrgangsbezogen organisierten Schulanfang verweilen kann oder ob dann nicht doch von Wiederholen gesprochen werden sollte. Den Schuleintritt betreffend hat das Land Berlin dem enormen öffentlichen und politischen Druck nachgegeben und das ab 2005/06 auf sechs Jahre und einen Monat gesenkte Schuleintrittsalter inzwischen auf sechs Jahre und vier Monate erhöht (SenBWF, 2004 i.d.F.v. 2016). Auch wurden die Hürden zur Einschulungsrückstellung bereits seit dem Schuljahr 2007/08 sukzessive abgebaut, sodass die

Anzahl zurückgestellter Kinder seitdem erneut kontinuierlich gestiegen ist (Autorengruppe Bildungsberichterstattung, 2016). Die integrative Beschulung aller Kinder in allen Fächern ist in der aktuellen Version des Schulgesetzes verankert (SenBWF, 2004).

Bei der Neukonzeption des Schulanfangs, nicht nur in Berlin, stellen die Senkung des Schuleintrittsalters mit dem Ziel einer langfristigen Senkung des Berufseintrittsalters einerseits und die Einführung des Jahrgangsübergreifenden Lernens zum besseren Umgang mit Heterogenität andererseits die wesentlichen Neuerungen dar. Sie stehen daher auch im Mittelpunkt der vorliegenden Arbeit und werden im Folgenden ausführlicher beschrieben.

2.4 Einführung des Jahrgangsübergreifenden Lernens (JÜL)

Die Umsetzung des Jahrgangsübergreifenden Lernens (JÜL), also des gemeinsamen Lernens von Schüler*innen aus mindestens zwei Jahrgangsstufen in einer Lerngruppe, begann in Berlin in einigen wenigen öffentlichen Schulen bereits in den 1990er Jahren. Auf Eigeninitiative erhielten ab dem Schuljahr 1993/94 bzw. 1994/95 zwei Grundschulen die Möglichkeit ihren Schulanfang jahrgangsübergreifend zu organisieren (Wilde, 2006). Ab dem Schuljahr 1999/00 konnten im Rahmen des Schulversuchs „Unterricht in Jahrgangsübergreifenden Lerngruppen (JÜL)“ weitere Schulen ihren Unterricht jahrgangsübergreifend organisieren (Wilde, 2008). Im Auftrag der Senatsverwaltung für Schule, Jugend und Sport wurde die Arbeitsstelle Bildungsforschung Primarstufe der Freien Universität mit der Untersuchung der Einführung und Erprobung von JÜL beauftragt. Unter der Leitung von Prof. Dr. Diemut Kucharz (Pädagogische Hochschule Weingarten) und unter Mitarbeit von Dr. Matthea Wagener (Freie Universität Berlin) wurde die Studie während der Schuljahre 2001/02 bis 2003/04 durchgeführt (Kucharz & Wagener, 2007, S. 7). Die Ergebnisse der Untersuchung wurden seitens der Senatsbildungsverwaltung insgesamt positiv bewertet (Abgeordnetenhaus Berlin, 2010b; Wilde, 2006) und dienten somit als wissenschaftliche Grundlage für die flächendeckende Einführung von JÜL im Schulanfang an Berliner Grundschulen.

Nach der Einführung der SAPH zum Schuljahr 2005/06 sollten mit Beginn des Schuljahres 2006/07 alle SAPH-Klassen in staatlichen Grundschulen in Berlin jahrgangsübergreifend organisiert werden. Diejenigen Lehrkräfte, die in diesem Schuljahr eine SAPH-Klasse übernehmen sollten, wurden im Rahmen von Fortbildungen vorbereitet (Wilde, 2006). Zwischen 2004 und Anfang 2007 wurden nach Angaben der Senatsbildungsverwaltung etwa 3.000 Lehrkräfte in zentralen Fortbildungen und (bezirkliche) Multiplikator*innen in

regionalen Fortbildungen und LISUM¹⁷-Konferenzen auf die Einführung von JÜL vorbereitet. Zudem hospitierten im Schuljahr 2005/06 über 3.000 Lehrkräfte an 19 Schulen des ehemaligen Schulversuchs „Jahrgangsübergreifendes Lernen“, die Lehrkräfte dieser Schulversuchsschulen führten etwas unter 100 Studientage und Informationsveranstaltungen zu JÜL und SAPH durch und circa 150 Erzieher*innen und Lehrkräfte nahmen an sechs- bis zwölfmonatigen Intensivkursen zur Einführung der SAPH teil (Abgeordnetenhaus Berlin, 2007c). So sollte neben der Qualifikation der beteiligten Einzelpersonen sichergestellt werden, dass in jedem der zwölf Berliner Bezirke Fachmultiplikator*innen zu JÜL als Ansprechperson und Leitung der Fachkonferenz „JÜL“, bei der jede Grundschule durch eine Lehrperson vertreten wäre, zur Verfügung stehen (Abgeordnetenhaus Berlin, 2008). Diese Fachmultiplikator*innen wurden von einer Moderatorin aus dem LISUM fachlich begleitet (ebd.). Darüber hinaus erhalten JÜL-Klassen in der Schulanfangsphase Unterstützung durch die Zuteilung von vier Erzieher*innenwochenstunden, plus zwei Lehrer*innenwochenstunden oder drei weitere Erzieher*innenwochenstunden oder 3.200 Euro Projektmittel (Abgeordnetenhaus Berlin, 2012b). Als nach den Wahlen zum Regierenden Bürgermeister 2006 Prof. Dr. E. Jürgen Zöllner (SPD) das Amt des zuständigen Senators von Klaus Böger (SPD) übernahm, wurde die Einführung von JÜL jedoch um ein Jahr verschoben: seitens der Bildungsadministration wurde die Verschiebung offiziell damit begründet, dass sie „den Wünschen von Schulen Rechnung“ trage, „die in ihren Vorbereitungen auf die Realisierung der Jahrgangsmischung noch nicht weit genug fortgeschritten sind, die in ihrem Personalentwicklungs- und Fortbildungskonzept für die Qualitätsentwicklung der SAPH mehr Zeit benötigen oder deren pädagogisches Konzept eine Mischung von drei Jahrgängen vorsieht.“ (Abgeordnetenhaus Berlin, 2007c). Einen weiteren Grund stellt die Tatsache dar, dass aufgrund der Senkung des Schuleintrittsalters (siehe Abschnitt 2.5) im Schuljahr 2005/06 einmalig ein halber Jahrgang mehr eingeschult wurde (Abgeordnetenhaus Berlin, 2004) und somit die (Alters-)Heterogenität dieses Jahrgangs ohnehin bereits größer war als in den Jahren zuvor. Die Schulen sollten nun mehr Zeit zur Vorbereitung erhalten. Jedoch bedeutete diese Verschiebung auch, dass einerseits die nun fortgebildeten Lehrkräfte in jahrgangsbezogenen Lerngruppen eingesetzt wurden sowie andererseits die Lehrkräfte für den folgenden Jahrgang aufgrund der Kurzfristigkeit der Entscheidung nicht mehr in diesem Umfang fortgebildet werden konnten und somit vergleichsweise unvorbereitet ihre neuen Klassen übernahmen (Wilde, 2014).

¹⁷ Landesinstitut für Schule und Medien Berlin-Brandenburg

Im Schuljahr 2008/09 trat die verpflichtende Umsetzung von JÜL für die Jahrgangsstufen der SAPH an allen öffentlichen Berliner Grundschulen schließlich in Kraft. Doch bereits ab 2008/09 erhielten die Schulen keinen Stichtag mehr, bis wann sie JÜL umzusetzen hatten. Sie sollten jedoch in Absprache mit ihrer regionalen Schulaufsicht einen Zeit- und Maßnahmenplan für die Umsetzung erarbeiten (Abgeordnetenhaus Berlin, 2010b). Ein Schuljahr später hatten nach Auskunft der Senatsbildungsverwaltung 85 % der Schulen JÜL eingeführt und alle übrigen Schulen die entsprechenden Zeit- und Maßnahmenpläne abgeschlossen (ebd.). Zwei weitere Jahre später, im Schuljahr 2011/12, wurde es Schulen, die JÜL bis dahin noch nicht umgesetzt hatten, ermöglicht, von der Einführung gänzlich abzusehen (Abgeordnetenhaus Berlin, 2011, Abgeordnetenhaus Berlin, 2012c) und ab 2012/13 stand es allen Schulen frei, ihre SAPH jahrgangsgemischt oder -übergreifend zu organisieren (Abgeordnetenhaus Berlin, 2012a, Abgeordnetenhaus Berlin, 2012b). Bei der letztlichen Aufhebung der Verpflichtung, aber auch bereits bei der verzögerten Einführung von JÜL kann davon ausgegangen werden, dass die zum Teil sehr kritische öffentliche Wahrnehmung und Berichtserstattung der regionalen und überregionalen Presse die politischen Entscheidungsprozesse zumindest anteilig beeinflusst hat. Als Argumente gegen JÜL wurden dabei, ähnlich wie schon bei der Neukonzeption der SAPH allgemein, die Überforderung der Lehrkräfte durch den erhöhten Arbeitsaufwand, mangelhafte Fortbildungen und fehlendes Personal für die Umsetzung (Abgeordnetenhaus Berlin, 2007e), aber auch befürchtete schwächere Leistungen der Schüler*innen (ebd.), fehlende Strukturen (Abgeordnetenhaus Berlin, 2006b) und die Herausforderung bei der Umsetzung in sehr heterogenen Klassen (Abgeordnetenhaus Berlin, 2006c) genannt. Auch wurde JÜL als „Verordnung von oben“ (Abgeordnetenhaus Berlin, 2014; Klesmann, 2011) oder „Experiment“ (Schneider, 2013) verstanden.

Im Schuljahr 2014/15 wurde nach Auskunft der Senatsbildungsverwaltung noch in 44 % der SAPH-Klassen jahrgangsübergreifend unterrichtet (Abgeordnetenhaus Berlin, 2015).

2.5 Senkung des Schuleintrittsalters

Wie auch andere Bundesländer, entschloss sich das Land Berlin im Anschluss an die Empfehlungen der KMK zum Schulanfang (KMK, 1997) und als Reaktion auf das vergleichsweise hohe Berufseinstiegsalter in Deutschland zu einer Senkung des Schuleintrittsalters. Entgegen der KMK-Empfehlungen sollte das Durchschnittsalter Berliner Schüler*innen nicht nur um drei, sondern sechs Monate herabgesetzt werden. Zudem wurde,

gemäß den KMK-Empfehlungen, eine Beschulung aller Schüler*innen angestrebt und auf Rückstellungen sollte verzichtet werden.

Zum Schuljahr 2005/06 wurde jedes Kind als schulfähig eingestuft, das bis zum 31. Dezember 2005 das sechste Lebensjahr vollendete (Abgeordnetenhaus Berlin, 2004; SenBWF, 2004). Diese Stichtagsverschiebung um sechs Monate – zuvor war der 30. Juni als Stichtag festgelegt – hatte die Senkung des mittleren Einschulungsalters von sechs Jahren und sieben Monate auf sechs Jahre und einen Monat zum Ziel und sollte somit dazu beitragen, das Alter Berliner Absolvent*innen langfristig zu senken. Parallel zu dieser Stichtagsverschiebung wurde, im Sinne einer Flexibilisierung des Schuleintritts, die Fähigkeitsfeststellung durch den Schularzt nur noch bei „Kann-Kindern“, die vorzeitig eingeschult werden sollten, sowie in Ausnahmefällen, z. B. bei stark lernbeeinträchtigten Kindern, durchgeführt. Für die große Mehrheit der Kinder galt nun einzig ihr Alter als Kriterium für die Schulpflicht. Dennoch sollte mit dieser Reform, nämlich durch die Einschulung aller Kinder, dem Versuch weg vom Konzept des schulreifen Kindes (Kern, 1951) hin zu einem Konzept der „kindfähigen“ Schule (Abgeordnetenhaus Berlin, 2004; Stipek, 2002), Rechnung getragen werden. In diesem Sinne wurde auch zunächst völlig auf Rückstellungen verzichtet. Doch bereits ab dem Schuljahr 2007/08 wurden die Regelungen zur Rückstellungen sukzessive gelockert, wenn auch zunächst nicht offiziell, und wurden seitdem mehr und mehr aufgeweicht, sodass spätestens seit dem Schuljahr 2011/12 Eltern auch offiziell wieder die Möglichkeit haben, ihr Kind um ein Jahr verzögert einzuschulen (Abgeordnetenhaus Berlin, 2010a). Als Reaktion auf die erneut steigenden Rückstellungsquoten, aber auch auf die anhaltende öffentliche Kritik an der sogenannten Früheinschulung (Hecht, 2010; Vieth-Entus, 2012, 2013a, 2013b), wurde zum Schuljahr 2017/18 die Verschiebung des Einschulungsstichtages auf den 30. September und somit die Anhebung des mittleren Einschulungsalters auf sechs Jahre und vier Monate beschlossen (SenBJF, 2017). Zudem haben Eltern heute die Möglichkeit, wie auch vor der Einschulungsreform, ihr Kind vorzeitig (bis zu einem halben Jahr) oder verzögert (mit rechtzeitiger schulärztlicher Untersuchung) einschulen zu lassen (ebd.).

2.6 Zusammenfassung

Seit den ersten Überlegungen zur Beschreibungen der Erwartungen an Schulanfänger*innen von Artur Kern (1951), die er unter dem Begriff Schulreife zusammenfasst, hat sich die Diskussion einerseits wegbewegt von einem alleinigen Fokus auf die für die Schule

notwendigen Fähigkeiten der Schulanfänger*innen und der Annahme, dass diese Fähigkeiten nicht verändert werden können, hin zu der Annahme, dass individuelle Fähigkeiten nicht nur gefördert werden können, sondern dass auch die Schule in der Lage sein muss, mit unterschiedlichen Lernvoraussetzungen der Schulanfänger*innen umzugehen (Abgeordnetenhaus Berlin, 2004). In Deutschland wird zur Beschreibung dieser Annahme der Begriff der Schulfähigkeit verwendet (Kammermeyer, 2000). Andererseits zeigt sich, nicht nur in der öffentlichen Debatte, eine anhaltende Fokussierung auf das Alter eines Kindes bei der Einschulung (Kerstan & Spiewak, 2008; Müller-Lissner, 2015; Stamm, 2013).

Aufbauend auf der Annahme einer Schulfähigkeit und als Antwort auf die Ergebnisse der internationalen Schulleistungstudien, wurde zu Beginn der 2000er Jahre die Neukonzeption des Schulanfangs auf den Weg gebracht (Abschnitt 2.2 f.). Kennzeichen der neuen SAPH in Berlin waren die pädagogische-curriculare Zusammenfassung der ersten beiden Jahrgangsstufen, die Herabsetzung des mittleren Einschulungsalters mithilfe einer Verschiebung des Einschulungstichtages (vom 30. Juni auf den 31. Dezember) und zunächst auch dem Verzicht auf die Möglichkeit zur Rückstellung von Schulanfänger*innen um ein Jahr (Abgeordnetenhaus Berlin, 2004). Mit der Einführung des Jahrgangsübergreifenden Lernens als ein Merkmal der neuen SAPH und der damit einhergehenden Möglichkeit einer flexiblen Verweildauer von ein bis drei Jahren sollte auf die steigende Heterogenität der Berliner Schülerschaft reagiert werden können (Abgeordnetenhaus Berlin, 2010c). Ziel war es dabei, dass Schüler*innen mit ihren unterschiedlichen Lernausgangslagen den Übergang in die dritte Jahrgangsstufe und somit in die meist jahrgangsbezogene Unterrichtsorganisation erfolgreich bewältigen können (Abgeordnetenhaus Berlin, 2004). Die beiden wesentlichen Neuerungen im Grundschulbereich, die Einführung von JÜL und die Herabsetzung des Schuleintrittsalters, wurden in Berlin zunächst mithilfe des neuen Schulgesetzes (SenBWF, 2004) rechtlich verbindlich eingeführt, jedoch auch aufgrund des großen öffentlichen und politischen Drucks zumindest teilweise wieder zurückgenommen. Für das Schuljahr 2017/18 bedeutet das, dass alle Kinder schulpflichtig werden, die bis zum 30. September 2017 das sechste Lebensjahr abgeschlossen hatten. Kinder, die bis zum 31. März 2018 das sechste Lebensjahr abschließen, sind sogenannte „Kann-Kinder“ und können auf Antrag der Eltern vorzeitig eingeschult werden. Schulpflichtige Kinder (zwischen 01.01.2011 und 30.09.2011 geboren) können auf Antrag ihrer Eltern und nach Besuch eines Schularztes um ein Jahr von der Einschulung zurückgestellt werden. Das Jahrgangsübergreifende Lernen wird im Schulgesetz für die Organisation des Unterrichts in der SAPH empfohlen. Schulen können aber in der

Schulkonferenz mit einer Zweidrittelmehrheit die jahrgangsbezogene Organisation ihrer SAPH beschließen.

3. Zielsetzung der Arbeit

Die vorangestellten Kapitel haben gezeigt, dass die Implementation politisch motivierter Bildungsreformen sowohl die Initiator*innen der Reform, hier die Senatsbildungsverwaltung (aktuell SenBJF) in Berlin, als auch die betroffenen Akteur*innen (Schulen, Schulleitungen, Lehrkräfte, Schüler*innen und ihre Eltern) vor große Herausforderungen stellt und als forschungstheoretisch komplexer Prozess anerkannt wird (Altrichter & Wiesinger, 2005; Goldenbaum, 2012; Gräsel & Parchmann, 2004; Hall & Hord, 2015). Vor allem wenn, wie in den hier untersuchten Reformen der Fall, die Reformen alle Grundschulen des Bundeslandes betreffen und vornehmlich top-down implementiert wurden.

Mit den Studien dieser Arbeit soll der Versuch unternommen werden den Implementationserfolg bei beiden umfassenden und sowohl öffentlich als auch politisch kontrovers diskutierten Reformen im Zuge der neu geschaffenen Berliner Schulanfangsphase, JÜL und Früheinschulung, abzuschätzen.

3.1 Studien zur Einführung des Jahrgangsübergreifenden Lernens (JÜL)

Der Erfolg der Implementation von JÜL in die SAPH wird mithilfe von zwei Fragestellungen untersucht. Dabei wird zunächst, im Anschluss an die theoretischen Überlegungen zum Implementationserfolg, beschrieben, wie JÜL in den Schulen umgesetzt wurde. Im Anschluss werden auf Grundlage dieser Beschreibungen und der spezifischen pädagogischen Anforderungen an die Umsetzung von JÜL Unterschiede der Schul- und Unterrichtsqualität zwischen Typen der JÜL-Implementation untersucht.

3.1.1 Studie I: Wie wurde das Jahrgangsübergreifende Lernen an Berliner Grundschulen umgesetzt?

Aufgrund der besonderen Situation in Berlin – Schulversuche zu JÜL, Umsetzungspflicht für alle Schulen, Rücknahme der Umsetzungspflicht (siehe Kapitel 2) – ist es hier, den von Coburn (2003) beschriebenen Erfolgsindikatoren von Implementationsprozessen folgend, zunächst notwendig zu beschreiben, wie JÜL über die Zeit in den Schulen, aber auch auf Klassenebene umgesetzt wurde, also in den Schulen verbreitet ist (*spread*). Haben sich die Schulen eher an die Vorgaben der Bildungsadministration gehalten und zwei Jahrgangsstufen zu einer jahrgangsübergreifenden Lerngruppe zusammengefasst (JÜL-2) oder sind sie dem (reform-) pädagogischen Ideal gefolgt (s. Abschnitt 4.2.1) und haben die offizielle Möglichkeit genutzt,

drei Jahrgangsstufen als Basis für eine Lerngruppe zusammenzufassen (JÜL-3; u. a. Laging, 2010)? Weiterhin soll im Anlehnung an die Diffusionstheorie von Rogers (2003) versucht werden, die Schulen, entsprechend ihrer JÜL-Umsetzung über die Zeit, bestimmten Typen zuzuordnen. Dabei wird davon ausgegangen, dass sich die Anwendungskategorien, die sich bei Rogers auf Individuen beziehen, auch auf Organisationen übertragen lassen. An dieser Stelle wird ebenfalls geprüft, inwiefern Schulen JÜL dauerhaft (*sustainability*) umsetzen (Coburn, 2003).

Im zweiten Teil der ersten Studie wird nach der längsschnittlichen die Perspektive auf den Querschnitt gelegt und versucht, einen vertiefenden Einblick zur Verbreitung vom JÜL innerhalb der Schulen zu erhalten. Dabei interessiert einerseits, in welchem Stundenumfang an den Schulen jahrgangsübergreifend unterrichtet wird, und andererseits, wie sich die Umsetzung in verschiedenen Unterrichtsfächern unterscheiden.

3.1.2 Studie II: Unterscheiden sich Schulen entsprechend ihrer Zuordnung zu einem Implementationstypen in ihrer Schul- und Unterrichtsqualität?

Vor dem Hintergrund der besonderen Bedeutung der Schul- und Unterrichtsqualität für die erfolgreiche Umsetzung von JÜL (Carle & Metzen, 2014; siehe Abschnitt 4.1) sollen im Anschluss an Studie I und unter Verwendung der identifizierten Typen der JÜL-Implementation (nach Rogers, 2003) Unterschiede zwischen den Typen und der Qualität auf Ebene der Schule und des Unterrichts überprüft werden. Diese Analysen sollen Hinweise über die Implementationstiefe (*depth*; Coburn, 2003) geben.

3.2 Studien zur Senkung des Schuleintrittsalters

Die Untersuchung des Implementationserfolgs der sogenannten Früheinschulung erfolgt im Rahmen von drei Studien. Das erklärte Ziel der Reform war die Senkung des mittleren Einschulungsalters. Studie III dieser Arbeit und die erste Studie zum Schuleintrittsalter widmet sich daher zunächst der Frage, ob dies in Berlin gelungen ist. Einen weiteren Indikator für den Implementationserfolg stellt die erfolgreiche Integration der nun früher eingeschulter Kinder in den Schulen dar. Daher wird in Studie IV am Beispiel der Schüler*innenleistungen die Auswirkungen der Reform untersucht (relative Alterseffekte). Die dritte Studie zum Schuleintrittsalter (Studie V) befasst sich, aufbauend auf Studie IV, genauer mit den in der Forschung zu Alterseffekten viel diskutierten sogenannten relativen Alterseffekten. Es steht

dabei die Frage im Mittelpunkt, ob sich relative Alterseffekte über verschiedene Schüler*innengruppen und -kohorten sowie über Klassen hinweg verallgemeinern lassen.

3.2.1 Studie III: Umsetzung der Senkung des Schuleintrittsalters

Mit der Verankerung des neuen Stichtags zur Einschulung im Schulgesetz (SenBWF, 2004) und der damit vorgezogenen Schulpflicht standen Berliner Grundschulen nicht vor der Entscheidung, ob sie die vergleichsweise jüngeren Kinder aufnehmen. Ab dem Schuljahr 2005/06 wurden alle Kinder, die bis zum 31. Dezember 2005 das sechste Lebensjahr abgeschlossen hatten, eingeschult. Studie III untersucht zunächst, ob die Reform wie gewollt umgesetzt wurde und das Einschulungsalter entsprechend der bildungspolitischen Vorgaben im Mittel gesenkt werden konnte. Um sicherzustellen, dass die weiteren Analysen zum Implementationserfolg keine systematischen Verzerrungen beinhalten, wird im zweiten Teil der Studie geprüft, ob sich systematische Unterschiede des mittleren Einschulungsalters zwischen den einzelnen SAPH-Klassen feststellen lassen. Mithilfe dieser Analysen sollen Rückschlüsse auf die Verbreitung (Coburn, 2003) der Einschulungsreform gezogen werden.

3.2.2 Studie IV: Wie wirkt sich die Senkung des Schuleintrittsalters auf die Schüler*innenleistung aus?

Mit Blick auf die Implementationstiefe kann als weiterer Indikator für den Implementationserfolg der Einschulungsreform die schulische Integration der nun früher eingeschulten Schüler*innen herangezogen werden. Studie IV untersucht dies mithilfe der schulischen Leistungen. Dabei werden die Leistungen der Schüler*innen einer Klassenstufe ihrem Alter entsprechend gegenübergestellt (relativer Alterseffekte). Dabei wird geprüft ob sich altersspezifisch systematische leistungsbezogene Unterschiede, vor allem zuungunsten der nun früher eingeschulten Schüler*innen, feststellen lassen. Zudem sollen Verteilungen der Schüler*innen entsprechend ihrem Alter zeigen, ob sich Anteile bestimmter Altersgruppen über die Zeit verringern und es somit Hinweise auf eine erhöhte Anzahl an Klassenwiederholungen gibt. Dies könnte ein Indikator für eine nicht erfolgreiche Integration der vergleichsweise Jüngeren darstellen.

3.2.3 Studie V: Lassen sich systematische Leistungsunterschiede zwischen vergleichsweise jüngeren und älteren Schüler*innen (relative Alterseffekte) über verschiedene Schüler*innengruppen, Jahrgänge, Kohorten und Klassen verallgemeinern?

Studie V nimmt erneut relative Alterseffekte, also systematische Leistungsunterschiede zwischen vergleichsweise jüngeren oder älteren Schüler*innen, in den Blick und stellt dabei die Generalisierbarkeit dieser Effekte in den Mittelpunkt der Untersuchungen. Die Verallgemeinerung von Forschungsergebnissen stellt eine wichtige Anforderung an die wissenschaftliche Praxis und die Voraussetzung gesicherten Erkenntnisgewinns dar (Cumming, 2012; Shadish, Cook & Campbell, 2002). Studie V widmet sich im Einzelnen der Generalisierbarkeit relativer Alterseffekte in den Lernbereichen Deutsch / Lesen und Mathematik über verschiedene Schüler*innengruppen, Jahrgangsstufen und Kohorten sowie Klassen und gibt so Aufschluss darüber, ob sich relative Alterseffekte vor und nach der Einschulungsreform verändert haben.

4. Studien zur Einführung des Jahrgangsübergreifenden Lernens

Die Frage nach der (optimalen) Zusammensetzung von Lerngruppen beschäftigt die Pädagogik seit jeher (Laging, 2010). Die Idee der Jahrgangsklasse, als heute am meisten verbreitete Form der Lerngruppenorganisation in deutschen Schulen, geht im Wesentlichen auf den tschechischen Philosophen und Pädagogen Comenius zurück (ebd.). Mitte des 17. Jahrhunderts sprach er sich dafür aus, dass in Schulen weitgehend derselbe Lehrstoff vermittelt werden sollte, mit dem Ziel, „der gesamten Jugend beiderlei Geschlechts eine Bildung in den Wissenschaften zu ermöglichen“ (Laging, 2010, S. 7). Ausgehend von der Annahme, „daß alle Schüler wie weißes Papier mit den immer gleichen Druckformen behandelt werden könnten, um so eine Vervielfältigung zu ermöglichen“ (ebd., S. 8), hielt er es für notwendig, die Methoden und die Organisation in der Schule zu vereinheitlichen. So sollten erstmalig Schüler*innen nur noch an einem Tag im Jahr eingeschult werden und nach ihrem Alter in sogenannten Jahrgangsklassen zusammen lernen. Weiterhin sollte jede Jahrgangsstufe ihr eigenes, im Laufe des Schuljahres zu bearbeitendes Lehrbuch erhalten, sodass ein gemeinsames Anfangen, Fortschreiten und letztlich auch Schließen des Schulunterrichts möglich werden sollte. Die dominierende Lehr- und Lernform dieser Gleichschrittigkeit war die Belehrung, wobei die Lehrkraft stets die Rolle der Lehrenden und die Schüler*innen die der Lernenden einnahmen. Trotz anhaltender Kritik an den Homogenisierungsbestrebungen in der Schule und recht erfolgreicher Alternativen, z. B. des Fachklassensystems nach H. A. Francke¹⁸, setzte sich die Jahrgangsklasse vor allem ab der zweiten Hälfte des 19. Jahrhunderts in den Volksschulen mehr und mehr durch¹⁹ (Laging, 2010, S. 9).

Obwohl sich die Jahrgangsklasse im Laufe des 20. Jahrhunderts im deutschen Schulsystem grundlegend etabliert hat, scheint die kritische Auseinandersetzung mit dieser Lernform innerhalb der Pädagogik spätestens seit dem Jahr 2000 wieder verstärkt stattzufinden. Hauptkritikpunkt stellt hier die nicht erreichte Homogenität der Klasse dar (vgl. Carle & Metzen, 2014). Laging (2010) verdeutlicht die Problematik der Homogenitätsbestrebungen am Beispiel der Alterszusammensetzung einer Jahrgangsklasse. Bereits im voranschreitenden 19. Jahrhundert wurde die Altershomogenität einer Klasse durch eine hohe Anzahl an Schüler*innen, die pflichtmäßig ein Schuljahr wiederholten, also sitzenblieben, infrage gestellt.

¹⁸ Hier sollte jede*r Lernende je nach Kenntnis in dem jeweiligen Fach in eine angemessene Klasse versetzt werden (vgl. Laging, 2010, S. 8f.).

¹⁹ U.a. auch deshalb, weil Widerstände der Bevölkerung gegen die Einführung der Jahrgangsklassen an Gymnasien 1834 in Preußen durch staatliche Verordnung gebrochen wurden (Laging, 2010, S. 9).

Und auch heute können Schüler*innen wiederholen oder überspringen ein Schuljahr oder werden vorzeitig eingeschult oder von der Einschulung zurückgestellt (Stanat et al., 2002). Daraus ergibt sich für viele Jahrgangsklassen ein Altersspektrum zwischen drei und fünf Jahren, wobei als „normal“ gilt, wer entsprechend dem Stichtag eingeschult wurde und jedes Jahr eine Klassenstufe aufsteigt. Ältere oder jüngere Schüler*innen werden mit einem „Makel des Versagens oder dem Schein der besonderen Eignung versehen“ (Laging, 2010, S. 10). Doch selbst wenn eine Jahrgangsklasse bezüglich des tatsächlichen Alters ihrer Lernenden vergleichsweise homogen ist, so hat ihre Lehrkraft, vor allem zu Schulbeginn, ohnehin mit zum Teil sehr verschiedenen Eingangsvoraussetzungen und damit einhergehend unterschiedlicher kognitiver und sozial-emotionaler Entwicklung ihrer Schüler*innen umzugehen. Dennoch suggeriert „das Konstrukt der Jahrgangsklasse (...)“, so fasst Carle (2014) zusammen, „offenbar mehr Homogenität als vorhanden ist oder dient als Vorwand für gleichschrittiges Vorgehen trotz dominierender Heterogenität“ (ebd., S. 10f.).

Zum verbesserten Umgang mit der Heterogenität in Schulen und Klassen wird immer wieder das altersgemischte bzw. jahrgangsübergreifende Lernen (JÜL) diskutiert. Ursprünglich, vor allem in ländlichen Regionen, die einzige Möglichkeit zum Umgang mit geringen Schüler*innenzahlen und daher auch heute von Lehrkräften und Eltern als rückschrittig empfunden, diente es der Reformpädagogik zu Beginn des 20. Jahrhunderts als Antwort auf die Unterschiedlichkeit der Lernenden. Vor allem die Konzepte von Maria Montessori, Berthold Otto und Peter Petersen können als direkte Grundlagen für die heute diskutierten Formen von JÜL verstanden werden (zusammenfassend: Laging, 2010). So schlagen Montessori und Petersen alters- und geschlechtsgemischte Lerngruppen über drei Altersjahrgänge explizit vor, da, so die Annahme, somit die Unterschiede groß genug sind, um den jeweils anderen als jeweils anders zu unterscheiden (ebd.). Hier galt und gilt die Unterschiedlichkeit der Kinder nicht als Last, sondern als Ressource, mit der es sich zu arbeiten lohnt. Aus diesen beiden Konzeptionen und Motivationen zum Jahrgangsübergreifenden Lernen heraus wird auch heute noch nach JÜL als eine Form der Unterrichtsorganisation, wie sie vor allem in ländlichen Regionen als Umgang mit geringen Schüler*innenzahlen praktiziert wird, und JÜL als umfassendes, die ganze Schule betreffendes Konzept, dem eine pädagogische Motivation zugrunde liegt, unterschieden. International werde diese beiden Konzepte als *multi-grade* und *multi-age* bezeichnet (Veenman, 1995).

In der aktuellen Auseinandersetzung wird JÜL vordergründig pädagogisch motiviert und im Zusammenhang mit dem Schulanfang diskutiert. Angesichts der vielerorts großen

Heterogenität insbesondere der Schulanfänger (Roßbach, 2010; SenBJW, 2013b) „wird eine jahrgangsgemischte Eingangsklasse mit flexibler Verweildauer als eine Möglichkeit angesehen, den Übergang von Familie und Kindergarten in die Schule für die Kinder erfolgreicher zu gestalten“ (Kucharz & Wagener, 2007, S. 10).

Das folgende Kapitel umfasst zwei Studien zum jahrgangsübergreifenden Lernen in der Berliner Schulanfangsphase. Ziel der Studien ist es, die Implementation von JÜL umfassend zu beschreiben: Studie I (Abschnitt 4.1) mit der Leitfrage, wie JÜL an Berliner Grundschulen implementiert werden konnte, sowie die sich daran anschließende Studie II (Abschnitt 4.2), die die Zusammenhänge von erfolgreicher JÜL-Implementation und Qualität einer Schule und ihres Unterrichts in den Blick nimmt. Der abschließende Abschnitt 4.3 fasst die wichtigsten Erkenntnisse des Kapitels zusammen.

4.1 Studie I: Wie wurde das Jahrgangsübergreifende Lernen an Berliner Grundschulen implementiert?²⁰²¹

Für die Erfassung des Implementationserfolgs zum Jahrgangsübergreifenden Lernen in Berlin ist es zunächst wichtig, umfassend zu beschreiben wie JÜL an den Schulen umgesetzt wurde. Hierzu wird in Abschnitt 4.1.1 zunächst beschrieben, vor welchen spezifischen organisatorischen (z. B. räumliche Möglichkeiten; Professionalisierung der Lehrkräfte; Zusammenarbeit in multiprofessionellen Teams) und pädagogisch-didaktischen (z. B. Umgang mit Heterogenität; kooperative Lernformen; neue Anforderungen an die Lehrperson und ihre Rolle) Herausforderungen Schulen stehen, die ihre bisher jahrgangsbezogen organisierten Schulanfangsklassen ab einem festgelegten Zeitraum in jahrgangsübergreifende Klassen überführen sollen. Vor dem Hintergrund dieser spezifischen Herausforderungen, aber auch der allgemeinen Bedingungen und Hindernisse in Implementationsprozessen (Kapitel 1) und der Situation in Berlin (Kapitel 2), widmet sich Studie I der Beschreibung, wie JÜL in der SAPH implementiert wurde. Nach der theoretischen Einführung werden die Forschungsfragen abgeleitet (Abschnitt 4.1.2) und Methoden (Abschnitt 4.1.3) und Ergebnisse (Abschnitt 4.1.4) dargestellt. Abschnitt 4.1.5 fasst die wichtigsten Erkenntnisse aus der Studie zusammen.

4.1.1 Theoretischer Hintergrund

In Berlin waren einer verpflichtenden Umsetzung von JÜL mehrere Modellversuche an Schulen vorausgegangen, von denen einige an der Reformpädagogik orientierte Schulen bereits ab dem Schuljahr 1992/93 JÜL praktizierten (Wilde, 2006). Weitere Schulversuche folgten in den Schuljahren bis 2004/05, durch die die Jahrgangsmischung erprobt werden sollte (Abgeordnetenhaus Berlin, 2003, Abgeordnetenhaus Berlin, 2013; Ramseger, 2004; Wilde, 2006). Die Modellversuche liefen nach Ansicht der Schulen und der Bildungsadministration erfolgreich und motivierten zur verpflichtenden, flächendeckenden Umsetzung des JÜL-Konzepts in der Schulanfangsphase (Abgeordnetenhaus Berlin, 2004) Die ursprünglich für das Schuljahr 2007/08 geplante verpflichtende Einführung wurde allerdings nach kontroversen Diskussionen im Senat und aufgrund des Widerstands der Schulen um ein Jahr auf das Schuljahr 2008/09 verschoben. Der Widerstand hielt an, sodass ab dem Schuljahr 2012/13 den Schulen

²⁰ Teile dieser Studie wurden gemeinsam mit Studie II als Fachartikel in der *Zeitschrift für Erziehungswissenschaft* eingereicht. Dieser befindet sich derzeit unter folgendem Titel in Begutachtung: Thoren, K. & Brunner, M. (2016). Implementation des Jahrgangsübergreifenden Lernens in der Berliner Schulanfangsphase: Langfristige Umsetzung = Hohe Qualität?.

²¹ Teile dieser Studie sind in einem (unveröffentlichten) Projektbericht enthalten.

freigestellt wurde, ob sie jahrgangsübergreifend oder -bezogen in der Schulanfangsphase unterrichten (Koalitionsvereinbarung zwischen SPD Landesverband Berlin und CDU Landesverband Berlin für die Legislaturperiode 2011-2016, 2011). In Tabelle 4 ist der Entwicklungsverlauf der Implementation für den Zeitraum 2004 bis 2014 mit den jeweiligen Vorgaben dargestellt.

Tabelle 4: Entwicklungsverlauf der JÜL-Implementation in Berlin

Jahr	Vorgaben für die Implementation von JÜL
2004	Festlegung der Jahrgangsstufen 1 und 2 als pädagogisch-curriculare Einheit der flexiblen SAPH ¹ ; Einführung von JÜL im Schuljahr 2007/08 in der SAPH
2005	Festlegung, dass die Umsetzung von JÜL auch außerhalb der SAPH möglich ist; Zusammenlegung von maximal drei Jahrgängen
2006	Verschiebung der verbindlichen Einführung von JÜL in der SAPH um ein Jahr auf das Schuljahr 2008/09
2008	Übergangsregelung: individuelle Vereinbarungen der Schulen mit der Schulaufsichtsbehörde zur Einführung von JÜL, SAPH ist möglich; kein Stichtag mehr für Einführung von JÜL
2011	Seit dem Schuljahr 2011/12: Schulen, die von JÜL abweichen wollen, benötigen ein pädagogisch-didaktisches Konzept und den Beschluss der Schulkonferenz
2012	Seit dem Schuljahr 2012/13: dauerhafte Möglichkeit, die SAPH jahrgangsbezogen zu gestalten mit einer Zweidrittelmehrheit in der Schulkonferenz
2014	Erweiterung der SAPH auf drei Jahrgänge möglich

¹ SAPH = Schulanfangsphase

Eines der wesentlichen Merkmale des Jahrgangsübergreifenden Lernens ist das Miteinander-Lernen von Kindern aus unterschiedlichen Alters-, Leistungs- und Interessengruppen (Carle, 2013, S. 4). Ein Hauptargument stellt dabei die Tatsache dar, dass Kinder mit sehr unterschiedlichen Voraussetzungen eingeschult werden (u. a. Kucharz & Wagener, 2007, Laging, 2010): einige von ihnen haben bereits für ein oder mehrere Jahre eine vorschulische Einrichtung besucht (Schober & Spieß, 2012; Schober & Stahl, 2014), andere haben bereits zu Hause von ihren Eltern oder Geschwistern etliches gelernt oder die Familie bildet eine lernanregende Umgebung für das Kind (OECD, 2011). Wiederum andere Kinder kommen aus Elternhäusern, deren Alltagssprache eine andere als Deutsch ist, und haben so häufig einen Nachteil gegenüber ihren Mitschüler*innen mit deutscher Alltagssprache (Bos, Tarelli, Bremerich-Vos & Schwippert, 2012; Bos, Wendt, Köller & Selter, 2012; Kristen, 2006). Und auch ohne die Unterschiede in der sogenannten sozioökonomischen Herkunft, bringen unterschiedliche Interessen der Kinder und die bereits angesprochene große Spannweite in der Alterszusammensetzung (Abgeordnetenhaus Berlin, 2004; SenBJW, 2013b) eine umfassende

Heterogenität in jede Klasse, mit der pädagogisch und didaktisch nicht nur in den ersten Schuljahren, aber vor allem dort, umgegangen werden muss. Eine jahrgangsgemischte Lerngruppe ermöglicht dieses gemeinsame Arbeiten, da, so die Annahme, die Unterschiedlichkeit der Kinder mehr in den Mittelpunkt gerückt wird, diese sich somit in ihrer Unterschiedlichkeit verstärkt wahrnehmen und ihre Lernvoraussetzungen und Interessen als gleichberechtigt nebeneinander begreifen (Carle & Metzen, 2014). Die Stärken der einzelnen Kinder kommen hier, nicht nur für Eltern und die Lehrkraft, vor allem auch für das Kind selbst wahrnehmbar, in den Blick. Die gestärkte Selbstwahrnehmung der eigenen Kompetenzen können und sollen die Schüler*innen in den interaktiven und kooperativen Lernprozess einbringen (Carle & Metzen, 2014). Weiterhin regt die größere Differenz zwischen den Schüler*innen das gegenseitige Helfen an (vgl. Burk, 1996; Carle & Metzen, 2014; Laging, 2010), ein auch aus psychologischer Sicht erwünschtes, prosoziales Verhalten (Bierhoff, 1990; Kucharz & Wagener, 2007) und daher ein weiteres Argument für jahrgangsgemischtes Unterrichten. Folglich sollte das Unterrichten in einer JÜL-Klasse geprägt sein durch Formen des Kooperativen Lernens und Helfens sowie der Öffnung des Unterrichts (Carle & Metzen, 2014; Kucharz & Wagener, 2007). Zentrale Ziele jahrgangsübergreifender Lernformen haben Goodlad und Anderson bereits 1987 (übersetzt von Carle & Metzen, 2014, S. 17) formuliert:

- Entwicklung von selbstgesteuert handelnden und autonomen Individuen
- Schule hilft bei der bestmöglichen Entwicklung individueller Potenziale
- Jedes Individuum ist einzigartig und muss würdevoll und respektvoll behandelt werden
- Allseitige Entwicklung: ästhetisch, physisch, emotional und intellektuell
- Alle involvierten Akteur*innen sind Ko-Lernende
- Schulatmosphäre unterstützt Lernfreude

Für die pädagogische motivierte, erfolgreiche Umsetzung von JÜL haben Carle und Metzen (2014) in ihrer Erfassung des Forschungsstandes im Auftrag des Grundschulverbands eine anschauliche Zusammenfassung der Wirkungsbedingungen für JÜL herausgearbeitet (S. 10), die vor allem auf den Grundsätzen reformpädagogischer Modellvorstellungen beruhen. Der Darstellung zufolge kann Unterricht, der tiefgreifendes Lernen provoziert und alle Kinder einschließt, erreicht werden durch: (a) gesellschaftliche Unterstützung, (b) personelle Ressourcen, (c) Routinen und Traditionen, (d) Schule als tragfähige Basis, (e) Zustimmung der Eltern, (f) räumliche Möglichkeiten, (g) qualifiziertes Personal, (h) Heterogenität als Wert, (i) Schärfung des diagnostischen Blicks, (j) Öffnungsgrad des Unterrichts, (k) Kooperation der Kinder untereinander sowie (l) Schul- und Unterrichtsentwicklung. Dabei gelten diese

Bedingungen nicht nur für Jahrgangübergreifendes Lernen, sondern sollten Grundlage jedes Lernens sein. Für eine erfolgreiche Implementation von JÜL in einer Schule ist die Berücksichtigung dieser Bedingungen als Merkmale für hohe Schul- und Unterrichtsqualität jedoch von besonderer Bedeutung (Carle & Metzen, 2014; Hanke, 2008; Wendt, 2007).

Insgesamt kann zusammengefasst werden, dass eine jahrgangsgemischte oder -übergreifende Lerngruppe keine wesentlich abweichenden Anforderungen an die Qualität der Prozesse in Schule und Unterricht stellt, als dies bei einer Jahrgangsklasse der Fall sein sollte. Der Unterschied scheint jedoch darin zu liegen, dass in einer JÜL-Klasse die Heterogenität der Schüler*innen und ihres Lernens derart in den Mittelpunkt gerückt wird, dass, zumindest theoretisch, es keine andere Möglichkeit gibt, als damit zu arbeiten.

Befunde zur Implementation von JÜL gibt es u. a. in Form von wissenschaftlichen Untersuchungen für die flexible Schuleingangsphase (FLEX) in Brandenburg (Liebers et al., 2008) oder den Transfer der Schuleingangsphase in Thüringen (Carle & Metzen, 2009). Beide Projekte haben die Einführung und Umsetzung der SAPH, in Brandenburg auch mit einem besonderen Augenmerk auf JÜL, umfassend begleitet und evaluiert und beide Studien heben die Vorteile des integrativen, altersgemischten Schulanfangs hervor. Die Untersuchung zum Schulanfang in Thüringen zielt auf die „flächendeckende Einführung der integrativen, klassenstufenübergreifenden und flexiblen Schuleingangsphase“ (Carle & Metzen, 2014, S. 12) ab und hat daher einen formativ-evaluierenden Charakter (u. a. Wottawa & Thierau, 2003). Während es sich bei FLEX eher um eine umfassende summative Evaluation eines landesweiten Modellversuchs handelt. Über die flächendeckende Einführung und Umsetzung der Schulanfangsphase allgemein und spezifisch des Jahrgangübergreifenden Lernens liegen, nach Wissen der Autorin, jedoch bislang keine Studien vor. Mit Blick auf Berlin soll im Folgenden erstmalig eine landesweite Einführung von JÜL untersucht werden.

4.1.2 Forschungsfragen

Aufgrund der beschriebenen Gegebenheiten im Vorfeld der Einführung von JÜL in Berlin sowie der bei Implementationen üblichen Herausforderungen in der Praxis (u.a. Hasselhorn et al., 2014) ist zu vermuten, dass der Zeitpunkt und die Art und Weise, wie an Berliner Grundschulen JÜL umgesetzt wird, zwischen den Schulen erheblich variiert. Es ist anzunehmen, dass ein Teil der Schulen bereits mit Beginn der Verpflichtung zu JÜL (ab dem Schuljahr 2008/09) diese Form der Unterrichtsorganisation implementiert, andere Schulen hingegen JÜL später umsetzen bzw. nach Aufhebung der Verpflichtung zum jahrgangsbezogenen Unterricht in der

Schulanfangsphase zurückkehren werden. Auch wenn Schulen JÜL anbieten, kann die Intensität in der Gestaltung zwischen den Schulen variieren. So ist es möglich, dass einige Schulen durchgängig in allen Unterrichtsstunden jahrgangsübergreifend unterrichten, andere Schulen wiederum lediglich stundenweise, nur in einzelnen Fächern oder in Projektphasen jahrgangsgemischt arbeiten. Daher wird in Studie I das Wie der JÜL-Implementation in Berlin in den Blick genommen. Dabei ist besonders interessant, an wie vielen Schulen und innerhalb der Schulen in welchem Umfang JÜL vor dem Hintergrund der Anforderungen an JÜL, der Herausforderungen von Implementationsprozessen allgemein (Kapitel 1) und der spezifischen Bedingungen in Berlin (Abschnitte 2.3 bis 2.5) umgesetzt wurde. Der erste Teil der Untersuchungen in Studie I fokussiert auf die Umsetzung von JÜL über die Zeit: (I.1) An wie vielen Berliner Grundschulen und in wie vielen Klassen wurde JÜL in den Schuljahren 2003/04 bis 2015/16 umgesetzt? Ausgehend von der Annahme, dass JÜL dann erfolgreich umgesetzt werden kann, wenn es nicht nur in einer Klasse eingeführt wird (Carle & Metzen, 2014), stellt sich weiter die Frage (I.2), wie viele Klassen der Schulanfangsphase innerhalb der Schulen über die Zeit jahrgangsübergreifend organisiert wurden? In der (reform-) pädagogischen Tradition gilt JÜL dann als sinnvoll, wenn in einer Lerngruppe drei Jahrgangsstufen miteinander lernen (Dübbers & Gewerkschaft Erziehung und Wissenschaft, 2005; GEW Berlin, 2006; Laging, 2010). Die rechtlichen Grundlagen in Berlin schrieben eine Zusammenfassung zweier Jahrgangsstufen (JÜL-2) als Mindestanforderung vor, ermöglichten aber auch Klassen, in denen Schüler*innen aus drei Jahrgangsstufen zusammen lernten (JÜL-3). Daher wird im Folgenden untersucht, wie sich JÜL-2 und JÜL-3 über die Jahre in Berliner Grundschulen verteilen (Frage I.3). Der Diffusionstheorie von Rogers (2003) folgend, wird angenommen, dass sich auch Schulen, und nicht nur Individuen wie von Rogers postuliert, bei der Implementation von Reformen entsprechend ihrer Innovationsbereitschaft in unterschiedliche Typen zusammenfassen lassen. Frage I.4 prüft, ob sich Typen der JÜL-Implementation identifizieren lassen und, falls ja, wie diese beschrieben werden können.

Der zweite Abschnitt der Studien zur JÜL-Implementation versucht einen tiefgreifenden Einblick zu geben, wie JÜL in der Schulanfangsphase umgesetzt wurde. Dabei wird geprüft, in welchem Stundenumfang an den Schulen jahrgangsübergreifend unterrichtet wird (Frage I.5) und wie sich die Umsetzungen in verschiedenen Unterrichtsfächern unterscheiden (Frage I.6).

4.1.3 Methode

Der folgende Abschnitt stellt die Methode der Studie I dar. Nach einer Beschreibung der Datengrundlage (Abschnitt 4.1.3.1) und der daraus definierten Stichprobe (Abschnitt 4.1.3.2) wird die Analysestrategie für die einzelnen Fragestellungen (Abschnitt 4.1.3.3) im Detail dargestellt.

4.1.3.1 Datengrundlage

Grundlage für die Analysen zur Frage der Implementation bilden Archivdaten aller Berliner Grundschulen aus der Schulstatistik, die von der Senatsverwaltung für Bildung, Jugend und Wissenschaft (SenBJW) zur Verfügung gestellt wurden, sowie Daten einer Zusatzerhebung im Rahmen der Vergleichsarbeiten in der dritten Jahrgangsstufe (VERA 3) im Schuljahr 2012/13. Im Zentrum stehen die Daten aller öffentlichen Grundschulen des Landes Berlin. Dabei liegt der Fokus auf den Jahrgangstufen 1 und 2 bzw. der Schulanfangsphase aus den Schuljahren 2003/04 bis 2014/15. Von einer JÜL-Klasse wird seitens der Senatsbildungsverwaltung dann gesprochen, wenn mindestens ein Drittel der Schüler*innen einer Klasse einer höheren oder niedrigeren Jahrgangsstufe zugeordnet ist als die übrigen Schüler*innen dieser Klasse. Im Fall der SAPH besteht eine JÜL-Klasse also mindestens zu einem Drittel aus Schüler*innen der Jahrgangsstufe 1, während die übrigen Schüler*innen der Jahrgangsstufe 2 angehören, oder umgekehrt.

4.1.3.1.1 Schulstatistik des Landes Berlin

Kurz nach Beginn eines jeden Schuljahres wird zum Stichtag 1. November von der SenBJW die sogenannte „Oktoberstatistik“ an Berliner Schulen erhoben. Die Statistik umfasst Informationen zur genauen Anzahl der Lehrkräfte und Schüler*innen in den einzelnen Schulen, Jahrgangsstufen und Klassen und bildet unter anderem die Grundlage für das frei zugängliche Schulportrait auf den Internetseiten der Senatsverwaltung (SenBJW, 2016).

Für diese Arbeit standen aggregierte Datensätze auf Klassenebene aus den Schuljahren 2003/04 bis 2015/16 zur Verfügung. Diese Datensätze enthalten Informationen zur amtlichen Berliner Schulnummer, der Schulart, des Schulträgers, zur Anzahl der Klassen je Klassenart sowie die Summe der Schüler*innen und ihre Unterscheidung nach Geschlecht und Herkunftssprache je Klasse bzw. Klassenart. Die Daten der Schulstatistik eröffnen demnach die Möglichkeit, Fragestellungen sowohl bezogen auf die einzelnen Schuljahre, also im Querschnitt, als auch

hinsichtlich der Entwicklung von JÜL über die Zeit zu untersuchen. In Tabelle 5 ist eine Übersicht der Daten abgebildet.

Tabelle 5: Übersicht zu den Daten der Schulstatistik für die Schulanfangsphase (Schulen in öffentlicher Trägerschaft)

Schuljahr	Jahrgangsstufe*	N _{Schulen}	N _{Klassen}	N _{Schüler *innen}	% weiblich	% _{ndH}
2003/04	1	374	1.069	24.549	49,7%	30,7%
	2	380	1.056	24.532	50,0%	30,7%
	3	382	1.026	23.797	49,0%	30,3%
2004/05	1	372	1.070	24.899	49,1%	30,8%
	2	374	1.073	24.983	49,7%	31,7%
	3	378	1.059	25.062	49,6%	30,9%
2005/06	1	364	1.340	32.113	48,3%	32,8%
	2	372	1.077	25.719	49,6%	31,6%
	3	374	1.084	26.150	49,6%	31,6%
2006/07	SAPH	367	2.378	56.428	48,6%	32,9%
	3	372	1.114	27.080	49,6%	31,7%
2007/08	SAPH	351	1.984	46.760	49,3%	34,3%
	3	370	1.436	33.885	48,4%	34,2%
2008/09	SAPH	331	1.783	41.944	49,3%	34,4%
	3	370	1.334	31.415	48,8%	35,3%
2009/10	SAPH	313	1.707	40.368	48,8%	35,0%
	3	372	1.355	32.083	49,6%	36,5%
2010/11	SAPH	316	1.756	40.731	49,1%	36,5%
	3	371	1.430	32.952	49,0%	37,0%
2011/12	SAPH	323	1.819	41.731	48,5%	37,6%
	3	374	1.439	33.106	49,0%	38,1%
2012/13	SAPH	325	1.869	42.718	48,6%	38,0%
	3	374	1.480	34.171	49,0%	39,3%
2013/14	SAPH	325	1.961	44.918	48,9%	39,0%
	3	376	1.480	34.245	48,6%	39,3%
2014/15	SAPH	339	2.078	47.714	48,4%	39,4%
	3	378	1.489	34.339	49,2%	41,6%
2015/16	SAPH	337	2.121	48.634	48,4%	40,9%
	3	378	1.584	36.324	48,8%	41,8%

* Ab dem Schuljahr 2005/06 wurden die Jahrgangsstufen 1 und 2 zur Schulanfangsphase (SAPH) zusammengefasst

4.1.3.1.2 Erweiterungsstudie

Es wird angenommen, dass Schulen, die JÜL mit ernsthaftem Interesse einführten, die gesamte Anzahl der Jahresunterrichtsstunden in den beiden Hauptfächern jahrgangsübergreifend

organisieren und nicht, wie es in einigen Berliner Schulen zu vermuten ist, lediglich innerhalb der Schulprojektwochen oder in einzelnen Fächern JÜL anboten, die nicht zu den Hauptfächern zählen. In welchem Stundenumfang JÜL an Berliner Grundschulen umgesetzt wurde, wird im Rahmen der Schulstatistik des Landes Berlin jedoch nicht erfasst. Daher wurde die Erweiterungsstudie in Absprache mit der in der Senatsbildungsverwaltung zuständigen Expertin für den Bereich Schulanfang, Dagmar Wilde, erarbeitet. Konkret wurden im Schuljahr 2012/13 in einer Faxabfrage 123 zufällig ausgewählte öffentliche Berliner Grundschulen, mit insgesamt 487 Klassen, zu ihrer Umsetzung von JÜL befragt. In der Faxabfrage wurden die Schulen gebeten anzugeben, in welchem Umfang in den Klassen und Lerngruppen der aktuellen 3. Jahrgangsstufe (2012/13) ihrer Schule im aktuellen sowie vergangenen Schuljahr (SAPH) in den Fächern Deutsch und Mathematik jahrgangsübergreifend unterrichtet wurde. Der genaue Wortlaut der Abfrage ist in Abbildung 3 ersichtlich.

Bitte markieren Sie die vier Skalen entsprechend für Ihre Klasse: **3a**

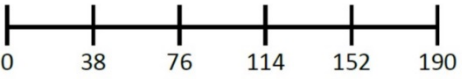
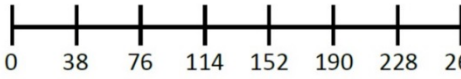
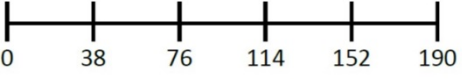
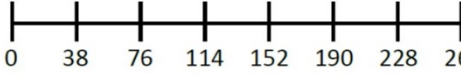
<p>Wie viele Unterrichtsstunden des Mathematikunterrichts wurde in dieser Klasse/Lerngruppe im Schuljahr 2011/12 jahrgangsübergreifend unterrichtet ...</p>  <p>Stunden im Schuljahr</p>	<p>Wie viele Unterrichtsstunden des Deutschunterrichts wurde in dieser Klasse/Lerngruppe im Schuljahr 2011/12 jahrgangsübergreifend unterrichtet ...</p>  <p>Stunden im Schuljahr</p>
<p>... und wie viele im Schuljahr 2012/13?</p>  <p>Stunden im Schuljahr</p>	<p>... und wie viele im Schuljahr 2012/13?</p>  <p>Stunden im Schuljahr</p>

Abbildung 3: Ausschnitt der Faxabfrage für Umsetzung von JÜL

Als Kriterien für die Ziehung der Stichprobe wurden die Stratifizierungsmerkmale Bezirk und Ergebnisse der Vergleichsarbeiten am Ende der 3. Jahrgangsstufe im Schuljahr 2012/13 (in Quartilen) herangezogen, um die Grundgesamtheit aller öffentlichen Berliner Grundschulen möglichst repräsentativ abzubilden. Die Population aller öffentlichen Berliner Grundschulen setzt sich im Schuljahr 2012/13 aus 374 Schulen mit insgesamt ca. 1.480 Klassen der 3. Jahrgangsstufe zusammen (Tabelle 5). Die Zusatzstudie fiel unter § 9 des Berliner Schulgesetzes (SenBWF, 2004). Die angeschriebenen Schulen waren somit zur

Studienteilnahme verpflichtet. Dennoch nahmen von den 123 ausgewählten Schulen drei Schulen nicht an der Befragung teil und drei weitere Schulen machten keine, nicht lesbare oder offensichtlich falsche Angaben und wurden deshalb aus den Analysen ausgeschlossen. So standen für die folgenden Auswertungen Daten von 117 öffentlichen Berliner Grundschulen mit insgesamt 460 Klassen zur Verfügung, was einer Rücklaufquote von rund 95 % entspricht. Tabelle 6 fasst die wesentlichen Merkmale der Erweiterungsstudie zusammen.

Tabelle 6: Übersicht zur Erweiterungsstudie

Schuljahr	N _{Schulen}	N _{Klassen}	Stunden im Schuljahr	M	SD
Mathe 2011/12	117	460	190	138,1	69,5
Mathe 2012/13	117	458	190	86,5	87,8
Deutsch 2011/12	117	458	266	193,1	93,3
Deutsch 2012/13	117	458	266	123,3	121,8

4.1.3.2 Stichprobe

Für die Beantwortung der Fragen I.1 bis I.3 werden zunächst die beschriebenen Daten der Schuljahre 2003/04 bis 2015/16 der amtlichen Schulstatistik des Landes Berlin verwendet. Tabelle 5 (S. 72) fasst die Merkmale dieser Daten zusammen. Für die in Frage I.4 zu identifizierenden Profile oder Typen der JÜL-Implementation wird ein Längsschnitt verwendet, der auf Grundlage der schulstatistischen Daten erstellt wurde. Dieser Längsschnitt umfasst auf Jahrgangsebene aggregierte Angaben für jede Schule und in jedem Schuljahr zu Anzahl der Gesamtklassen sowie unterschiedlicher Klassenarten (jahrgangsbezogen, jahrgangsübergreifend, sonstige), Schüler*innenzahlen insgesamt und in den einzelnen Klassenarten, sowie die Verteilung der Schüler*innen nach Geschlecht und Herkunft. Da es bei der Identifikation der Profile vordergründig um die Beschreibung der Implementation vor dem Hintergrund der sich verändernden schulpolitischen Vorgaben (Ver-/Entpflichtung) geht, werden in dieser Untersuchung Angaben aus den Schuljahren 2007/08 bis 2015/16 berücksichtigt. Nicht berücksichtigt wurden Schulen, die in den ersten drei Messzeitpunkten (2007/08 bis 2010/11) oder in den letzten drei Messzeitpunkten (2012/13 bis 2015/16) vollständig fehlende Werte aufweisen, da hier angenommen wurde, dass diese Schulen entweder gerade erst gegründet wurden und somit zu Beginn der JÜL-Implementation noch gar nicht existierten oder diese Schulen wurden geschlossen und es konnte nicht mehr geprüft werden, ob sie JÜL langfristig umgesetzt hatten.

Für die Analysen I.5 und I.6 zur JÜL-Umsetzung in den einzelnen Jahrgangsstufen und den Fächern Deutsch und Mathematik wurde auf die beschriebenen Daten der Erweiterungsstudie zurückgegriffen (Abschnitt 4.1.3.1.2).

4.1.3.3 Analysestrategie

In Anerkennung der Relevanz praktischer Bedeutsamkeit empirischer Befunde einerseits und der Herausforderungen in Zusammenhang mit Signifikanzprüfungen andererseits (vgl. Cumming, 2012) stellen in den Analysen der vorliegenden Dissertation Effektgrößen und Konfidenzintervalle einen wesentlichen Schwerpunkt dar. In Kombination mit signifikanzprüfenden Verfahren soll somit sowohl der empirische als auch praktische Erkenntnisgewinn gesichert werden.

Der erste Teil der Forschungsfragen zu Studie I (Fragen I.1 bis I.3), der das Beschreiben der Implementation von JÜL auf Schul- und Klassenebene zum Ziel hat, wird mithilfe der Verteilungen von Häufigkeiten und Anteilsberechnungen beantwortet. Die praktische Bedeutsamkeit möglicher Verschiebungen in den Verteilungen (z. B. eine starke Zunahme der Anteile an JÜL-Klassen mit Beginn des Verpflichtungszeitraums im Vergleich zum Schuljahr davor) kann mit der Effektstärke Cohens h (Cohen, 1988, 179 ff.) überprüft werden. Dabei wird ab einem Wert von $h \geq .20$ von einem kleinen, ab $h \geq .50$ von einem mittleren und bei $h \geq .80$ von einem großen Effekt ausgegangen (ebd.).

Für die Prüfung möglicher Profile oder Typen der Implementation (I.4) werden latente Profilanalysen in Mplus (Version 7.3, Muthén & Muthén, 1998-2014) durchgeführt. Dieses Verfahren ist eine besondere Art der latenten Klassenanalyse, einer Methode zur Identifikation von a priori unbekanntem Subpopulationen (latenten Klassen) in den Zusammenhängen zwischen den Items oder Indikatoren (Geiser, 2010). Es handelt sich um eine Methode zur Datenreduktion (ebd.), die aufgrund ihres latenten Ansatzes als messfehlerbereinigt angenommen wird (Collins & Lanza, 2010). Hierbei geht die Ausprägung einer oder mehrerer (kontinuierlicher) Indikatorvariablen zu jedem Messzeitpunkt in die Schätzung der latenten Klassen ein. Als Indikator in der vorliegenden Studie wurde der Anteil an jahrgangsübergreifend unterrichteten Klassen an der Gesamtklassenanzahl der Schulanfangsphase je Schuljahr verwendet: Ein Anteil von 0 bedeutet demnach, dass keine SAPH-Klasse an einer bestimmten Schule im entsprechenden Schuljahr jahrgangsübergreifend unterrichtet wurde, ein Anteil von 1 bedeutet, dass alle SAPH-Klassen einer Schule in einem

Schuljahr jahrgangsübergreifend unterrichtet wurden. Mit diesem quantitativen Indikator zu JÜL kann somit ein wichtiger Aspekt der Unterrichtsorganisation erfasst werden. Dem Vorgehen von Collins und Lanza (2010) folgend, wurde die Anzahl der latenten Klassen schrittweise erhöht und jedes Modell jeweils gegen das Modell getestet, das eine latente Klasse weniger hat. Für die Auswahl des Modells mit dem besten Fit wurden, der einschlägigen Literatur folgend (Collins & Lanza, 2010; Nylund, Asparouhov & Muthén, 2007), folgende Kennwerte herangezogen: *Akaike's Information Criterion* (AIC; möglichst klein), (*sample size adjusted*) *Bayesian Information Criterion* (SABIC/BIC; möglichst klein); *Entropy* (ein Gütemaß für die Sicherheit, mit der Schulen den latenten Profilen zugeordnet werden können; sollte möglichst nahe an 1 sein), *Lo-Mendell-Rubin likelihood ratio test* (LMR-LRT; wenn nicht signifikant, ist das Modell mit der geringeren Klassenzahl zu bevorzugen) und *bootstrap likelihood ratio test* (BLRT; wenn nicht signifikant, ist das Modell mit der geringeren Klassenzahl zu bevorzugen). Zudem wurden theoretische Überlegungen, vor allem unter Berücksichtigung des Zeitraums zur verpflichtenden Umsetzung von JÜL und bereits existierenden Schulversuchsschulen, bei der Modellauswahl berücksichtigt.

Im letzten Abschnitt von Studie I steht der Umfang der JÜL-Umsetzung in den Jahrgangsstufen 2 und 3 sowie den Fächern Deutsch und Mathematik im Zentrum. Zunächst wird mithilfe der Anteile der jahrgangsübergreifend unterrichteten Unterrichtsstunden an der Gesamtzahl der Unterrichtsstunden in einem Schuljahr die Verteilung des Umfangs beschrieben (I.5). Anschließend werden die Unterschiede bzw. Gemeinsamkeiten in der JÜL-Umsetzung in den Jahrgangsstufen 2 und 3 sowie den Fächern Deutsch und Mathematik (I.6) mittels bivariater Korrelationen überprüft. Zur Einschätzung der statistischen und praktischen Bedeutsamkeit der Unterschiede in der Umsetzung werden darüber hinaus t-Tests für abhängige Stichproben (vgl. Eid, Gollwitzer & Schmitt, 2015) durchgeführt und die jeweilige Effektstärke Cohens *d* berechnet (Cohen, 1988).

4.1.4 Ergebnisse

Der folgende Abschnitt stellt die Ergebnisse der Studie I dar. Die Darstellung erfolgt entlang der in Abschnitt 4.2.3 formulierten Forschungsfragen.

I.1 An wie vielen Berliner Grundschulen und in wie vielen Klassen wurde JÜL in den Schuljahren 2003/04 bis 2015/16 umgesetzt?

Zur Beantwortung der formulierten Forschungsfragen wurde größtenteils deskriptiv vorgegangen (siehe Abschnitt 4.2.4.3). Tabelle 7 stellt für die Schuljahre 2003/04 bis 2015/16 dar, an wie vielen Grundschulen Berlins mindestens eine SAPH-Klasse jahrgangsübergreifend organisiert wurde. Weiterhin ist der Anteil der JÜL-Klassen an der Gesamtklassenanzahl der SAPH beschrieben. In der gesamten Tabelle wurden sowohl Klassen berücksichtigt, die die erste und zweite Jahrgangsstufe übergreifend organisierten (JÜL-2), was den Mindestanforderungen der Senatsbildungsverwaltung und der „typischen“ Schulanfangsphase entsprach, als auch Klassen, die die Jahrgänge eins bis drei zusammenfassten (JÜL-3) und somit über die Mindestanforderungen hinausgingen.

Tabelle 7: Anteil der JÜL-Klassen in Berliner Schulen und Klassen

Schuljahr	N _{Schulen}	Schulen _{JÜL} ¹	N _{Klassen} ²	Klassen _{JÜL} ³	Unterrichtsorganisation
2003/04	381	2,1%	2.178	4,4%	Jahrgangsklassen sind die Regel, JÜL auf Antrag möglich
2004/05	376	2,1%	2.198	4,3%	
2005/06	373	2,1%	2.489	4,9%	
2006/07	371	12,4%	2.478	9,9%	
2007/08	370	45,1%	2.260	42,4%	
2008/09	371	67,1%	2.267	63,2%	Zeitraum der verpflichtenden JÜL-Umsetzung
2009/10	373	81,5%	2.309	75,4%	
2010/11	377	87,0%	2.439	79,9%	
2011/12	379	85,8%	2.507	78,9%	
2012/13	378	77,0%	2.552	70,3%	Sowohl Jahrgangs- als auch JÜL-Klassen möglich
2013/14	378	63,5%	2.614	57,0%	
2014/15	383	59,3%	2.704	53,0%	
2015/16	383	58,0%	2.770	51,8%	

¹ Anteil der Schulen mit mindestens einer JÜL-Klasse in der SAPH

² Schulen hatten die Möglichkeit, über die ersten beiden Jahrgangsstufen (JÜL-2) hinaus auch die SAPH und die Jahrgangsstufe 3 zu einer Lerngruppe/Klasse (JÜL-3) zusammenfassen. Diese Klassen wurden daher ebenfalls in die Analysen einbezogen.

³ Anteil der JÜL-Klassen an der Gesamtzahl der Klassen der Schulanfangsphase; einschließlich Klassen mit JÜL-3

Die Ergebnisse zeigen erwartungsgemäß auf, dass es bereits vor der verpflichtenden Einführung von JÜL Schulen gab, nämlich solche aus den Modellversuchen (Abgeordnetenhaus Berlin, 2003; 2013), die mindestens einen Teil ihrer ersten beiden (oder drei) Jahrgangsstufen übergreifend unterrichteten. Zu erkennen ist weiterhin, dass der Anteil an Schulen mit JÜL-Klassen und der Anteil der JÜL-Klassen insgesamt bereits im Schuljahr 2007/08, also dem Jahr, in dem die verpflichtende Umsetzung ursprünglich beginnen sollten, deutlich ansteigt: von 12,4 % der Schulen und 9,9 % der Klassen im Schuljahr davor auf 45,1 % bzw. 42,4 %. Zum Schuljahr 2008/09, dem eigentlichen Beginn des Verpflichtungszeitraums, steigen die Anteile

auf circa zwei Drittel der Schulen bzw. Klassen (67,1 % bzw. 63,2 %) und erreichen im Schuljahr 2011/12 ihren Höchststand. Das bedeutet, dass maximal in 87 % der Schulen zumindest eine Klasse und gesamt maximal 79,9 % aller Klassen jahrgangsübergreifend organisiert wurde/n. Danach wurden beide Anteile sukzessive kleiner, sodass zum letzten Messzeitpunkt, dem Schuljahr 2015/16, JÜL noch in 58 % der Schulen bzw. 51,8 % der Klassen praktiziert wurde. Interessant ist außerdem das Schuljahr 2006/07: hier wussten die Schulen bereits, dass sie zukünftig ihre Klassen der Schulanfangsphase verpflichtend jahrgangsübergreifend organisieren sollten. Immerhin 10 % schienen den Vorlauf nutzen zu wollen und begannen, zumindest teilweise, bereits mit der Umsetzung von JÜL.

I.2 Wie viele Klassen der Schulanfangsphase innerhalb der Schulen wurden über die Zeit jahrgangsübergreifend organisiert?

Der Annahme folgend, dass JÜL nur dann erfolgreich implementiert werden kann, wenn es in mehr als einer Klasse umgesetzt wird (Carle & Metzen, 2014), stellt die folgende Abbildung (Abbildung 4) die Anteile der Schulen, die JÜL in einer, in mehr als einer, jedoch nicht allen oder in allen ihren SAPH-Klassen umsetzten (Frage I.2). Es wurden also nur die Schulen berücksichtigt, die in ihrer SAPH mindestens eine Klasse jahrgangsübergreifend organisiert haben. Dabei zeigt sich, dass, mit Ausnahme des Schuljahres 2006/07, stets über 60 % der JÜL-Schulen diese Form der Unterrichtsorganisation in all ihren Klassen umgesetzt haben. Dieser Anteil stieg während des Verpflichtungszeitraums auf bis zu 81 % der Schulen und scheint sich nach der Aufhebung der Verpflichtung auf 65 % bis 70 % einzupendeln. Im Gegensatz dazu richtete lediglich im Laufe des Schulversuchs zu JÜL eine vergleichsweise größere Anzahl an Schulen (2003/04: 13 %) eine einzelne jahrgangsübergreifende Lerngruppe ein. In den folgenden Schuljahren lag dieser Anteil zum allergrößten Teil zwischen 0 % und 1 %. Der

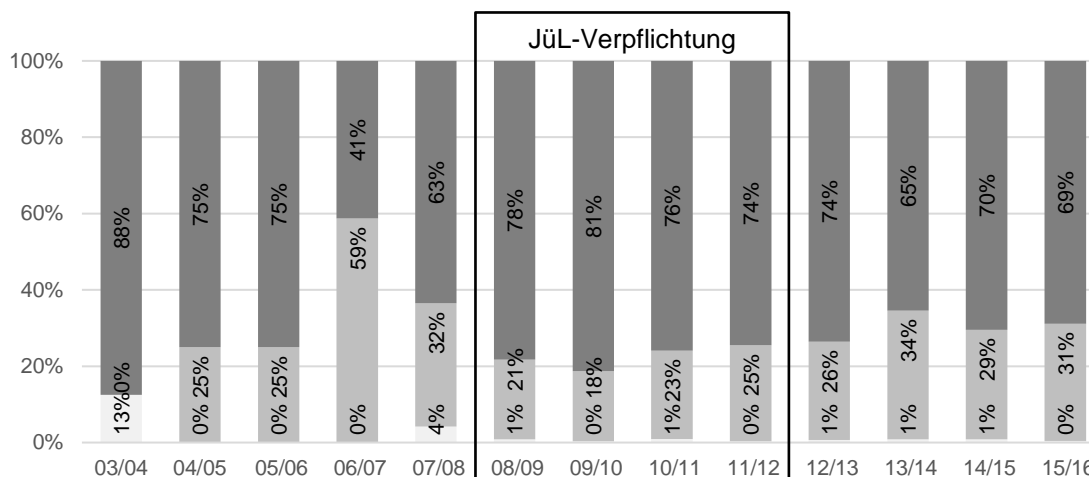


Abbildung 4: Verteilung der Schulen nach der Anzahl ihrer JüL-Klassen

Anteil an Schulen, die mehr als eine, aber nicht alle ihrer SAPH-Klassen jahrgangsübergreifend organisierten bewegt sich in den meisten Schuljahren um die 25 % bis 30 %. Auffällig ist hier erneut das Schuljahr (2006/07), in dem sich die verpflichtende Einführung von JÜL bereits abzeichnete und einige Schulen JÜL zumindest bereits teilweise einführten.

I.3 Wie verteilen sich JÜL-2 und JÜL-3 in Berliner Grundschulen über die Jahre?

Anschließend wurde untersucht, inwieweit Schulen, die mindestens eine JÜL-Klasse in der Schulanfangsphase eingerichtet haben, zwei (JÜL-2) oder drei (JÜL-3) Jahrgangsstufen in diesen Klassen zusammen unterrichten. Abbildung 5 bildet die Verteilung von JÜL-2 und JÜL-3 Klassen über die Zeit hinweg ab. In den Schuljahren 2003/04 bis 2005/06, in denen JÜL nur im Rahmen des Schulversuchs umgesetzt wurde, haben alle am Schulversuch teilnehmenden Schulen in ihren JÜL-Klassen drei Jahrgangsstufen gemeinsam unterrichtet. Ab dem Zeitpunkt, ab dem die verpflichtende Einführung an allen Schulen absehbar war (2006/07), steigt der Anteil der JÜL-2-Klassen von 0 % auf 67 %. In den nachfolgenden Schuljahren ist eine leichte Verschiebung der Verteilung zugunsten von JÜL-3 auf 59 % im Schuljahr 2015/16 zu erkennen. Mithilfe der Effektgröße Cohens h (Cohen, 1988) lässt sich seit Beginn der verpflichtenden Umsetzung von JÜL eine Verschiebung von 69 % im Jahr 2008/09 auf 59 % in 2015/16 mit $h = .21$ als kleiner Effekt einschätzen.

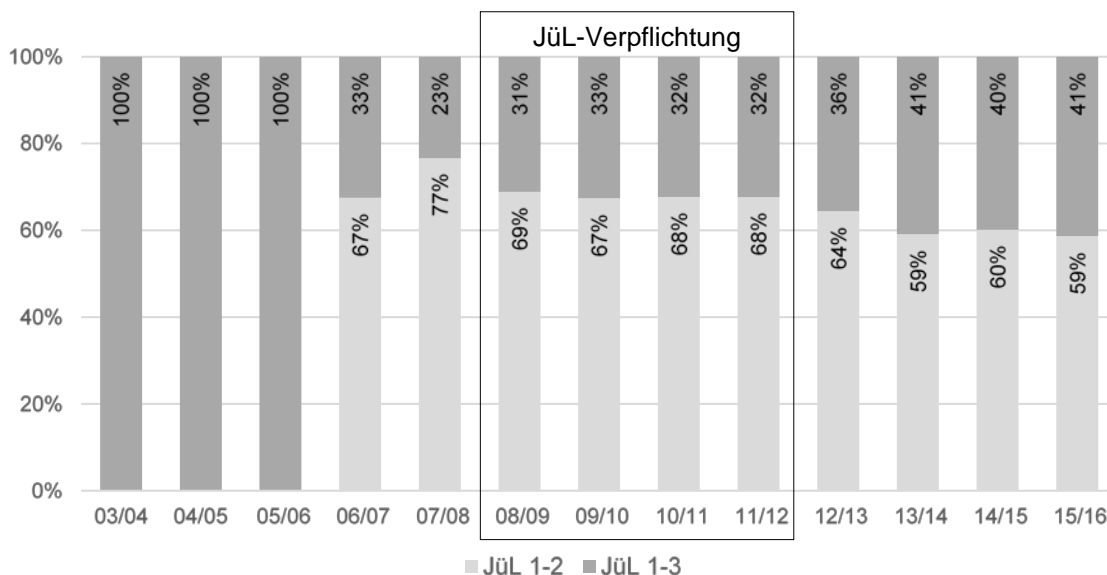


Abbildung 5: Verteilung der JÜL-2 und JÜL-3-Klassen innerhalb der Schulen, die JÜL in mindestens einer ihrer Klassen der Schulanfangsphase umsetzen.

1.4 Können Typen der JÜL-Implementation identifiziert werden und, falls ja, wie können diese beschrieben werden?

Da die bisherigen Untersuchungen lediglich über die Schulen hinweg aufzeigen konnten, wie JÜL implementiert wurde, soll im Folgenden versucht werden, den Verlauf der JÜL-Implementation in den einzelnen Schulen zu beschreiben. In Anlehnung an die Diffusionstheorie von Rogers (2003), wonach Schulen verschiedenen Typen der Implementation zugeordnet werden können, wurden zu diesem Zweck latente Profilanalysen durchgeführt, womit eine Beschreibung der Implementationsverläufe mithilfe von Profilen ermöglicht wird. Als Indikator zur Bestimmung der latenten Profile wurde jeweils der Anteil der JÜL-Klassen an der SAPH-Gesamtklassenanzahl für die Jahre 2007/08 bis 2015/16 verwendet. Zur Erinnerung: zur Komplexitätsreduktion und Effizienzsteigerung des Verfahrens wurden die Schuljahre, die vor dem (geplanten und tatsächlichen) Beginn des Verpflichtungszeitraums lagen (2003/04 bis 2006/07), in diesen Analysen nicht berücksichtigt.

Die Fit-Werte der unterschiedlichen latenten Klassenlösungen sind in Tabelle 8 dargestellt.

Tabelle 8: Fit-Statistiken der latenten Profilanalyse zur JÜL-Implementation

k	LL	SCF	#fp	AIC	BIC	SABIC	p LMR-LRT	p BLRT	Entropy
2	-15357,1	1,50	28	30770,3	30878,8	30790,0	,001	,001	,993
3	-14592,4	1,63	38	29260,7	29408,1	29287,5	,000	,001	,993
4	-14318,3	1,75	48	28732,6	28918,7	28766,5	,001	,001	,995
5	-14320,3	1,69	58	28756,7	28981,6	28797,6	,024	,026	,976
6	-14058,8	1,63	68	28253,5	28517,2	28301,5	,004	,004	,980
7	-13921,9	1,95	78	27999,9	28302,4	28054,9	,101	,106	,993
8	-13870,2	1,88	88	27916,4	28257,6	27978,4	,251	,258	,995

k: Anzahl latenter Profile im Modell; LL: log likelihood; SCF: scaling correction factor for MLR; #fp: number of free parameters; AIC: Akaike's Information Criterion; SABIC/BIC: (sample size adjusted) Bayesian Information Criterion; Entropy; LMR-LRT: Lo-Mendell-Rubin likelihood ratio test; BLRT: bootstrap likelihood ratio test; das ausgewählte Modell ist fett gedruckt.

Ausgehend von diesen Fit-Statistiken wurde ein Modell gewählt, das sechs Typen der JÜL-Implementation für die SAPH an Berliner Grundschulen unterscheidet: Obwohl die Entropy-Werte für einige alternative Modelle näher bei 1 lagen, spricht für dieses Modell, dass die Werte für AIC, BIC und SABIC jeweils kleiner waren als korrespondierende Werte für Modelle mit weniger latenten Profilen ($2 \leq k \leq 5$). Zudem waren der LMR-LRT- sowie der BLRT-Test für sukzessive Modellvergleiche bis zur Lösung mit sechs Profilen statistisch signifikant, was dieses Modell empirisch stützte. Die Modelle mit mehr latenten Profilen ($k \geq 7$) wiesen

vergleichsweise zwar jeweils nochmals kleinere Werte für AIC, BIC und SABIC auf. Jedoch zeigten der LMR-LRT- sowie der pBLRT-Test keine statistisch signifikante Verbesserung gegenüber dem Modell mit sechs latenten Profilen auf, was letztlich ausschlagend für Entscheidung für dieses Modell war. Im Anschluss wurde die empirische Entscheidung entlang der theoretischen Vorüberlegungen geprüft und bestätigt.

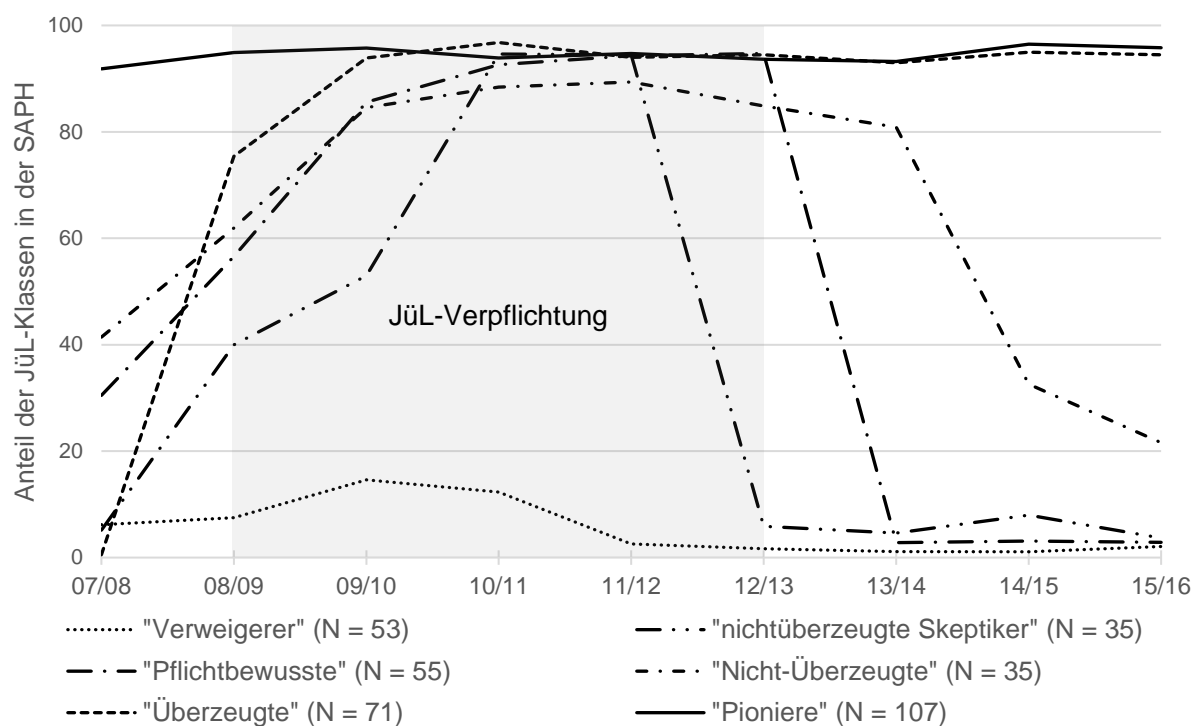


Abbildung 6: Typen der JÜL-Implementation²²

Die sechs Profiltypen lassen sich wie folgt beschreiben (Abbildung 6): Dem Typ „Verweigerer“ gehören die Schulen an, die zu keinen der neun Messzeitpunkte in mehr als 18 % ihrer Klassen (i. d. R. nicht mehr als eine Klasse) der Schulanfangsphase angegeben haben, dass diese jahrgangsübergreifend organisiert wird. Sie machen 15 % der untersuchten 356 Schulen aus. Die übrigen 85 % der Schulen haben zu mindestens zwei der neun Messzeitpunkte den Großteil der untersuchten Klassen jahrgangsübergreifend organisiert. Die „Pflichtbewussten“ (ca. 15 %) stellen dabei den Typus der Schulen dar, die mit der verpflichtenden Einführung von JÜL mit der Umsetzung begonnen haben und, sobald sie die offizielle Möglichkeit zur jahrgangsbezogenen Organisation erhielten (2012/13), auch zu dieser Form der Unterrichtsorganisation zurückkehrten. Insgesamt hat jede dritte Schule in Berlin zwar mit der

²² © Thoren, K. & Brunner, M. eingereicht zur Publikation in Zeitschrift für Erziehungswissenschaft, 2017, abgedruckt mit freundlicher Genehmigung von Springer Nature

Umsetzung von JÜL begonnen, ist jedoch bis spätestens zum Schuljahr 2015/16 zum jahrgangsbezogenen Lernen zurückgekehrt. Auf der anderen Seite haben 20 % der Schulen JÜL mit Beginn der Verpflichtung eingeführt und sind bis zum Schuljahr 2015/16 auch dabei geblieben („Überzeugte“). Weitere 30 % haben bereits vor Beginn der Verpflichtung ihre Schulanfangsphase jahrgangsübergreifend organisiert und sind ebenfalls bisher dabei geblieben („Pioniere“), sodass JÜL an der Hälfte der untersuchten Schulen, zumindest als Form der Unterrichtsorganisation, dauerhaft (Coburn, 2003) umgesetzt wird. Da diese Schulen JÜL auch nach der Aufhebung der Verpflichtung weiter umsetzen, kann angenommen werden, dass diese Schulen sich mit der Reform identifizieren (*shift in reform ownership*, Coburn, 2003).

Der zweite Teil der in dieser Studie untersuchten Implementation von JÜL beschäftigt sich mit der Frage, in welchem Umfang JÜL an den Schulen implementiert wurde. Dabei wird geprüft, in welchem Stundenumfang an den Schulen jahrgangsübergreifend unterrichtet wurde (Frage I.5) und wie sich die Umsetzung in verschiedenen Unterrichtsfächern unterscheiden (Frage I.6).

I.5 In welchem Stundenumfang wurde in Berliner Grundschulen jahrgangsübergreifend unterrichtet?

Die Darstellung der Verteilungen zeigt sowohl für das Fach Deutsch (Abbildung 7) als auch für das Fach Mathematik (Abbildung 8), dass in den Schuljahren 2011/12 (2. Jahrgangsstufe) und 2012/13 (3. Jahrgangsstufe) ein Großteil der Schulen JÜL entweder in keiner oder in allen Unterrichtsstunden umgesetzt hat. Erwartungsgemäß wurde in der zweiten Jahrgangsstufe in den meisten Unterrichtsstunden jahrgangsübergreifend unterrichtet. In der dritten Jahrgangsstufe waren es weit weniger Unterrichtsstunden, da die Schulen zwar drei Jahrgänge zu einer JÜL-Klasse zusammenfassen konnten, aber zu keinem Zeitpunkt mussten. Verpflichtend war lediglich das Zusammenfassen zweier Jahrgänge.

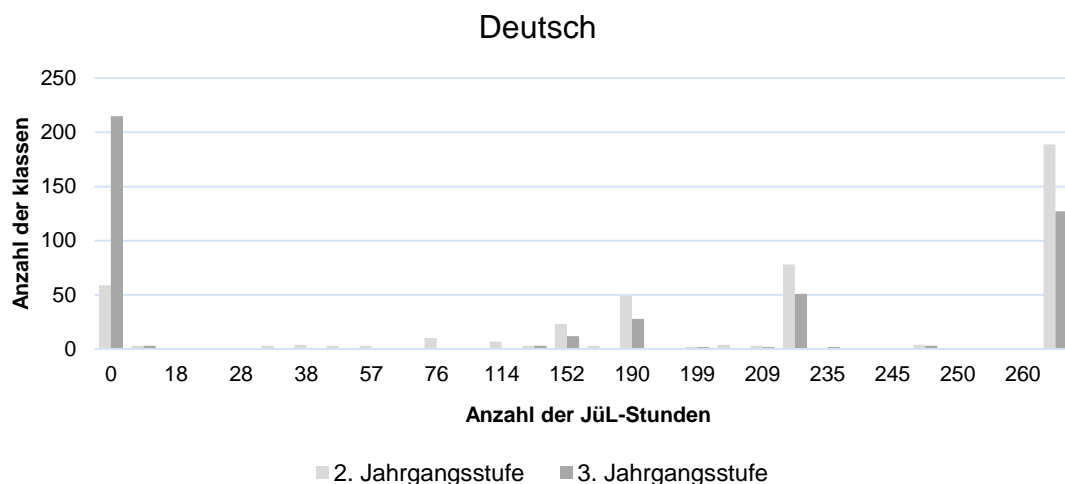


Abbildung 7: Anteile der jahrgangsübergreifend unterrichteten Stunden an den Jahresunterrichtsstunden für das Fach Deutsch

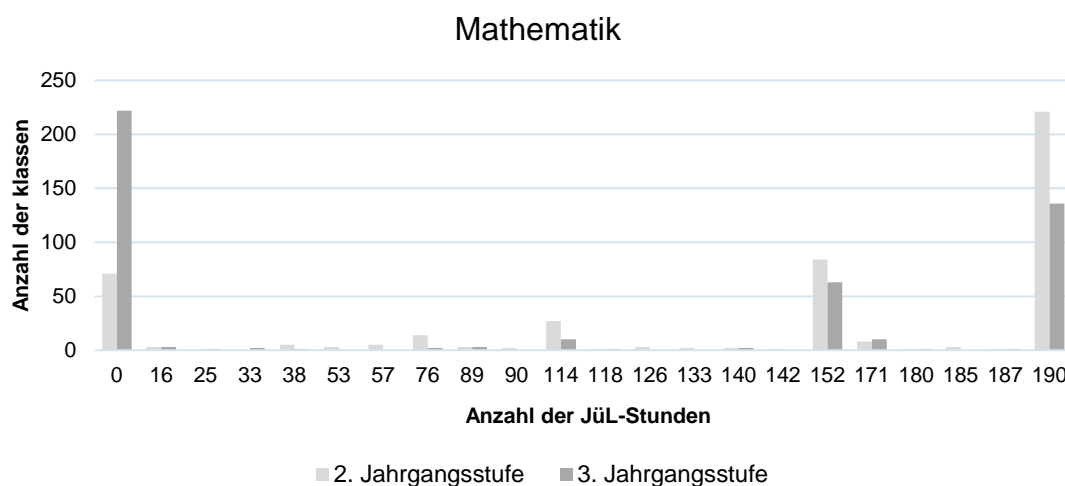


Abbildung 8: Anteile der jahrgangsübergreifend unterrichteten Stunden an den Jahresunterrichtsstunden für das Fach Mathematik

Der Vergleich der Mittelwerte innerhalb der Fächer und über die Jahrgangsstufen hinweg mittels t-Tests für abhängige Stichproben bestätigt, dass in Jahrgangsstufe 3 im Mittel eine signifikant geringere Anzahl der Unterrichtsstunden jahrgangsübergreifend organisiert wurde (Tabelle 9). Cohens d zeigt dabei $d = ,64$ (Deutsch) bzw. $d = ,65$ (Mathematik) an (Cohen, 1988).

Tabelle 9: t-Tests für abhängige Stichproben über die Jahrgangsstufen 2 und 3

	2. Jahrgangsstufe		3. Jahrgangsstufe		df	t	p	d
	M ¹	SD	M ¹	SD				
Deutsch	193,08	93,34	123,33	121,82	450	14,37	,00	,64
Mathematik	138,05	69,50	86,52	87,76	453	13,91	,00	,65

¹ Anzahl der jahrgangsübergreifend unterrichteten Stunden an der Gesamtzahl der Jahresunterrichtsstunden je Fach

I.6 Wie unterscheidet sich die JÜL-Umsetzung bezogen auf die Jahresunterrichtsstunden in verschiedenen Unterrichtsfächern?

Die Ergebnisse zur Prüfung von Forschungsfrage I.6 werden im Folgenden dargestellt (Abbildung 9). Die Punktediagramme, in denen die jahrgangsübergreifend durchgeführten Unterrichtsstunden in den Fächern Deutsch und Mathematik einander gegenübergestellt sind, zeigen sowohl für die Jahrgangsstufe 2 als auch für die Jahrgangsstufe 3, dass ein Großteil der untersuchten Schulen bezogen auf den Stundenumfang JÜL einheitlich für beide Fächer umsetzt. Die sehr hohen Korrelationskoeffizienten verdeutlichen diesen Befund.

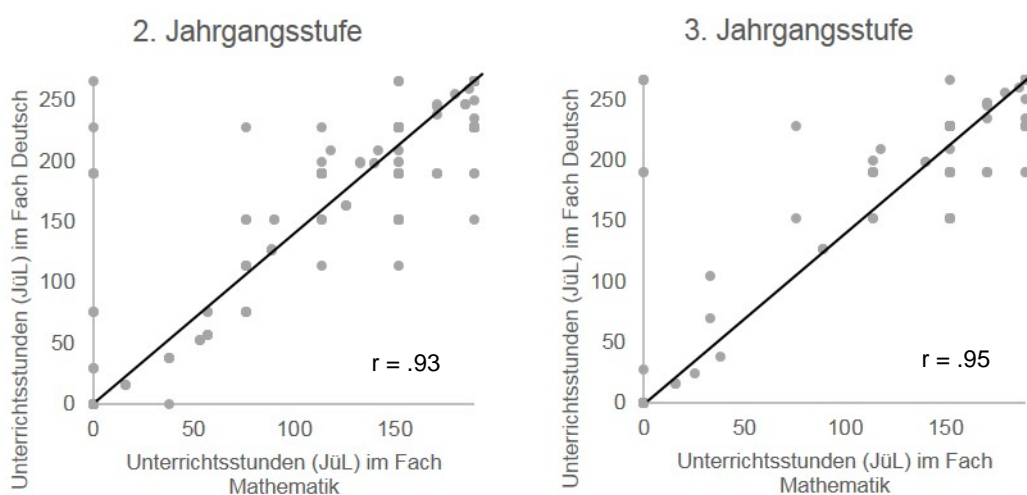


Abbildung 9: Streudiagramm zu den jahrgangsübergreifend unterrichteten Stunden in den Fächern Deutsch und Mathematik in der zweiten und dritten Jahrgangsstufe; Korrelationskoeffizient r

Auch die Gegenüberstellung der Anteile der mittleren JÜL-Unterrichtsstunden an den Gesamtjahresunterrichtsstunden jeweils für die Jahrgangsstufen 2 und 3 mittels t-Tests für abhängige Stichproben und Cohens d zeigt, dass keine statistisch bedeutsamen Unterschiede bezogen auf die Jahrgangsmischung zwischen den Fächern Deutsch und Mathematik, festgestellt werden konnten (Tabelle 10).

Tabelle 10: Übersicht zur den jahrgangsübergreifend unterrichteten Stunden

	Deutsch		Mathematik		df	t	p	d
	M% ¹	SD%	M% ¹	SD%				
2, Jahrgangsstufe	72,59	35,09	72,66	36,58	456	0,26	,79	,00
3, Jahrgangsstufe	46,37	45,80	45,54	46,19	457	1,46	,14	,02

¹ % Anteil der jahrgangsübergreifend unterrichteten Stunden an der Gesamtzahl der Jahresunterrichtsstunden je Fach

4.1.5 Zusammenfassung

Die Ergebnisse von Studie I zeigen auf, dass JÜL in der Berliner Schulanfangsphase bereits vor der verpflichtenden Einführung an einigen wenigen Schulen umgesetzt wurde. Alle dieser Schulen setzten JÜL im Rahmen des Schulversuchs in einem Großteil ihrer SAPH-Klassen um und fassten jeweils drei Jahrgangsstufen zu einer jahrgangsübergreifenden Lerngruppe zusammen. Mit Beginn der seitens der Bildungspolitik beschlossenen verpflichtenden Einführung von JÜL in allen Klassen des Schulanfangs (Schuljahr 2008/09) begann auch der größte Teil der Schulen, JÜL umzusetzen, jedoch gab es zu jeder Zeit Schulen, die entweder JÜL gar nicht einführten oder nicht in allen ihren SAPH-Klassen jahrgangsübergreifend unterrichteten. Mit Rücknahme der verpflichtenden JÜL-Umsetzung nahm auch die tatsächliche Umsetzung ab und es ist seitdem (Schuljahr 2012/13) ein sukzessiver Rückgang zu konstatieren. Mit Blick auf die Unterscheidung von JÜL-2 und JÜL-3, also ob in einer JÜL-Gruppe zwei oder, dem (reform-)pädagogischen Ideal folgend, drei Jahrgänge gemeinsam unterrichtet werden, wurde festgestellt, dass während des Verpflichtungszeitraums etwa ein Drittel JÜL-3 umsetzte. Seit dem Ende der Verpflichtung ist ein leicht steigender Trend der JÜL-3-Klassen zu verzeichnen. Bei dem allgemein geringer werdenden Anteil der Schulen und Klassen, die JÜL umsetzen, sind es also vor allem JÜL-2-Klassen, die zu einer jahrgangsbezogenen Organisation zurückkehren. Weiterhin gab es kaum Schulen, die JÜL lediglich in einer ihrer SAPH-Klassen umgesetzt haben, da es für Schulen nicht praktikabel zu sein scheint, eine einzelne JÜL-Klasse einzurichten.

Mit dem Ziel der Systematisierung der JÜL-Implementation wurde – der Diffusionstheorie von Rogers (2003) folgend – überprüft, ob sich Schulen nach Typen der Implementation unterscheiden lassen. Dabei konnten sechs Typen der JÜL-Implementation identifiziert werden. Diese Typen beschreiben Schulen, die zu keinem der neun untersuchten Messzeitpunkte in (mehr als) einer Klasse jahrgangsübergreifend unterrichtet haben („Verweigerer“), Schulen, die JÜL einerseits um das Schuljahr 2008/09 (Beginn des Verpflichtungszeitraums) im Großteil ihrer SAPH-Klassen umsetzte, jedoch andererseits zum Teil bereits vor Aufhebung der Verpflichtung wieder zu einer jahrgangsbezogenen Organisation zurückkehrten („nichtüberzeugte Skeptiker“, „Pflichtbewusste“ und „Nicht-Überzeugte“), sowie Schulen, die entweder bereits vor der verpflichtenden Einführung (z. B. im Rahmen der Schulversuche) jahrgangsübergreifend unterrichteten („Pioniere“) oder aber mit Beginn der Verpflichtung mit JÜL begannen („Überzeugte“) und bis zum letzten Messzeitpunkt auch in der Mehrzahl ihrer SAPH-Klassen dabei geblieben sind. Diese sechs Typen entsprechen nicht nur empirisch,

sondern vor allem auch theoretisch dem, was zu erwarten war und stellen eine fundierte Grundlage für weitere Analysen zur Implementation dar (siehe Studie II).

Der letzte Abschnitt der Studie I befasst sich damit, in welchem Stundenumfang an den Schulen in einem Schuljahr (2012/13) jahrgangsübergreifend unterrichtet wurde und wie sich die Umsetzungen in verschiedenen Unterrichtsfächern unterschieden. Hier konnte aufgezeigt werden, dass es zum einen eine Vielzahl an Klassen gab, in denen in der zweiten, nicht jedoch in der dritten Jahrgangsstufe jahrgangsübergreifend unterrichtet wurde, was die Befunde aus Frage I.3 bestätigt. Zum anderen konnte dargestellt werden, dass die überwiegende Mehrheit der Schulen, die für eine SAPH-Klasse angibt, diese jahrgangsübergreifend zu unterrichten, dies auch in jeder bis maximal jeder zweiten Unterrichtsstunde der Fächer Deutsch und Mathematik umsetzte. Weiterhin wurde innerhalb der Jahrgangsstufen zwei und drei gezeigt, dass die große Mehrheit der Schulen JÜL in beiden Fächern (Deutsch und Mathematik) in gleichem Umfang umsetzt. Befürchtungen, dass JÜL nur in einem geringen Anteil der Jahresunterrichtsstunden und in Fächern, die nicht zu den Hauptfächern zählen, umgesetzt wurde, konnten somit entkräftet werden.

4.2 Studie II: Unterscheiden sich Schulen entsprechend ihrer Zuordnung zu einem Implementationstypen in ihrer Schul- und Unterrichtsqualität?²³

Carle und Metzen (2014) haben in ihrer Zusammenstellung des Forschungsstandes zu JÜL auf die Bedeutung verschiedener Wirkbedingungen für die erfolgreiche Umsetzung von JÜL hingewiesen (Abschnitt 4.1.1). Wesentliche Teile dieser Wirkbedingungen beschreiben Merkmale der Schul- und Unterrichtsqualität, die in diesem Zusammenhang ein wichtiges Kriterium zu sein scheinen. Studie II nimmt daher die Qualität auf diesen beiden Ebenen in den Blick. Nach einer theoretischen Einführung (Abschnitt 4.2.1) werden die Forschungsfragen abgeleitet (Abschnitt 4.2.2) und Methoden (Abschnitt 4.2.3) und Ergebnisse (Abschnitt 4.2.4) dargestellt. Abschnitt 4.2.5 fasst die wichtigsten Erkenntnisse aus Studie II zusammen.

4.2.1 Theoretischer Hintergrund

„Für die Förderung der Lern- und Bildungsprozesse von Kindern [ist] nicht primär die Organisationsform des Unterrichts, sondern insbesondere dessen pädagogisch-fachdidaktische Gestaltung entscheidend“ (Hanke, 2008, S. 22). Auch die Ausführungen von Carle & Metzen (2014), die in Abschnitt 4.1.1 zusammengefasst sind, verweisen auf die besondere Relevanz von Schul- und Unterrichtsqualität bei der Einführung und Umsetzung von JÜL. Wendt (2007) formuliert vor diesem Hintergrund fünf Maxime zu Schulentwicklung, Strukturen und Qualität bei der Einführung von JÜL: (1) Berücksichtigung der kognitiven und emotionalen Aspekte des Übergangs sowie Transparenz der geplanten Organisationsschritte; (2) Einführung als Schul- und Unterrichtsentwicklung im Team; (3) Umstrukturierung der Unterrichtsmethodik im Konsens des Kollegiums und nach gemeinsam vereinbarten Grundsätzen; (4) Einfordern von Supervision und Coaching, u. a. für Persönlichkeitsarbeit, Teambildung, Unterrichts- und Organisationsveränderungen; (5) interne und externe Evaluation.

Auch Ditton (2000, 2002) hebt in seinem Strukturmodell zur Evaluation und Qualitätssicherung im Bildungsbereich (Abbildung 10) in Anlehnung an das Modell von Scheerens und Bosker (1997) die Qualität der einzelnen Bildungseinrichtungen und die Qualität der Lehr- und Lernsituation als primäre Merkmale, die einen direkten Einfluss auf (langfristige) Wirkungen haben, hervor.

²³ Teile der Studie I wurden gemeinsam mit dieser Studie als Fachartikel in *der Zeitschrift für Erziehungswissenschaft* eingereicht. Dieser befindet sich derzeit unter folgendem Titel in Begutachtung: Thoren, K. & Brunner, M. (eingereicht). Implementation des Jahrgangsübergreifenden Lernens in der Berliner Schulanfangsphase: Langfristige Umsetzung = Hohe Qualität?

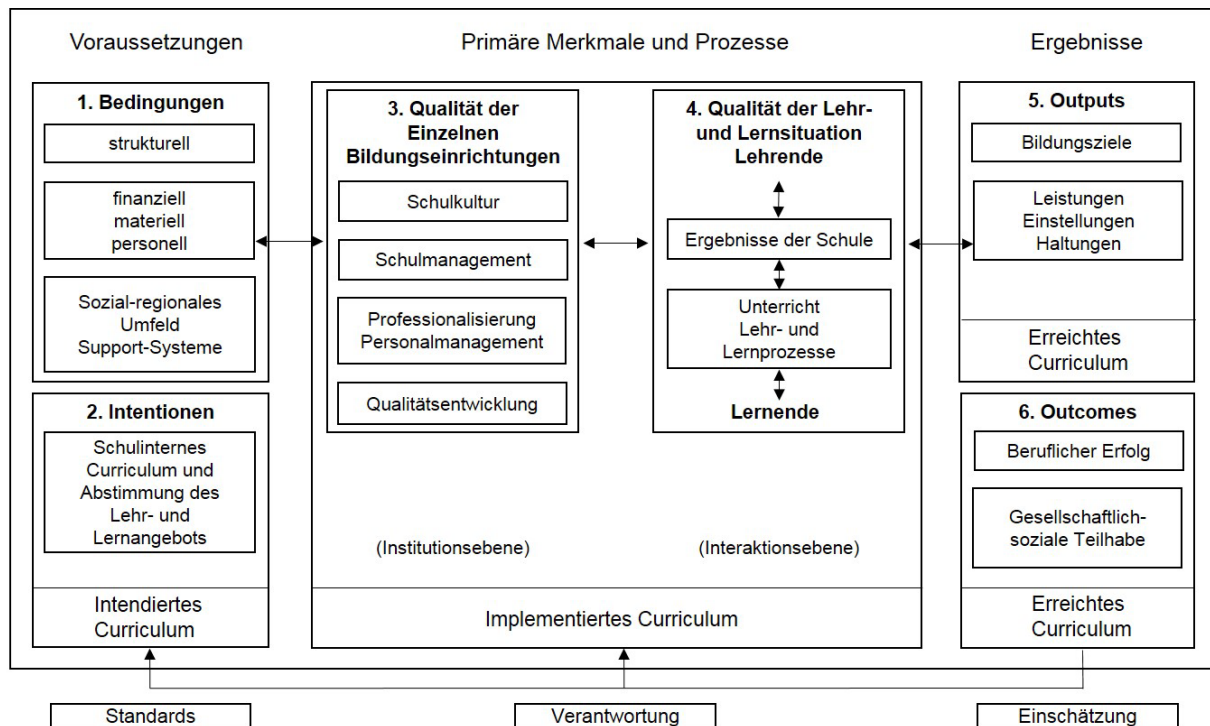


Abbildung 10: Strukturmodell zu Evaluation und Qualitätssicherung im Bildungswesen (Ditton, 2000; 2002)

In der Implementationsforschung wurden auf Schulebene die Rolle der Schulleitung (Gräsel & Parchmann, 2004) und der Schulkultur (Holtappels, 2005) als bedeutsame Prädiktoren für den Implementationserfolg empirisch belegt (Abschnitt 1.4). Eine besondere Rolle kommt hier auch den Lehrkräften zu, die sowohl auf Schul- als auch auf Unterrichtsebene agieren und deren Akzeptanz (Hall & Hord, 2015) und Qualifikation (Odden, 1991) eine Voraussetzung für erfolgreiche Implementation ist. Zentrale Kriterien der Unterrichtsqualität, speziell für die Bewertung der Umsetzung der Schulanfangsphase und des Jahrgangsübergreifenden Lernens, identifizieren Carle und Metzen (2008b) für die Evaluation des FLEX-Modellvorhabens in Brandenburg: Lernatmosphäre, sozialer Umgang, Kooperation der Kinder, sinnvolle Aufgaben, Adaptivität, Elementarisierung, Strukturiertheit, Effizienz, Leistungssog, Lernumgebung, Unterrichtsabläufe, Unterrichtsformate, Steuerungselemente und diagnostisch fundierte Unterrichtsplanung. Die Ergebnisse ihrer Videostudie zeigten, dass für eine erfahrene FLEX-Klasse besonders hohe Bewertungen der Lernatmosphäre und des sozialen Umgangs (vgl. ebd. S. 131) festgestellt werden konnten. Vergleichsweise niedriger bewertet wurde die Kooperation der Kinder, die sehr von den gestellten Aufgaben abhing. Insgesamt wurden in dieser Klasse jedoch die meisten Aspekte sehr gut bewertet. Im Gegensatz dazu zeigte eine wenig erfahrene FLEX-Klasse noch großes Entwicklungspotenzial. So fielen hier, abgesehen von den Aspekten sinnvolle Aufgaben und Elementarisierung, alle Bewertungen schlechter aus als in der

erfahrenen FLEX-Klasse. Da aufgrund des Umfangs der Studie lediglich vier Klassen untersucht wurden, stellt sich jedoch die Frage nach der Übertragbarkeit der Ergebnisse auf die Gesamtheit der FLEX-Klassen in Brandenburg.

4.2.2 Forschungsfragen

Vor dem Hintergrund der Relevanz ausgeprägter Schul- und Unterrichtsqualität für eine erfolgreiche Umsetzung von JÜL wird in Studie II überprüft, ob Schulen, die JÜL langfristig umsetzen, sich durch Qualität in diesen Bereichen von den Schulen unterscheiden, die JÜL nicht langfristig umgesetzt haben (II.1). Es wird dabei davon ausgegangen, dass Schulen, die JÜL langfristig umsetzen, eine höhere Qualität zeigen als die übrigen Schulen. Im Anschluss an die reformpädagogischen Motive für JÜL (Abschnitte 4.1 und 4.1.1) und Betonung der Mischung dreier Jahrgänge (JÜL-3) als Ideal, gegenüber der in Berlin verpflichtenden Mischung von zwei Jahrgängen (JÜL-2) mit Möglichkeit zu JÜL-3, sollen weiterhin Unterschiede der Schul- und Unterrichtsqualität zwischen JÜL-2- und JÜL-3-Klassen untersucht werden (Frage II.2). Für die Unterscheidung der Schulen mit langfristiger und nicht langfristiger Umsetzung wird auf die in der ersten Studie identifizierten Typen der JÜL-Implementation (Frage I.4) zurückgegriffen.

4.2.3 Methode

Der folgende Abschnitt stellt die Methode der Studie II dar. Nach einer Beschreibung der Datengrundlage (4.2.3.1) und der daraus definierten Stichprobe (4.2.3.2) wird die Analysestrategie für die einzelnen Fragestellungen (4.2.3.3) im Detail dargestellt.

4.2.3.1 Datengrundlage

Im Rahmen von Studie II zur Untersuchung der Schul- und Unterrichtsqualität standen die Daten aus der zweiten Runde (Schuljahre 2011/12–2016/17) der Berliner Schulinspektion zur Verfügung. Die Schulinspektion ist Teil der externen Evaluation (SenBJW, 2011) und der Strategie zur Qualitätssicherung in Berliner Schulen. Grundlage der Konzeption ist der Berliner Handlungsrahmen für Schulqualität, der sechs Qualitätsbereiche definiert: (1) Lehr- und Lernprozesse, (2) Schulkultur, (3) Schulmanagement, (4) Professionalität der Lehrkräfte, (5) Qualitätsentwicklung und (6) Ergebnisse und Wirkungen (SenBWF, 2012). Schulen werden in der Regel alle fünf Jahre besucht und nach folgendem Verfahren einheitlich inspiziert: (1) Datenerhebung vorhandener Zahlen und Statistiken, (2) Analyse des Schulprogramms und

weiterer wichtiger Dokumente, (3) Onlinebefragungen von Lehrkräften und weiteren pädagogischen Mitarbeiterinnen und Mitarbeitern sowie der Schüler- und Elternschaft im Vorfeld der Inspektion, (4) Schulrundgang zur Erfassung der Standortbedingungen, (5) Unterrichtsbeobachtungen auf der Grundlage eines Unterrichtsbeobachtungsbogens, (6) Interviews mit Lehrkräften, Schulleitung, pädagogischem Personal, Schülerinnen und Schülern, Eltern und (7) Gespräche mit dem nichtpädagogischen Personal (SenBWF, 2012).

Tabelle 11: Deskriptive Übersicht: Qualitätsmerkmale der Schulinspektion

Qualitätsmerkmal	N	M	SD
Schulqualität			
Arbeits- und Kommunikationskultur im Kollegium	282	3,43	0,76
Beteiligung der Schülerinnen und Schüler und der Eltern	282	3,61	0,59
Evaluation schulischer Entwicklungsvorhaben	278	3,01	0,95
Fortschreibung des Schulprogramms	282	2,81	1,01
Personalentwicklung und Personaleinsatz	280	2,85	0,87
Schulleistungsdaten und Schullaufbahn	281	3,48	0,71
Schulleitungshandeln und Qualitätsmanagement	281	3,12	0,89
Schulleitungshandeln und Schulgemeinschaft	272	3,57	0,78
Soziales Klima und soziales Lernen in der Schule	281	2,91	0,86
Unterrichtsqualität			
Fachimmanentes, fachübergreifendes und fächerverbindendes Lernen	282	3,17	0,38
Förderung von Selbstvertrauen und Selbsteinschätzung	282	3,11	0,34
Innere Differenzierung	282	1,74	0,71
Kooperation des pädagogischen Personals	282	3,63	0,61
Kooperatives Lernen	282	1,52	0,6
Lehr- und Lernzeit	282	3,95	0,24
Leistungs- und Anstrengungsbereitschaft	282	3,71	0,51
Lern- und Arbeitsbedingungen	282	4,00	0,06
Methoden- und Medienwahl	282	3,67	0,54
Pädagogisches Klima im Unterricht	282	3,96	0,22
Problemorientiertes Lernen	282	1,19	0,43
Schulinternes Curriculum und Abstimmung des Lehr- und Lernangebots	282	3,21	0,77
Selbstständiges Lernen	282	1,38	0,56
Sprach- und Kommunikationsförderung	282	2,67	0,62
Strukturierung und transparente Zielausrichtung	282	3,73	0,44
Systematische Unterstützung, Förderung und Beratung	282	3,61	0,59
Verhalten im Unterricht	282	3,99	0,12

N: Anzahl der Schulen. In jedes Qualitätsmerkmal gehen zahlreiche Indikatoren ein, die jeweils auf einer vier-stufigen Skala von 1 bis 4 bewertet werden, wobei 4 die beste Bewertung repräsentiert. Die Beurteilung eines Qualitätsmerkmals für eine bestimmte Schule basiert auf dem Durchschnittswert über die relevanten Indikatoren. Die Angaben zu *N*, *M* und *SD* beziehen sich auf diese Durchschnittswerte der Einzelschulen.

Die Ergebnisse aus der Schulinspektion werden für jede Schule jeweils in einem Bericht zusammengefasst. Eine Kurzversion dieses Berichts wird auf dem Portrait der jeweiligen Schule im Schulverzeichnis der Senatsverwaltung für Bildung, Jugend und Wissenschaft digital

veröffentlicht (SenBJW, 2016). Die für die Analysen dieser Studie verwendeten quantitativen Merkmale umfassen alle Qualitätsmerkmale, die an den Schulen im Rahmen der Unterrichtsbeobachtungen verpflichtend geprüft wurden (Tabelle 11). Die Merkmale wurden von den Inspektor*innen auf einer vierstufigen Skala bewertet. In Anlehnung an das Qualitätsmodell von Ditton (2002) konnten sie konzeptionell der Unterrichtsqualität oder der Schulqualität zugeordnet werden.

4.2.3.2 Stichprobe

Zum Zeitpunkt der vorliegenden Studie lagen veröffentlichte Daten der Schulinspektion von 282 Berliner Grundschulen vor. Das entspricht einem Anteil von 79 % der in dem Längsschnitt berücksichtigten Schulen. Die Auswahl, welche Schulen zu einem bestimmten Zeitpunkt inspiziert werden, erfolgt zufällig. Tabelle 12 bestätigt, dass bereits inspizierte Schulen sich in ihren Merkmalen (z. B. Geschlechterverteilung, Anteil der Schüler*innen mit nichtdeutscher Herkunftssprache) nicht bedeutsam von der Gesamtheit der öffentlichen Grundschulen in Berlin unterscheiden.

Tabelle 12: Gegenüberstellung der Daten der Schulstatistik und Schulinspektion

	Nschule	% weiblich	%ndh
SAPH Berlin	398	48.65	36.52
Schulinspektion	282	48.85	37.63

4.2.3.3 Analysestrategie

Für die Untersuchung der Unterschiede in der Schul- und Unterrichtsqualität zwischen den zuvor identifizierten Implementationstypen (II.1) wurden Varianzanalysen (ANOVA) in R (R Core Team, 2016) gerechnet: Als Faktor wurde die latente Klassenvariable und das jeweilige Qualitätsmerkmal als abhängige Variable eingesetzt. Zur Prüfung der Varianzhomogenität wurde der Levene-Test durchgeführt (Eid et al., 2015) und Prüfstatistiken bei einer Verletzung der Annahmen mittels Welch-Test korrigiert (Maxwell & Delaney, 2004). Zur Feststellung der Unterschiede zwischen den Implementationstypen wurden Post-hoc-Analysen nach Scheffé (Eid et al., 2015) durchgeführt. Als Effektstärkemaß wurde η^2 (Cohen, 1988) berechnet.

Zur Untersuchung der Unterschiede in der Schul- und Unterrichtsqualität zwischen den Schulen, die JÜL für zwei bzw. drei Jahrgänge umsetzen, wurden t-Tests für unabhängige Stichproben und Cohens d (Cohen, 1988) berechnet (II.2). Zur Vergleichbarkeit der Effekte der

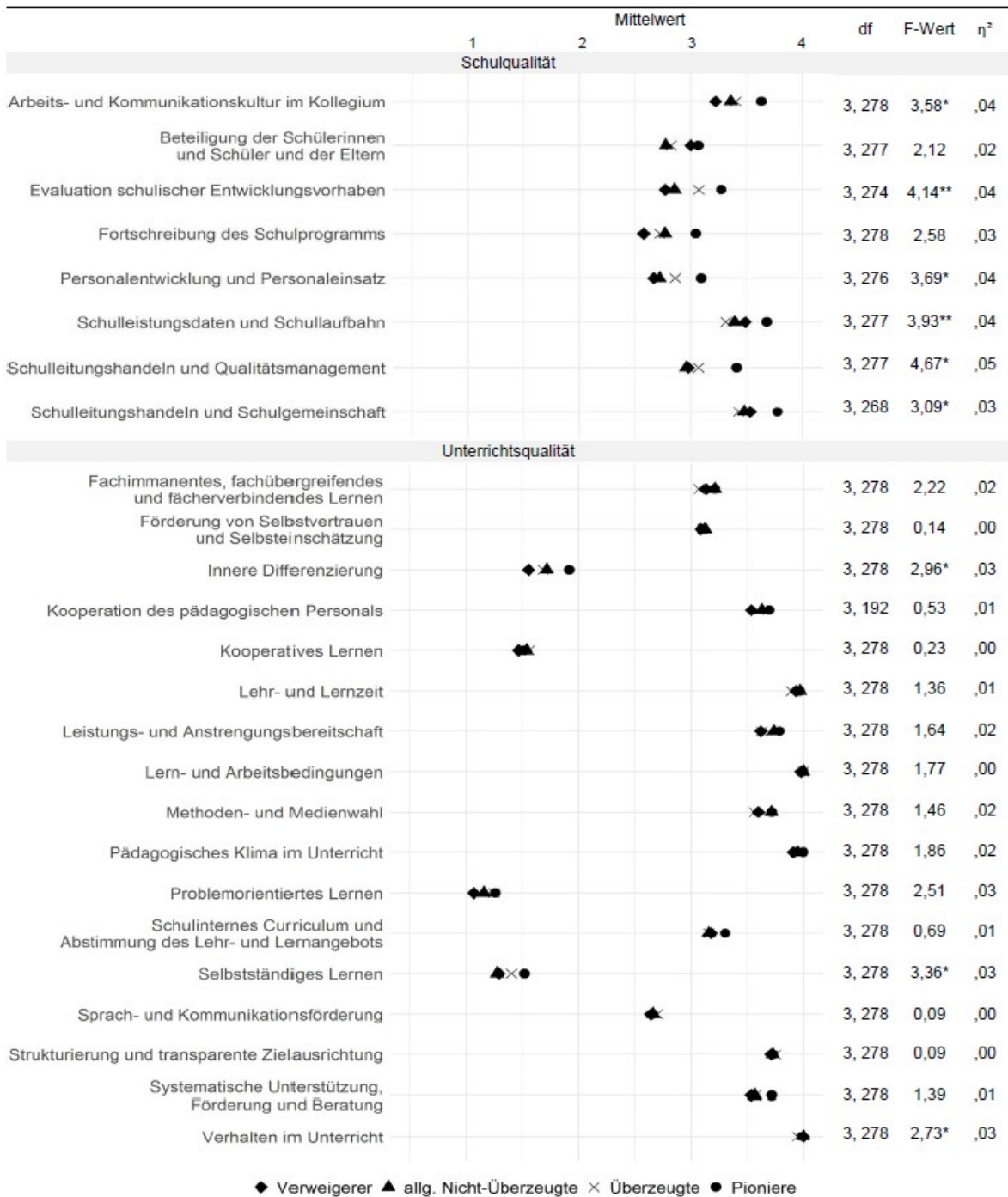
beiden Teilfragestellungen zur zweiten Teilstudie wird zudem η^2 berichtet. Für die Unterscheidung zwischen den beiden JÜL-Arten wurden Schulen, die zu mindestens vier Messzeitpunkten in ihren JÜL-Klassen zwei Jahrgänge zusammenfassten und entweder dem Typus „Pioniere“ oder „Überzeugte“ zugeordnet wurden, als JÜL-2-Schulen gekennzeichnet. Schulen, die zu mindestens vier Messzeitpunkten in ihren JÜL-Klassen drei Jahrgänge zusammenfassten und entweder dem Typus „Pioniere“ oder „Überzeugte“ zugeordnet wurden, wurden demzufolge als JÜL-3-Schulen markiert.

4.2.4 Ergebnisse

Im Rahmen der zweiten Teilstudie wurde die identifizierte Typologie der JÜL-Implementation verwendet, um Unterschiede in der Schul- und Unterrichtsqualität zwischen den einzelnen Typen zu untersuchen. Ziel war es dabei, neben der im ersten Teil untersuchten Verbreitung und Nachhaltigkeit die Tiefe der JÜL-Implementation in Berlin in den Fokus zu nehmen.

II.1 Unterscheiden sich Schulen, die JÜL langfristig umsetzen, durch eine höhere Schul- und Unterrichtsqualität von Schulen, die JÜL nicht langfristig umgesetzt haben?

Für die im Folgenden dargestellten Ergebnisse der Varianzanalysen wurden die Typen „Nicht-Überzeugte“, „Pflichtbewusste“, „nichtüberzeugte Skeptiker“ zu einer Gruppe unter dem Namen „allg. Nicht-Überzeugte“ zusammengefasst. Gründe hierfür sind einerseits die geringe Anzahl an Schulen, die den Gruppen „Nicht-Überzeugte“ und „nichtüberzeugte Skeptiker“ zugeordnet werden können (jeweils 35). Andererseits ist eine Zusammenfassung theoretisch sinnvoll, da alle drei Typen das Jahrgangsübergreifende Lernen zunächst eingeführt und im Laufe von maximal sechs Schuljahren wieder abgeschafft haben. Eine Übersicht der varianzanalytischen Ergebnisse mit den in Studie I identifizierten sechs latenten Klassen, den für Studie II zusammengefassten vier latenten Klassen sowie eine Gegenüberstellung der drei zusammengefassten latenten Klassen der „allg. Nicht-Überzeugten“ findet sich im Anhang der Arbeit (Tab. A- 1 bis Tab. A- 4)



df: Freiheitsgrade (Zwischen den Gruppen, Innerhalb der Gruppen); nach Cohen (1988) wird für $,01 < \eta^2 < ,06$ von einem kleinen Effekt, $\eta^2 < ,14$ von einem mittleren und $\eta^2 > ,14$ von einem großen Effekt ausgegangen

Abbildung 11: Ergebnisse der ANOVAs zu Unterschieden der Schul- und Unterrichtsqualität zwischen den Typen der JÜL-Implementation²⁴

df: Freiheitsgrade (zwischen den Gruppen, innerhalb der Gruppen); nach Cohen (1988): $01 < \eta^2 < ,06$ kleiner, $\eta^2 < ,14$ mittlerer und $\eta^2 > ,14$ großer Effekt.

²⁴ © Thoren, K. & Brunner, M. eingereicht zur Publikation in Zeitschrift für Erziehungswissenschaft, 2017, abgedruckt mit freundlicher Genehmigung von Springer Nature

Die Ergebnisse der Varianzanalysen (Abbildung 11) zeigen, dass sich auf der Ebene der Schule für sechs von acht Merkmalen statistisch bedeutsame Unterschiede feststellen lassen: Arbeits- und Kommunikationskultur im Kollegium, Evaluation schulischer Entwicklungsvorhaben, Personalentwicklung und Personaleinsatz, Schulleistungsdaten und Schullaufbahn, Schulleitungshandeln und Qualitätsmanagement sowie Schulleitungshandeln und Schulgemeinschaft. Statistisch bedeutsame Unterschiede zwischen den einzelnen Gruppen waren in den Post-Hoc-Analysen zugunsten der „Pioniere“ festzustellen. Zwischen „Pionieren“ und „allg. Nicht-Überzeugten“ sind Unterschiede für die Merkmale Evaluation schulischer Entwicklungsvorhaben, Personalentwicklung und Personaleinsatz sowie Schulleitungshandeln und Qualitätsmanagement festzustellen. „Pioniere“ und „Verweigerer“ unterscheiden sich in den Merkmalen Arbeits- und Kommunikationskultur im Kollegium und Evaluation schulischer Entwicklungsvorhaben und „Pioniere“ und „Überzeugte“ in dem Merkmal Schulleistungsdaten und Schullaufbahn. Auf der Ebene des Unterrichts sind für drei von 17 Merkmalen statistisch bedeutsame Unterschiede zwischen den vier JÜL-Typen festzustellen. Im Einzelnen sind das die Merkmale Innere Differenzierung, Selbstständiges Lernen und Verhalten im Unterricht. Die Post-Hoc-Analysen zeigen jedoch lediglich beim Selbstständigen Lernen statistisch bedeutsame Unterschiede zugunsten der „Pioniere“ im Vergleich zu den „allg. Nicht-Überzeugten“. Zur Einschätzung der praktischen Bedeutsamkeit der Unterschiede wurde η^2 als Effektmaß berechnet. Die Größe des Effekts rangiert zwischen $\eta^2 = ,03$ und $\eta^2 = ,05$. Es können also 3–5 % der Varianz eines Qualitätsmerkmals durch die Zugehörigkeit zu einem Implementationstyp erklärt werden. Nach Cohen (1988) sind diese Effekte als klein zu bezeichnen.

II.2 Unterscheiden sich Schulen, die JÜL langfristig umsetzen, in ihrer Schul- und Unterrichtsqualität danach, ob sie in ihren JÜL-Klassen zwei oder drei Jahrgänge zusammenfassen?

Fragestellung II.2 bezieht sich nur auf die Schulen, die JÜL langfristig in ihrem Unterricht umsetzen („Überzeugte“ und „Pioniere“), und untersucht, ob sich die Unterrichtsqualität von Schulen, die zwei Jahrgänge zusammenfassen (JÜL-2), von derjenigen von Schulen unterscheidet, die drei Jahrgänge zusammenfassen (JÜL-3). Zunächst sind in Tabelle 13 die Verteilungen von JÜL-2 und JÜL-3 über die Typen „Überzeugte“ und „Pioniere“ abgebildet.

Tabelle 13: Verteilung der JÜL-2- und JÜL-3-Schulen über die Implementationstypen „Überzeugte“ und „Pioniere“ (absolute Zahlen)

	„Überzeugte“	„Pioniere“	Gesamt	Cohens <i>h</i> („Pioniere“ vs. „Überzeugte“)
JÜL2	53 (73 %)	72 (69 %)	125	-,09
JÜL3	20 (27 %)	33 (31 %)	53	,09
Gesamt	73 (100 %)	105 (100 %)	178	-,36

h: Effektgröße für Unterschiede in Prozentwerten, nach Cohen (1988, 179 ff.) wird ab $h \geq .20$ von einem kleinen Effekt ausgegangen.

Es zeigt sich dabei kein bedeutsamer Unterschied in der Verteilung der JÜL-Arten über die Implementationstypen (gemessen an der Effektstärke *h*; Cohen, 1988). Die Ergebnisse der Unterschiedsanalysen (Abbildung 12) spiegeln die Annahmen zum Teil wider. Für 13 von 17 Qualitätsmerkmalen auf der Ebene des Unterrichts sind bedeutsame Unterschiede zwischen JÜL-2- und JÜL-3-Schulen festzustellen. Die Unterschiede fallen dabei zugunsten der JÜL-3-Schulen aus, die demnach für diese Qualitätsmerkmale des Unterrichts höhere Bewertungen durch die Schulinspektion erhielten. Kleine Effekte zugunsten von JÜL-3 wurden für die Merkmale Kooperation des pädagogischen Personals, Methoden- und Medienwahl, Pädagogisches Klima im Unterricht, Schulinternes Curriculum und Abstimmung des Lehr- und Lernangebots, Strukturierung und transparente Zielausrichtung und Verhalten im Unterricht festgestellt. Mittlere Effekte konnten für die Merkmale Fachimmanentes, fachübergreifendes und fächerverbindendes Lernen, Kooperatives Lernen, Problemorientiertes Lernen, Selbstständiges Lernen, Sprach- und Kommunikationsförderung und Systematische Unterstützung, Förderung und Beratung aufgezeigt werden. Ein großer Unterschied zeigte sich auf der Unterrichtsebene für das Merkmal Innere Differenzierung ($d = ,8$).

Für die Qualität auf der Ebene der Schule wurden vier kleine Effekte (*Arbeits- und Kommunikationskultur im Kollegium, Beteiligung der Schülerinnen und Schüler und der Eltern, Evaluation schulischer Entwicklungsvorhaben, Personalentwicklung und Personaleinsatz*) zugunsten der JÜL-3-Schulen unter insgesamt acht Merkmalen festgestellt. Der einzige bedeutsame Unterschied zugunsten der JÜL-2-Schulen ist mit einem kleinen Effekt für das Merkmal *Schulleistungsdaten und Schullaufbahn* festzustellen.

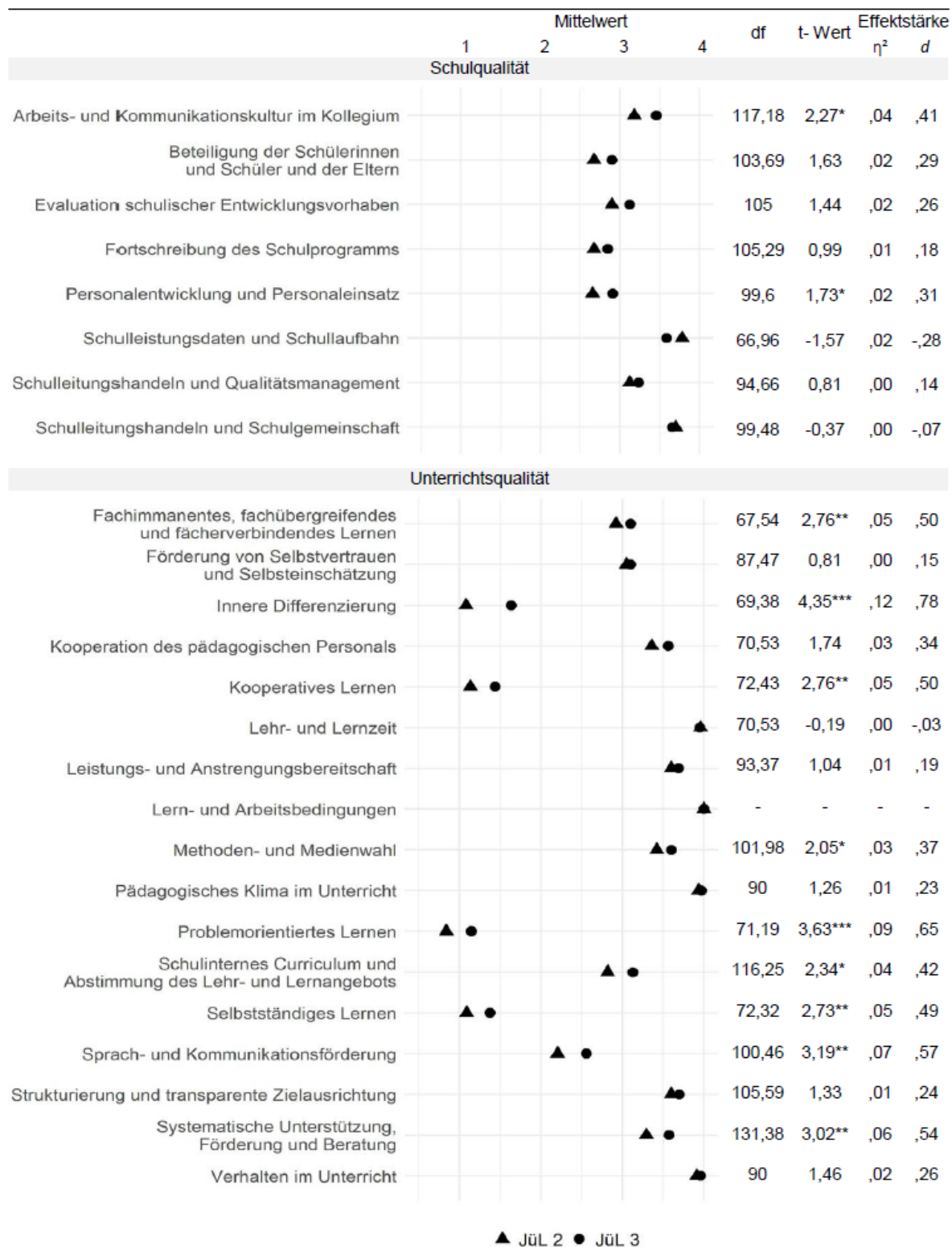


Abbildung 12: Ergebnisse der Unterschiedsanalysen (t-Tests, η^2 , Cohens d) zu Unterschieden der Schul- und Unterrichtsqualität zwischen den JÜL-2- und JÜL-3-Schulen²⁵

Nach Cohen (1988): $01 < \eta^2 < ,06$ kleiner, $\eta^2 < ,14$ mittlerer und $\eta^2 > ,14$ großer Effekt; $,2 < d < ,5$ kleiner, $d < ,8$ mittlerer und $d > ,8$ großer Effekt; ein negatives Vorzeichen bei d kennzeichnet einen Effekt zugunsten von JÜL-2.

²⁵ © Thoren, K. & Brunner, M. eingereicht zur Publikation in Zeitschrift für Erziehungswissenschaft, 2017, abgedruckt mit freundlicher Genehmigung von Springer Nature

4.2.5 Zusammenfassung

Im Anschluss an die Identifikation der Implementationstypen wurde überprüft, inwiefern sich Unterschiede in der Schul- und Unterrichtsqualität zwischen den Implementationstypen feststellen lassen. Mithilfe von Daten der Berliner Schulinspektion konnte gezeigt werden, dass Schulen, die JÜL bereits länger praktizieren („Pioniere“), vor allem auf der Ebene der „Schulqualität“ zu besseren Ergebnissen kommen. Diese Schulen setzen JÜL nicht nur bereits länger an ihren Schulen um, sondern haben gegenüber den „Überzeugten“ auch den Vorteil, dass sie mit hoher Wahrscheinlichkeit aus eigenem Antrieb die Einführung von JÜL in ihrer Schulanfangsphase entschieden und vorangetrieben haben, sich also bereits mit der Innovation identifizieren (Coburn 2003; Gräsel 2010). Die „Pioniere“ zeigen bedeutsam bessere Bewertungen im Bereich der Schulleitung (Schulleitungshandeln und Qualitätsmanagement; Schulleitungshandeln und Schulgemeinschaft) und des Personals (Arbeits- und Kommunikationskultur im Kollegium; Personalentwicklung und Personaleinsatz). Diese Ergebnisse stützen bisherige Befunde und Annahmen, die die besondere Relevanz der Schulleitung (Gräsel 2010; Goldenbaum 2012) und der Zusammenarbeit des übrigen Kollegiums (Kucharz und Wagener 2007) in Implementationsprozessen hervorheben. Die Ergebnisse hinsichtlich der Unterschiede in der Unterrichtsqualität bestätigen bisherige Erkenntnisse und Annahmen nur zum Teil. So zeigen sich Unterschiede für die Merkmale Verhalten im Unterricht, Innere Differenzierung und Selbstständiges Lernen. Vor allem die letzten beiden Merkmale waren aus der Literatur heraus zu erwarten. Für das von Carle und Metzen (2014) hervorgehobene Kooperative Lernen lassen sich jedoch keine Unterschiede zwischen den Typen feststellen. Weiterhin ist auffällig, dass die Merkmale Innere Differenzierung und Selbstständiges Lernen in Berliner Grundschulen insgesamt vergleichsweise schwach bewertet werden. Es ist somit bei allen Implementationstypen noch viel „Luft nach oben“. Insgesamt können 3–5 % der Varianz in den Merkmalen der Schulqualität und 2–3 % der Varianz in den Merkmalen der Unterrichtsqualität mit der Zugehörigkeit zu einem der Implementationstypen erklärt werden. Nach Cohen (1988) werden diese Effekte als klein eingeordnet. Werden jedoch Ergebnisse aus Untersuchungen wie die der Hamburger Schulinspektion (2015) berücksichtigt, wonach der Schularnterschied zwischen Gymnasium und Sekundarschulen lediglich 3 % der Varianz in der Unterrichtsqualität zwischen Schulen erklärt, relativiert sich dieser Eindruck und die Unterschiede zwischen den Typen der JÜL-Implementation gewinnen an Bedeutung. Insgesamt bestätigen diese Ergebnisse die Annahme, dass es vor allem die JÜL-Typen sind, die JÜL in der Breite und nachhaltig

umsetzen, die eine höhere Qualität vor allem auf der Schulebene zeigen. Auf der Unterrichtsebene zeigen sich, anders als angenommen, nur geringe Unterschiede.

Im folgenden Teil der Studie wurden JÜL-2- und JÜL-3-Schulen gegenübergestellt. Dabei wurde angenommen, dass JÜL-3-Schulen eine größere pädagogische Motivation zur Umsetzung von JÜL kennzeichnet, da sie über die rechtlichen Vorgaben hinweg und dem reformpädagogischen Ideal folgend, drei anstelle von zwei Jahrgängen zusammenfassen. Die Ergebnisse zeigen auf der Ebene des Unterrichts deutliche Unterschiede (bei 13 von 17 Merkmalen): So lassen sich auch die von Carle und Metzen (2014) angenommenen Unterschiede im Kooperativen Lernen mit einer mittleren Effektgröße ($d = ,5$) feststellen. Ein besonders großer Effekt war für das Merkmal Innere Differenzierung ($d = ,8$) zu erkennen. Insgesamt zeigten sich für sieben Merkmale auf der Unterrichtsebene mittlere bis große Effekte. Die Befunde bestätigen Annahmen über die besondere Bedeutung der Unterrichtsqualität für die Umsetzung von JÜL (u. a. Carle und Metzen 2014). Dennoch muss auch hier hinzugefügt werden, dass diese Merkmale insgesamt vergleichsweise schwach bewertet werden und auch bei den JÜL-3-Schulen nicht über einen Mittelwert von $M = 1,2$ (Range 1–4) hinauskommen. Die Qualität der Schulen unterscheidet sich zwischen JÜL-2- und JÜL-3-Schulen in fünf von acht Merkmalen, wobei die Effekte insgesamt als klein eingestuft werden können.

Insgesamt weisen die Ergebnisse darauf hin, dass nur wenige Schulen JÜL tiefgreifend, d.h. inklusive der implizierten Normen und Werte (Coburn, 2003) sowie unter Anwendung wesentlicher Komponenten (z. B. kooperatives Lernen, innere Differenzierung), implementiert haben.

4.3 Fazit der Studien zum JÜL

Die beiden Studien zur Einführung des Jahrgangsübergreifenden Lernens in der Schulanfangsphase in Berlin hatten zum Ziel, den Implementationserfolg dieser Reform zu beschreiben. Studie I untersuchte in diesem Zusammenhang, angelegt an die implementationstheoretischen Ausführungen von Coburn (2003) und auf der Grundlage schulstatistischer Daten und einer eigenen repräsentativen Erweiterungsstudie, wie JÜL in Berliner Grundschulen umgesetzt wurde. Insgesamt zeigten die Ergebnisse, dass JÜL zwar nicht an allen Schulen umgesetzt wurde, jedoch die Schulen, die es umsetzten, dies meist auch für einen Großteil ihrer SAPH-Klassen und Jahresunterrichtsstunden sowie über Fächer hinweg einheitlich taten. Coburn (2003) betont stets auch die qualitative Komponente ihrer Merkmale

des Implementationserfolgs (*depth, shift in reform ownership, spread, sustainability*), die im Rahmen der ersten Studie nicht untersucht werden konnten. Doch auf der quantitativ beschreibenden Ebene liefern die Ergebnisse Hinweise dafür, dass JÜL an etlichen Schulen erfolgreich implementiert werden konnte. Den Annahmen von Rogers (2003) folgend, ließ sich die Verbreitung der Reform auch anhand von Typen der JÜL-Implementation beschreiben, die sowohl seine idealtypischen Anwendungskategorien als auch die spezielle Situation in Berlin erkennen ließen. Daran erschließend hatte Studie II zum Ziel, eine qualitative Komponente in die Untersuchungen einzubringen und somit Hinweise über die Tiefe (Coburn, 2003) der JÜL-Implementation zu erhalten. Es wurde daher untersucht, ob sich die Typen der JÜL-Implementation in ihrer Schul- und Unterrichtsqualität voneinander unterscheiden. Die Ergebnisse zeigten, dass Schulen, die JÜL nach quantitativen Kriterien erfolgreich umsetzen, zwar auf Ebene der Schule zum Teil eine höhere Qualität zeigen, auf der Ebene des Unterrichts jedoch kaum Unterschiede zwischen den Typen gibt. Weiterhin konnte gezeigt werden, dass die erfolgreichen JÜL-Typen („Pionieren“ und „Überzeugte“), die in ihren Klassen drei Jahrgänge mischen, in vielen Bereichen der Unterrichtsqualität höhere Bewertungen erhielten, als diejenigen, die zwei Jahrgänge zusammenfassten. Diese Befunde können als Hinweise für die vor allem seitens der Reformpädagogik beschriebenen Vorteile von JÜL-3 gesehen werden. Dennoch ist dabei zu berücksichtigen, dass gerade für die als wesentlich postulierten Merkmale der JÜL-Umsetzung (Carle & Metzen, 2014), wie Innere Differenzierung, Kooperatives und Selbstständiges Lernen, Berliner Schulen insgesamt eher schwach eingeschätzt werden und somit auch bei den JÜL-3-Schulen noch Entwicklungsbedarfe zu konstatieren sind.

5. Studien zur Senkung des Schuleintrittsalters

Die Grundannahme, dass es einen Unterschied macht in welchem Alter ein Kind eingeschult wird, gibt der Diskussion um das richtige Einschulungsalter nicht erst seit Veröffentlichung der PISA-Ergebnisse (OECD, 2001b; 2002) stets wiederkehrenden Aufwind (Kerstan & Spiewak, 2008; Stamm, 2013). Die verschiedenen Positionen und Entwicklungen dieses Diskurses wurden in Kapitel 2 (Abschnitt 2.1) dargestellt. Bezogen auf das Alter, mit dem ein Kind eingeschult werden sollte, formulierte die KMK (1997) in ihren Empfehlungen zum Schulanfang: ²⁶

Angesichts des im internationalen Vergleich hohen durchschnittlichen Einschulungsalters der Kinder in Deutschland stimmt die Kultusministerkonferenz darüber ein, einerseits Maßnahmen zur Reduktion der teilweise hohen Zurückstellungsquoten zu ergreifen, zum anderen Eltern bzw. Erziehungsberechtigte zu ermutigen, von der Möglichkeit der vorzeitigen Einschulung Gebrauch zu machen, wenn die erforderlichen Voraussetzungen gegeben sind. (KMK, 1997, S. 2)

Weiterhin schlagen die Kultusminister*innen einen Stichtag zwischen dem 30. Juni und dem 30. September vor (ebd.). Bis dato wurden Kinder in allen Bundesländern dann eingeschult, wenn sie bis zum 30. Juni das 6. Lebensjahr abgeschlossen hatten. Vorzeitige Einschulungen waren auf Antrag und meist auch mit einer Bescheinigung des Schularztes in einem Zeitraum von drei, in Ausnahmefällen sechs Monaten nach dem Stichtag möglich. Rückstellungen waren ebenfalls in Absprache mit dem Schularzt, aber auch der zukünftigen Schule und ggf. Erzieher*innen der besuchten vorschulischen Einrichtung und den Eltern möglich. Hier konnte ein dem Alter nach schulpflichtiges Kind um maximal ein Schuljahr zurückgestellt werden (Berthold, 2008a; Faust, 2006).

Das Land Berlin entschied sich auf Grundlage der Empfehlungen der KMK (1997) und als Reaktion auf das vergleichsweise hohe Berufseintrittsalter deutscher Schüler*innen (Kastirke, 2006) für eine Verschiebung des Einschulungsstichtages auf den 31. Dezember, sodass Berliner Kinder im Mittel sechs Monate früher eingeschult werden sollten als zuvor (SenBWF, 2004, SenBWF, 2005). Das durchschnittliche Alter bei der Einschulung sollte somit von sechs Jahren und sieben Monaten auf sechs Jahre und einen Monat herabgesetzt werden. Wie in Kapitel 2

²⁶ Teile der folgenden Studien wurden bereits in Form eines Ergebnisberichts unter dem Titel Heinig, E*., Thoren, K.* & Brunner, M. (2014), „Erste Ergebnisse zur Evaluation der Früheinschulung in Berlin“ publiziert; *beide Autorinnen haben gleichermaßen zu diesem Bericht beigetragen

dargestellt, wurde die Stichtagsverschiebung gemeinsam mit der integrativen Schulanfangsphase zum Schuljahr 2005/06 eingeführt. Der Annahme folgend, dass Schüler*innen nicht reif oder fähig genug für die Schule sein müssen, sondern Schulen mit den unterschiedlichen Eingangsvoraussetzungen der Schüler*innen umgehen sollen (Abgeordnetenhaus Berlin, 2004), gab es ab diesem Schuljahr keine offizielle Möglichkeit mehr, sein Kind von der Einschulung zurückzustellen. Im Zusammenspiel mit der flexiblen Verweildauer (Abschnitt 2.3) und der drei Jahre später folgenden Einführung des Jahrgangsübergreifenden Lernens sollten Schulen fortan ihre bis dato weit verbreiteten Homogenisierungsbestrebungen aufgeben und die Heterogenität ihrer Schüler*innen als Chance bergreifen (u. a. Prenzel, 2006).

Das folgende Kapitel widmet sich dem Implementationserfolg der sogenannten Früheinschulung in Berlin und stellt die Wirkung dieser bildungspolitischen Reform auf die schulischen Leistungen der Schüler*innen in das Zentrum der Untersuchungen. Die Untersuchung des Implementationserfolgs ist im Folgenden in drei Studien gegliedert: Studie III beschreibt die Umsetzung der Früheinschulung in Berlin (Abschnitt 5.2). In Studie IV (Abschnitt 5.3) werden die mittleren Leistungen der Schüler*innen entsprechend ihrem Alter dargestellt und zueinander in Beziehung gesetzt. Die dritte Studie zum Schuleintrittsalter (Studie V, Abschnitt 5.4) untersucht sogenannte relative Alterseffekt (siehe auch Abschnitt 5.1) und deren Generalisierbarkeit über Schüler*innengruppen, Jahrgangsstufen und Kohorten sowie Klassen.

5.1 Studie III: Wie wurde die Senkung des Schuleintrittsalters an Berliner Grundschulen umgesetzt?

Als notwendige Voraussetzung für die Evaluation des Implementationserfolgs bei der Senkung des Schuleintrittsalters muss zunächst die Umsetzung der Reform überprüft werden. Dabei gilt die Verbreitung einer Reform oder Innovation bereits als ein Merkmal des Implementationserfolgs (Abschnitt 1.4; Coburn, 2003). Studie III dieser Arbeit nimmt die Implementation der Früheinschulung in den Blick. Neben der mittleren Senkung des Schuleintrittsalters in ganz Berlin soll auch geprüft werden, inwieweit das Alter in allen Klassen der Schulanfangsphase gesenkt wurde. Der theoretische Rahmen dieser Studie ist in Abschnitt 5.1.1 beschrieben. Anschließend werden die Forschungsfragen abgeleitet (Abschnitt 5.1.2) und Methoden (Abschnitt 5.1.3) und Ergebnisse (Abschnitt 5.1.4) dargestellt. Abschnitt 5.1.5 fasst die wichtigsten Erkenntnisse der Studie III zusammen.

5.1.1 Theoretischer Hintergrund

Als Grundlage der Untersuchung werden die formulierten Ziele der Bildungsadministration angeführt, das mittlere Schuleintrittsalter Berliner Schüler*innen zum Schuljahr um sechs Monate, von sechs Jahre und sieben Monate auf sechs Jahren und einen Monat, zu senken. Weiterhin besteht in der Bildungsforschung Konsens über die Relevanz der Berücksichtigung der Mehrebenenstruktur im Bildungssystem, wie sie theoretisch postuliert (Fend, 2009; Gräsel, 2014) und empirisch mit entsprechenden Verfahren modelliert (Raudenbush & Bryk, 2002; Snijders & Bosker, 2012) wird.

5.1.2 Forschungsfragen

Für die erste Studie zur Evaluation des Implementationserfolgs der sogenannten Früheinschulung ergeben sich vor allem aus den im Land Berlin definierten rechtlichen Bedingungen und unter Berücksichtigung der Mehrebenenstruktur im Bildungswesen zwei Fragestellungen: (III.1) Wie hoch ist das mittlere Einschulungsalter in den drei Kohorten? Und entsprechen sowohl die Höhe des Alters als auch die Altersspannweite den bildungspolitischen Erwartungen? Sowie (III.2) Variiert das mittlere Einschulungsalter zwischen den Klassen? – Vor allem die zweite Fragestellung soll sicherstellen, dass für die weiteren Analysen systematische Verzerrungen, z. B. durch die Häufung der vergleichsweise jüngeren oder älteren Schüler*innen in einer Klasse, ausgeschlossen oder ggf. berücksichtigt werden können.

5.1.3 Methode

Der folgende Abschnitt stellt die Methode für die folgenden Untersuchungen dar. Dabei werden zunächst das Studiendesign (Abschnitt 5.2.4.1), die Instrumente (Abschnitt 5.2.4.2) sowie die Stichprobe (Abschnitt 5.2.4.3) vorgestellt. Diese Abschnitte sind weitestgehend auch für die Studien IV und V zutreffend. Abschließend wird die Analysestrategie für Studie III (Abschnitt 5.2.4.4.) vorgestellt.

5.1.3.1 Studiendesign

Die Studien zu den Auswirkungen der Senkung des Schuleintrittsalters werden auf Grundlage dreier Jahrgänge oder Kohorten durchgeführt: (1) einer Hauptkohorte sowie (2) der Vergleichskohorte I und (3) der Vergleichskohorte II (Abbildung 13).

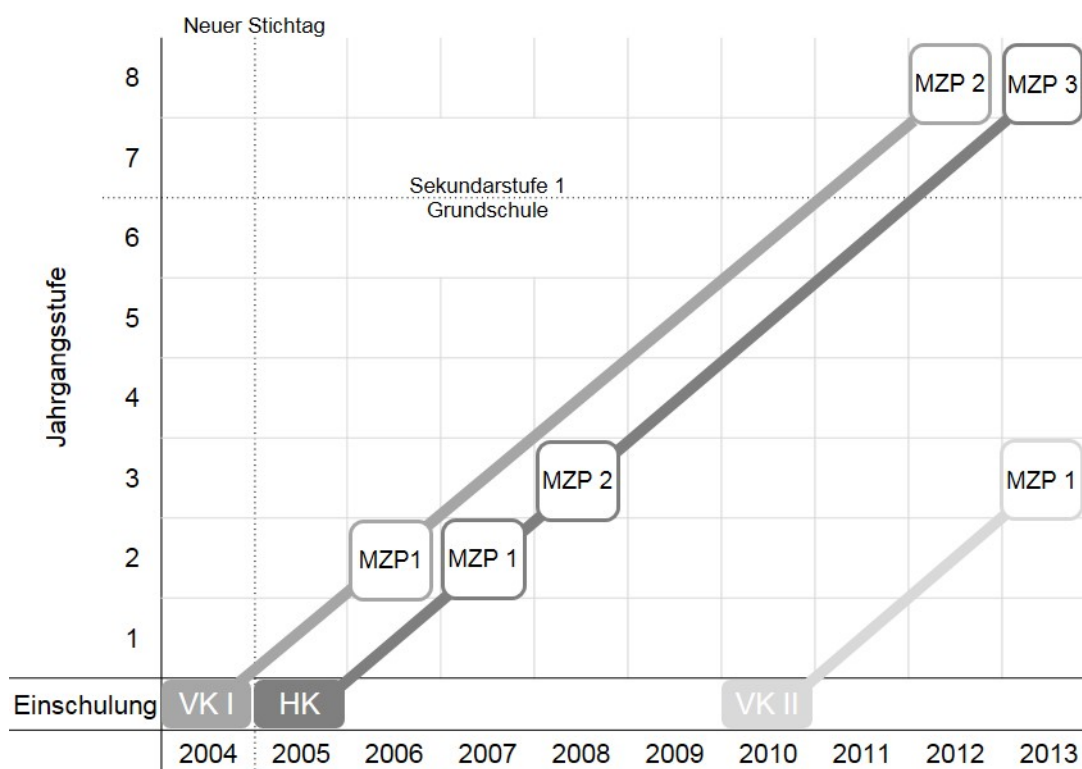


Abbildung 13: Multikohorten-Sequenz-Design der Studien zur Senkung des Schuleintrittsalters

Die Hauptkohorte (HK) stellt dabei den Jahrgang dar, der als Erstes von der Senkung des Schuleintrittsalters betroffen war. Diese Einschulungskohorte ist besonders, da sie sich aus Schüler*innen, die nach der alten Stichtagsregelung, und zugleich aus denjenigen, die nach der neuen Stichtagsregelung eingeschult wurden, zusammensetzt. Dieser Umstand schafft die Bedingungen eines natürlichen Experiments (Shadish et al., 2002) und ermöglicht eine direkte

Gegenüberstellung von Schüler*innen, die im Schuljahr 2005/06 letztmalig planmäßig eingeschult wurden, da sie nach dem alten Stichtag (30. Juni) die Schule begannen, mit Schüler*innen, die aufgrund des neuen Stichtags (31. Dezember) eingeschult wurden. Sie setzt sich daher aus anderthalb Jahrgängen zusammen.

Für die Schüler*innen der Hauptkohorte liegen Daten aus den Orientierungsarbeiten der zweiten Jahrgangsstufe (OA 2 2006/07; MZP 1), den Vergleichsarbeiten der dritten Jahrgangsstufe (VERA 3 2007/08; MZP 2) und den Vergleichsarbeiten der achten Jahrgangsstufe (VERA 8 2012/13; MZP 3) vor. Die Datenkonstellation bei der Hauptkohorte entspricht der einer Trendstudie (Schnell, Hill & Esser, 2005). Bei dieser Form der Längsschnittstudie wird dieselbe Schülerkohorte über mehrere Testzeitpunkte untersucht, allerdings können die einzelnen Testergebnisse nur auf der Ebene des Jahrgangs bzw. definierter Subgruppen (z. B. Mädchen oder Schüler*innen mit einem bestimmten Einschulungsalter), nicht aber auf Individualebene zugeordnet werden. Dies ist der Tatsache geschuldet, dass es sich bei den verwendeten Datensätzen um Archivdaten handelt, die nicht für Untersuchungen im Längsschnitt konzipiert wurden. Eine nachträgliche Verknüpfung der Daten auf Individualebene über die Zeit ist daher nicht möglich.

Um Aufschluss über die Stabilität der auf Grundlage der Hauptkohorte erhaltenen Erkenntnisse zu gewinnen, wurden Daten von zwei Vergleichskohorten herangezogen: (VK I) eine Vergleichskohorte, die ein Jahr vor der Stichtagsverschiebung eingeschult wurde und daher den letzten Jahrgang, mit dem Stichtag 30. Juni, bezeichnet. Für diese Kohorte standen Daten aus den Orientierungsarbeiten der zweiten Jahrgangsstufe (OA 2 2005/06; MZP 1) und den Vergleichsarbeiten der achten Jahrgangsstufe (VERA 8 2011/12; MZP 2) zur Verfügung. Eine weitere Vergleichskohorte (VK II) sollte ermöglichen, die Auswirkungen der Stichtagsverschiebung über die Zeit zu betrachten. Sie beschreibt einen Jahrgang, der sechs Jahre nach der Stichtagsverschiebung eingeschult wurde. Dabei standen Daten der Vergleichsarbeiten der dritten Jahrgangsstufe (VERA 3 2012/13; MZP 1) zur Verfügung.

5.1.3.2 Instrumente

Die verwendete Datengrundlage setzt sich aus groß angelegten, repräsentativen Schulleistungsstudien zusammen: zum einen die Orientierungsarbeiten, die zwischen 2005 und 2007 in der zweiten Jahrgangsstufe an Berliner Schulen durchgeführt wurden, sowie zum

anderen die Vergleichsarbeiten, die seit 2008 in der dritten Jahrgangsstufe und seit 2010 auch in der achten Jahrgangsstufe in Berlin implementiert sind.

5.1.3.2.1 Orientierungsarbeiten

Die Orientierungsarbeiten (OA 2) sind in den Schuljahren 2003/04 und 2006/07 in der zweiten Jahrgangsstufe an Berliner Grundschulen durchgeführte standardisierte Leistungstests in den Fächern Deutsch und Mathematik. Die Aufgaben wurden vom Staatsinstitut für Schulqualität und Bildungsforschung (ISB, 2016) in Bayern in Kooperation mit der Ludwig-Maximilians-Universität München entwickelt und in Bayern pilotiert. Für die Analysen in dieser Dissertation standen die Datensätze aus den Schuljahren 2005/06 und 2006/07 zur Verfügung. Neben den Ergebnissen der Leistungstests enthalten die Datensätze u. a. auch Angaben zu Schulart, Geschlecht, Migrationshintergrund und Geburtsdatum. Einen Überblick über die Daten ist in Tabelle 14 dargestellt.

Tabelle 14: Übersicht zu den Orientierungsarbeiten – Vergleichskohorte I und Hauptkohorte

Schuljahr	N _{Schulen}	N _{Klassen}	N _{Schüler*innen}	% weiblich	% _{ndH} ¹
2005/06	375	1.130	25.223	49,3	32,5
2006/07	373	1.555	31.323	48,7	33,9

¹ ndH: Schüler*innen mit nichtdeutscher Herkunftssprache

5.1.3.2.2 Vergleichsarbeiten

Die jährlich bundesweit durchgeführten standardisierten Verfahren zur Leistungsmessung (Vergleichsarbeiten, kurz VERA) sind neben der Teilnahme an internationalen Schulleistungsstudien wie PISA, TIMSS und IGLU, dem zentralen Überprüfen des Erreichens der Bildungsstandards (Ländervergleich) durch das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) sowie der regelmäßigen nationalen Bildungsberichterstattung eine der vier Säulen der „Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring“ (KMK, 2006). Mit der Strategie reagierten die Kultusminister*innen der Länder auf das erwartungswidrig schlechte Abschneiden deutscher Schüler*innen bei der ersten PISA-Studie im Jahr 2000 (OECD, 2001b). Die Vergleichsarbeiten dienen hier als Instrument der internen Evaluation, welches den Schulen und ihren Klassen neben einer Bestandsaufnahme der Leistungen ihrer Schüler*innen vor allem auch zur Implementation sowohl der Bildungsstandards als auch der Rahmenlehrpläne sowie der datengestützten Schul- und Unterrichtsentwicklung dienen soll (ISQ, 2017; KMK, 2006, KMK, 2015). Die Aufgaben und Testhefte der Vergleichsarbeiten werden vom IQB entwickelt und umfangreich pilotiert.

Tabelle 15: Übersicht zu den Vergleichsarbeiten der 3. Jahrgangsstufe (VERA 3) – Hauptkohorte und Vergleichskohorte II

Schuljahr	N _{Schulen}	N _{Klassen}	N _{Schüler*innen}	% weiblich	% _{ndH} ¹
2007/08	372	1.433	28.205	48,8	28,0
2012/13	376	1.454	22.949	49,0	32,2

¹ ndH: Schüler*innen mit nichtdeutscher Herkunftssprache

Für die Untersuchungen im Rahmen der Dissertation wurden VERA-Datensätze der dritten Jahrgangsstufe aus den Schuljahren 2007/08 und 2012/13 (Tabelle 15) sowie der achten Jahrgangsstufe aus den Schuljahren 2011/12 und 2012/13 (Tabelle 16) verwendet.

Tabelle 16: Übersicht zu den Vergleichsarbeiten der 8. Jahrgangsstufe (VERA 8) - Vergleichskohorte I und Hauptkohorte

Schuljahr	N _{Schulen}	N _{Klassen}	N _{Schüler*innen}	% weiblich	% _{ndH} ¹
2011/12	218	1.008	24.985	49,4%	31,0%
2012/13	220	1.126	28.370	48,6%	30,5%

¹ ndH: Schüler*innen mit nichtdeutscher Herkunftssprache

5.1.3.2.3 Erweiterungsstudien

Für drei der vier verwendeten VERA-Datensätze lagen keine Angaben zum Geburtsdatum oder Einschulungsalter der Schüler*innen vor. Diese Angaben wurden in Kooperation mit dem Institut für Schulqualität der Länder Berlin und Brandenburg e. V. (ISQ) in zusätzlichen, repräsentativen Stichproben erhoben. Die Zusatzstudie fiel unter § 9 des Berliner Schulgesetzes (SenBWF, 2004). Die angeschriebenen Schulen waren somit zur Studienteilnahme verpflichtet. Als Kriterien für die Ziehung der Stichproben wurden die Stratifizierungsmerkmale Bezirk und Ergebnisse der Vergleichsarbeiten am Ende der 3. bzw. 8. Jahrgangsstufe im Schuljahr 2011/12 bzw. 2012/13 (in Quartilen) sowie in der 8. Jahrgangsstufe die Schulart herangezogen, um die Grundgesamtheit aller öffentlichen Berliner Schulen möglichst repräsentativ abzubilden und etwaige Stichprobenfehler möglichst klein zu halten. Für die Erhebung im Schuljahr 2012/13 wurden die ausgewählten Schulen im Anschluss an VERA 8 (Februar 2013) und VERA 3 (Mai 2013) dann in einer Faxabfrage gebeten, die mithilfe des ISQ-Portals erstellten Schüler*innenlisten der ausgewählten Klassen um die Angaben Geburtsmonat und -jahr sowie das Jahr der Einschulung zu ergänzen (Abbildung 14) und die ausgefüllten Dokumente per E-Mail oder Fax an das ISQ zurückzusenden.

VERA-8-Klassenliste der Klasse 08E

Schüler- nummer	Geschlecht	Herkunfts- sprache	Geburtsmonat	Geburtsjahr	Jahr der Ein- schulung
1	w	andere			
2	w	deutsch			
3	m	türkisch			

Abbildung 14: Ausschnitt der Klassenliste aus dem ISQ-Portal

Für die im Schuljahr 2012/13 durchgeführten Erweiterungsstudien lagen die Rücklaufquoten mit jeweils 95 % (VERA 3) bzw. 99 % (VERA 8) im akzeptablen Bereich. Die Altersangaben für VERA 8 im Schuljahr 2011/12 wurden rückwirkend im Jahr 2014 erhoben. Da die Rekonstruktion der gewünschten Angaben einen höheren Aufwand bedeutete als in den Erweiterungsstudien zuvor, wurde im Vorhinein eine Machbarkeitsstudie mit elf Schulen durchgeführt. Diese Schulen wurden von Testleiter*innen besucht, um ein geeignetes Verfahren für die Datenerhebung zu erarbeiten. Nach Auswertung dieser Vorstudie wurden weitere Testleiter*innen für die Datenerhebung rekrutiert. Sie nahmen Kontakt zu den Schulen der repräsentativen Stichprobe auf und vereinbarten mit ihnen das Vorgehen bei der Erhebung. Die Schulen konnten sich dabei zwischen zwei Varianten der Erhebung entscheiden: (1) Mitarbeiter*innen des ISQ vereinbarten mit der Schule einen Termin, um die Daten vor Ort zu erheben; (2) die Schule nutzt die vorbereiteten Klassenlisten aus dem ISQ-Portal, ergänzt die gewünschten Altersangaben selbstständig und sendet die Daten anschließend per E-Mail oder Fax an das ISQ zurück.

Trotz der Erfahrungen aus der Machbarkeitsstudie und den Wahlmöglichkeiten der Schulen konnte für diese Erhebung eine vergleichsweise geringere Rücklaufquote von 93 % erreicht werden. Eine Übersicht über die in den Erweiterungsstudien erfassten Daten ist Tabelle 17 zu entnehmen.

Tabelle 17: Übersicht zu den Erweiterungsstudien

Schuljahr	N _{Schulen}	N _{Klassen}	N _{Schüler*innen}	% weiblich	% _{ndH} ¹
VERA 3					
2012/13	117	117	2.167	49,4	30,3
VERA 8					
2011/12	88	282	6.655	48,7	31,8
2012/13	118	130	3.058	49,4	30,8

¹ ndH: Schüler*innen mit nichtdeutscher Herkunftssprache

5.1.3.3 Stichprobe

Für die Untersuchungen standen die im Studiendesign (Abschnitt 5.1.3.1) dargestellten Datensätze zur Verfügung. Zentral für die Studien zur Senkung des Schuleintrittsalters ist der Vergleich derjenigen Kinder, die durch die Stichtagsänderung früher eingeschult wurden, mit denjenigen Kindern, die nach der alten Regelung eingeschult worden sind. Um diese Vergleiche durchzuführen, wurde auf Grundlage des Einschulungsalters zunächst die sogenannte Altersstichprobe definiert. Sie umfasst alle Schüler*innen, die laut Einschulungsverordnung (SenBWF, 2004, SenBWF, 2005) zur Einschulung verpflichtet waren und weder vorzeitig eingeschult noch von der Einschulung zurückgestellt wurden. Für die Hauptkohorte ergibt sich aufgrund der Stichtagsverschiebung daher eine Begrenzung auf die Schüler*innen, die zum Zeitpunkt der Einschulung zwischen 68 (fünf Jahre und acht Monate) und 85 Monaten (sieben Jahre und einen Monat) alt waren. Tabelle 18 fasst die Kennzeichen aller Schüler*innen sowie die der Altersstichprobe zusammen.

Tabelle 18: Übersicht der Datengrundlage

Schuljahr	Jahrgangsstufe	N _{Schulen}	N _{Klassen}	Gesamt			Altersstichprobe		
				N _{Schüler*innen}	% weiblich	% _{ndH}	N _{Schüler*innen}	% weiblich	% _{ndH}
Hauptkohorte									
2006/07	2	373	1.555	31.323	48,7%	33,9%	28.662	49,0%	33,1%
2007/08	3	372	1.433	28.205	48,8%	28,0%	25.500	49,2%	27,3%
2012/13	8	118	130	3.058	49,4%	30,8%	2.397	50,2%	28,4%
Vergleichskohorte I									
2005/06	2	375	1.130	25.223	49,3%	32,5%	18.767	48,6%	32,1%
2011/12	8	88	282	6.655	48,7%	31,8%	4.050	48,9%	32,0%
Vergleichskohorte II									
2012/13	3	117	117	2.167	49,4%	30,3%	1.570	49,7%	26,6%

5.1.3.4 Analysestrategie

Für die Untersuchungen in Studie III kamen deskriptive und multivariate Verfahren zum Einsatz, die im Folgenden beschrieben werden.

Grundlegend für die folgenden Untersuchungen ist die Annahme der Gleichverteilung der Schüler*innen über die zwölf möglichen Geburtsmonate, da sonst nicht davon ausgegangen werden kann, dass eine Verschiebung des Einschulungsstichtages um sechs Monate auch zu einer Änderung, in diesem Falle also zu einer Reduktion des mittleren Einschulungsalters um

sechs Monate führt. Daher wird in einer Vorstudie (III.0) zunächst die Verteilung der Geburten in Berlin über die Geburtsmonate betrachtet. Dies ist notwendig, weil eine Gleichverteilung der Geburten eine Bedingung für die Gleichverteilung der Einschulungen darstellt. Mithilfe des χ^2 -Tests (Eid et al., 2015, 209 f.) wird geprüft, ob sich die Verteilungen der Geburten in Berlin statistisch bedeutsam von der Gleichverteilung über alle Monate hinweg unterscheiden (Alternativhypothese). Anschließend werden die Monatsverteilungen entsprechend der Einschulungen in den drei Kohorten mittels χ^2 -Test mit der Geburtenverteilung verglichen. Sollten Unterschiede zwischen der Geburten- und der Einschulungsverteilung festgestellt werden, kann die Bedeutsamkeit dieser Unterschied unter Anwendung von Cohens h (Cohen, 1988) eingeschätzt werden.

Für die Betrachtung der Entwicklung des mittleren Einschulungsalters über die Messzeitpunkte und Kohorten wurde deskriptiv vorgegangen (Mittelwert, Standardabweichung). Zur Prüfung der Varianz des mittleren Einschulungsalters über Schulen und Klassen hinweg wurden Intraklassenkorrelationen (ICC) berechnet. Diese geben „das Verhältnis der Varianz zwischen [...] Clustern [...] zur Gesamtvarianz an“ (Geiser, 2010, S. 204). Zur Bildung von Clustern dient, in Anlehnung an die Erkenntnisse der Forschung zur Schuleffektivität (Reynolds et al., 2014; Scheerens & Bosker, 1997), die Zuordnung zu einer Klasse. Auf die Spezifikation eines Modells mit drei Ebenen wird aufgrund der Tatsache, dass in einigen Stichproben der Erweiterungsstudien lediglich eine Klasse je Schule erfasst wurde, sowie aus Gründen der Effizienz der Modelle verzichtet. Folglich wird zunächst ein sogenanntes Nullmodell mit zwei Ebenen im Paket lme4 (Version 1.1-8; Bates, Mächler, Bolker & Walker, 2015) in R (R Core Team, 2016) spezifiziert:

$$\gamma_{ij} = \gamma_{00} + u_{0j} + r_{ij}$$

Dabei wurde das Einschulungsalter als abhängige Variable für jede*n Schüler*in i in Klasse j (γ_{ij}) festgelegt. γ_{00} definiert den Gesamtmittelwert über alle Klassen (*grand mean*) und die Terme r_{ij} und u_{0j} stehen für die Residualwerte der jeweiligen Ebenen (Schüler*in, Klasse). Dem von Hox (2010) empfohlenen Vorgehen folgend, werden Intraklassenkorrelation anschließend wie folgt berechnet:

$$\text{ICC: } \rho = \frac{\sigma_{u_0}^2}{\sigma_{u_0}^2 + \sigma_e^2}$$

$\sigma_{u_0}^2$ steht dabei für die Varianz zwischen den Klassen und σ_e^2 für die Varianz innerhalb der Klassen. Ab einem Wert von $\rho = ,05$ kann nach Hox (2010) und Snijders und Bosker (2012) von einer bedeutsamen Varianz zwischen Klassen ausgegangen werden.

5.1.4 Ergebnisse

Dieser Abschnitt stellt die Ergebnisse der einzelnen Studien und ihrer jeweiligen Fragestellungen detailliert dar.

III.0 Vorstudie

Mithilfe des χ^2 -Tests wird geprüft, ob sich die Geburten in Berlin über die zwölf Monate hinweg gleich verteilen (Nullhypothese). Tabelle 19 gibt eine Übersicht über die tatsächlichen (n_j) und erwarteten (ε_{j0}) Geburtenanteile je Monat sowie den daraus errechneten χ^2 -Wert.

Tabelle 19: erwartete und tatsächliche Verteilung der Geburten über die Monate eines Jahres, sowie der jeweilige χ^2 -Wert

Monat	1998			1999		2004	
	ε_{j0}	n_j	$\frac{(n_j - \varepsilon_{j0})^2}{\varepsilon_{j0}}$	n_j	$\frac{(n_j - \varepsilon_{j0})^2}{\varepsilon_{j0}}$	n_j	$\frac{(n_j - \varepsilon_{j0})^2}{\varepsilon_{j0}}$
Januar	8,33	8,16	0,00	8,93	0,04	8,31	0,00
Februar	8,33	7,41	0,10	7,30	0,13	7,68	0,05
März	8,33	8,49	0,00	8,16	0,00	8,07	0,01
April	8,33	8,18	0,00	7,71	0,05	8,22	0,00
Mai	8,33	8,27	0,00	8,02	0,01	7,97	0,02
Juni	8,33	7,83	0,03	8,49	0,00	8,56	0,01
Juli	8,33	8,91	0,04	9,01	0,05	9,31	0,11
August	8,33	9,23	0,10	8,42	0,00	8,91	0,04
September	8,33	9,03	0,06	8,71	0,02	8,93	0,04
Oktober	8,33	8,44	0,00	8,64	0,01	8,27	0,00
November	8,33	8,07	0,01	7,81	0,03	7,82	0,03
Dezember	8,33	7,98	0,01	8,81	0,03	7,95	0,02
χ^2			0,36		0,38		0,33

Die dargestellten χ^2 -Werte liegen deutlich unter dem kritischen χ^2 -Wert $\chi^2_{\text{krit}} = 19,765$ (Eid et al., 2015, 1000 f.), sodass eine Gleichverteilung der Geburten über die Monate für Berlin weiterhin angenommen wird. Die Angaben zu den Geburten beziehen sich auf die Jahrgänge 1998, 1999 und 2004. Lediglich der Jahrgang 2004 ist auch Teil des Studiendesigns für die Untersuchungen zur Senkung des Schuleintrittsalters (vgl. Abschnitt 5.1.3.1). Da für die Prüfung der Gleichverteilung jedoch keine anderen Jahrgänge zur Verfügung standen und nicht

anzunehmen ist, dass sich die geprüften Jahrgänge von anderen systematisch unterscheiden, wird auch für die in den Studien III bis V relevanten Jahrgänge 2005 und 2010 angenommen, dass diese gleich verteilt sind.

Im Anschluss an die Gleichverteilungen der Geburten in Berlin über die Monate eines Jahres wurde geprüft, inwieweit auch die Geburtsmonate der Schulanfänger*innen gleich verteilt sind.

Tabelle 20: erwartete und tatsächliche Verteilung der Schulanfänger*innen entsprechend ihres Geburtsmonats, sowie der jeweilige χ^2 -Wert

Monat	Vergleichskohorte I			Hauptkohorte		Vergleichskohorte II	
	ϵ_{j0}	n_j	$\frac{(n_j - \epsilon_{j0})^2}{\epsilon_{j0}}$	n_j	$\frac{(n_j - \epsilon_{j0})^2}{\epsilon_{j0}}$	n_j	$\frac{(n_j - \epsilon_{j0})^2}{\epsilon_{j0}}$
Januar	8,33	7,99	0,01	7,28	0,13	8,12	0,01
Februar	8,33	7,38	0,11	6,58	0,37	6,36	0,47
März	8,33	7,80	0,03	6,69	0,32	8,35	0,00
April	8,33	7,93	0,02	6,33	0,48	9,28	0,11
Mai	8,33	8,14	0,00	6,79	0,29	7,70	0,05
Juni	8,33	8,05	0,01	6,56	0,38	8,86	0,03
Juli	8,33	9,63	0,20	9,20	0,09	9,42	0,14
August	8,33	9,54	0,18	9,92	0,30	8,49	0,00
September	8,33	8,78	0,02	10,38	0,50	9,00	0,05
Oktober	8,33	8,61	0,01	10,08	0,37	8,82	0,03
November	8,33	7,76	0,04	9,79	0,26	7,47	0,09
Dezember	8,33	8,39	0,00	10,39	0,51	8,12	0,01
χ^2			0,64		3,99		0,98

Ungleichverteilungen können zum Beispiel durch vorzeitiges oder verzögertes Einschulen entstehen und in den folgenden Analysen zu Verzerrungen führen. Tabelle 20 stellt die tatsächlichen und erwarteten Anteile der Geburtsmonate der Schulanfänger*innen jeweils für den ersten Messzeitpunkt der drei Kohorten gegenüber. Auch hier liegt der jeweilige χ^2 -Wert der drei Kohorten deutlich unter dem kritischen χ^2 -Wert $\chi^2_{\text{krit}} = 19,765$, sodass eine Gleichverteilung der Geburtsmonate bei der Einschulung für die folgenden Analysen angenommen wird.

III.1 Wie hoch ist das mittlere Einschulungsalter in den drei Kohorten?

Zur Untersuchung der Implementation der sogenannten Früheinschulung wird zunächst die Entwicklung des mittleren Einschulungsalters betrachtet. Tabelle 21 bildet Mittelwerte und Standardabweichungen der Gesamt- sowie der Altersstichprobe für die Hauptkohorte ab und stellt diese den Vergleichskohorten gegenüber.

Tabelle 21: Mittleres Schuleintrittsalter nach Jahrgangsstufe und Kohorte (Mittelwert und Standardabweichung)

Schuljahr	Jahrgangsstufe	Alle Schüler*innen		Altersstichprobe	
		M	SD	M	SD
Hauptkohorte					
2006/07	2	76,6	6,1	76,0	4,9
2007/08	3	76,9	6,3	76,2	4,8
2012/13	8	76,6	5,8	76,3	4,8
Vergleichskohorte I					
2005/06	2	79,0	5,7	79,3	3,3
2011/12	8	78,8	5,7	79,4	3,4
Vergleichskohorte II					
2012/13	3	73,9	4,6	73,7	3,3

In der zweiten Jahrgangsstufe im Schuljahr 2006/07 beträgt das durchschnittliche Einschulungsalter für die Stichprobe aller Schüler*innen 76,6 Monate, was in etwa sechs Jahren und fünf Monaten entspricht. Für die Altersstichprobe desselben Messzeitpunktes beträgt das mittlere Einschulungsalter 76 Monate, was sechs Jahren und vier Monaten entspricht. Da in der Hauptkohorte der Stichtag um ein halbes Jahr verschoben wurde, die Kohorte aber noch die Gruppe der Schüler*innen enthält, die nach der alten Stichtagsregelung eingeschult wurden, entspricht die Senkung des mittleren Alters um drei Monate in der Altersstichprobe den Erwartungen. Die Streuung des Einschulungsalters beträgt 6,1 Monate für die Gesamt- und 4,9 Monate für die Altersstichprobe. Für die weiteren Messzeitpunkte der Hauptkohorte (Jahrgangsstufen 3 und 8) zeigt sich ein sehr ähnliches Bild. Bei angenommener Gleichverteilung der Schüler*innen über die einzelnen Geburtsmonate und unter Berücksichtigung der regulären Altersspannweite dieser Kohorte von 18 Monaten können sowohl Mittelwerte als auch Streuungen als erwartungskonform eingeschätzt werden. Einzig die Streuungen der Altersstichprobe sind etwas geringer als erwartet.

Der Ergebnisse der beiden Vergleichskohorten zeigen sich als ebenso erwartungsgemäß: vor der Senkung des Schuleintrittsalters (Vergleichskohorte I) beträgt das durchschnittliche

Einschulungsalter 79 bzw. 78,7 Monate, also sechs Jahre und sieben Monate, und nach der Senkung (Vergleichskohorte II) 73,9 Monate, was sechs Jahren und zwei Monaten entspricht. Auffällig sind die Standardabweichungen der beiden Kohorten. Vor allem die Gesamtstichprobe der Vergleichskohorte I streut vergleichsweise weit um den Mittelwert, was darauf schließen lässt, dass es in dieser Kohorte eine hohe Anzahl an vorzeitigen Einschulungen und Rückstellungen gab (hier beträgt die reguläre Altersspannweite zwölf Monate). Für die Gesamtstichprobe der Vergleichskohorte II zeigt sich die Streuung vergleichsweise niedriger, sodass scheinbar mehr Kinder fristgerecht eingeschult wurden als noch vor der Stichtagsverschiebung. Die Streuungen der Altersstichprobe sind mit 3,3 bis 3,4 Monaten in etwa erwartungsgemäß hoch.

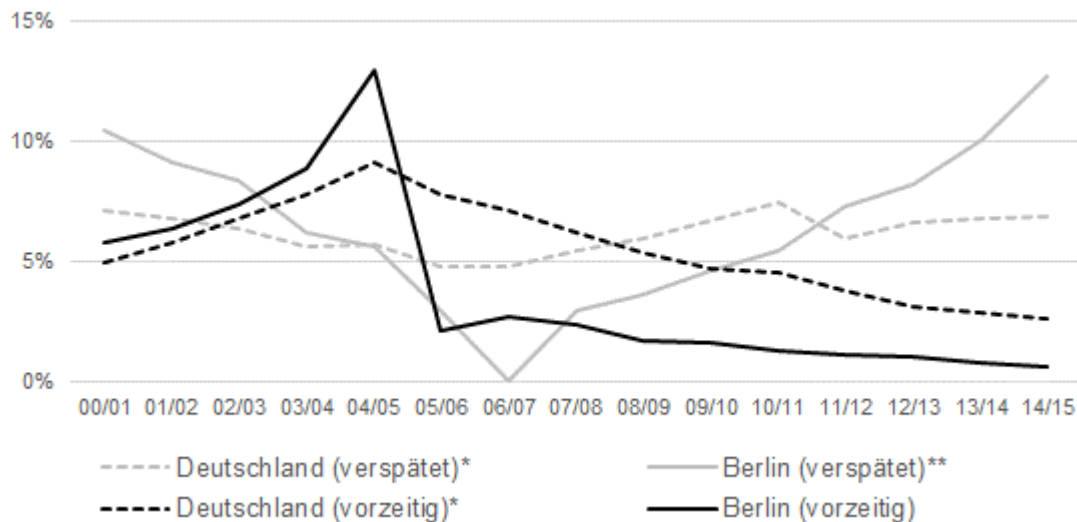


Abbildung 15: Entwicklung der vorzeitigen und verzögerten Einschulungen in Deutschland und Berlin (2000-2014)

* Angaben 2001 ohne Bremen.

** Verspätete Einschulungen wurden 2006 in Berlin unter "Sonstige und ohne Angabe" erfasst.

Die weitere Betrachtung der amtlichen Statistik (Abbildung 15) bestätigt, dass in der Hauptkohorte (Einschulung 2005/06) eine vergleichsweise geringe Anzahl der Schüler*innen vorzeitig oder verspätet eingeschult wurden. Für die Vergleichskohorte I zeigt sich eine sehr hohe Anzahl vorzeitiger Einschulungen und eine dem deutschen Durchschnitt entsprechende Zahl verzögerter Einschulungen. Bei den verzögerten Einschulungen ist, seit Eltern wieder die Möglichkeiten erhielten ihr Kind von der Einschulung zurückzustellen, eine steigende Tendenz festzustellen.

III.2 Variiert das mittlere Einschulungsalter zwischen den Klassen?

Im Anschluss an die Entwicklung des mittleren Einschulungsalters wurde geprüft, ob es empirische Hinweise auf eine Ungleichverteilung des mittleren Einschulungsalters in den Klassen der Berliner Schulen gibt. Wenn auch nicht sehr wahrscheinlich, so könnte es der Fall sein, dass sich in einzelnen Klassen besonders viele jüngere oder ältere Schüler*innen befinden, was, wenn systematisch umgesetzt, zu Verzerrungen in den Ergebnissen der Studien IV und V führen könnte. Als Maß für die Prüfung der Gleichverteilung wurden Intraklassenkorrelationen (ICC) berechnet.

Tabelle 22: Intraklassenkorrelationen (ICC) für das Schuleintrittsalter der drei Kohorten

Jahrgangsstufe	HK	VK I	VK II
	ICC		
2	,02	,01	-
3	,03	-	,00
8	,01	,02	-

Tabelle 22 zeigt die ICCs für die einzelnen Messzeitpunkte der Kohorten. Die Ergebnisse bestätigen, dass der Anteil der Varianz des Einschulungsalters, der durch die Klassenebene aufgeklärt wird, in allen Kohorten und zu allen Messzeitpunkten maximal 3 % beträgt. Nach der Einschätzung von Hox (2010) entspricht dies keinem bedeutsamen Anteil, sodass, bezogen die Verteilung des Einschulungsalters zwischen den einzelnen Klassen, keine Verzerrungen für die weiteren Analysen zu erwarten sind.

5.1.5 Zusammenfassung

In Studie II wurde untersucht, wie die Schulen die Früheinschulung umgesetzt haben. Die Ergebnisse zeigten, dass sich das Einschulungsalter von sechs Jahren und sieben Monaten vor der Reform auf sechs Jahre und zwei Monate reduziert hat. Erklärtes Ziel der Senatsbildungsverwaltung war ein Einschulungsalter von sechs Jahren und einem Monat. Diese Differenz von einem Monat kann mit der erneut steigenden Anzahl von der Einschulung zurückgestellter Schüler*innen erklärt werden. Darüber hinaus konnte immerhin fünf Jahre nach der Reform, im Vergleich zu davor, die mittlere Streuung des Alters reduziert werden, was auch auf eine vergleichsweise geringere Zahl an Rückstellungen schließen lässt.

Abschließend wurde geprüft, ob es Hinweise darauf gibt, dass sich Schüler*innen entsprechend ihrem Alter nicht gleichmäßig auf die untersuchten Klassen verteilen (*spread*; Coburn, 2003).

Mithilfe von Intraklassenkorrelationen konnte gezeigt werden, dass diesbezüglich keine bedeutsamen Verzerrungen in den weiteren Analysen zu erwarten sind.

5.2 Studie IV: Wie wirkt sich die Senkung des Schuleintrittsalters auf die Schüler*innenleistung aus?

Eine wesentliche Funktion schulischer Bildung stellt das Erreichen von Abschlüssen und das Einordnen von Schüler*innen entlang ihrer Leistungen dar. Fend (2009, 49 ff.) bezeichnet dies als Allokationsfunktion und grenzt sie von der Funktion der Enkulturation (Kulturelle Teilhabe und Identität), der Qualifikation (Berufsrelevante Fähigkeiten) und der Integration (Soziale Identität und politische Teilhabe) ab. Die Analysen der Studie IV haben demnach zum Ziel den Implementationserfolg der Einschulungsreform anhand der altersspezifischen schulischen Leistung zu evaluieren. Im Zentrum steht dabei die Prüfung der erfolgreichen Integration der nun früher eingeschulten Schüler*innen in die Schulen. Nach einer theoretischen Einführung (Abschnitt 5.2.1) werden die Forschungsfragen abgeleitet (Abschnitt 5.2.2) und Methoden (Abschnitt 5.2.3) und Ergebnisse (Abschnitt 5.2.4) dargestellt. Abschnitt 5.2.5 fasst die wichtigsten Erkenntnisse aus Studie IV zusammen.

5.2.1 Theoretischer Hintergrund

Eine der größten Sorgen seitens der Öffentlichkeit und der Eltern im Zusammenhang mit der Senkung des Schuleintrittsalters in Berlin war die möglicherweise mangelnde Schulfähigkeit und daraus resultierendes schulisches Versagen der nun früher eingeschulten Kinder (u. a. Hecht, 2010; Vieth-Entus, 2013a)). Der Implementationserfolg dieser Reform kann demnach mit Blick auf die altersbezogene Leistung untersucht werden. Erfolg würde folglich dann konstatiert werden, wenn keine systematische leistungsbezogene Benachteiligung der durch die Reform früher eingeschulten Kinder festgestellt werden kann. In der Forschungsliteratur wird diese Gegenüberstellung altersspezifischer Leistungen auch Alterseffekte genannt. Viele Studien, die Alterseffekte untersuchen, zeigen, dass zumindest zu Beginn der Schullaufbahn ältere Schüler*innen bessere Leistungen zeigen als ihre jüngeren Mitschüler*innen (Black, Devereux & Salvanes, 2011). Die Überprüfung der Auswirkung des Einschulungsalters von Schüler*innen auf ihre Leistungen ist jedoch von verschiedenen Effekten beeinflusst. Zum einen ist es das absolute Alter, das älteren Schüler*innen aufgrund ihrer durchschnittlich größeren Reife und Erfahrungen möglicherweise zu höheren Bildungserträgen verhilft (Bedard & Dhuey, 2008; Stipek, 2002). Unterschiede in den Leistungen können also daher rühren, dass

Kinder mit einem höheren Einschulungsalter wirklich einen Vorteil haben, weil sie mit einem höheren Alter eingeschult wurden als ihre Mitschüler*innen. Weiterhin können Leistungsunterschiede aber auch auf ein höheres Alter zum Testzeitpunkt zurückgeführt werden (Black et al., 2011). Diese beiden Effekte sind meist kollinear und lassen sich empirisch nur schwer voneinander trennen. Studien, die die Trennung der beiden Effekte vornehmen, indem sie Kinder gleichen Alters untersuchen, sehen sich dann mit der Tatsache konfrontiert, dass Kinder unterschiedlich lang beschult wurden. Dieser Effekt kann als Schulzeiteffekt (Black et al., 2011; Cahan & Cohen, 1989) bezeichnet werden. Neben dem absoluten Alter, dem Testalter und der absolvierten Schulzeit wirken hier auch immer sogenannte relative Alterseffekte. Diese entstehen, weil auch ohne vorzeitige oder verspätete Einschulungen die Altersspannweite innerhalb einer Klasse zwölf Monate beträgt und das Durchschnittsalter der Mitschüler*innen und die individuelle relative Altersposition sich direkt oder indirekt auf die schulische Entwicklung eines Kindes auswirken (vgl. Gold, Duzy, Rauch & Murcia, 2012). Die Untersuchung relativer Alterseffekte bringt den Vorteil mit sich, dass dabei die natürlich auftretende Altersspannweite innerhalb schulischer Lerngruppen berücksichtigt wird. Da mit der Stichtagsverschiebung im Schuljahr 2005/06 Kinder sowohl nach der alten als auch nach der neuen Einschulungsverordnung eingeschult wurden und auch keine Möglichkeit zur Rückstellung eines Kindes von der Einschulung bestand, wurden die Bedingungen eines natürlichen Experiments (Shadish et al., 2002) geschaffen, die die Betrachtung relativer Alterseffekte in dieser Studie besonders interessant machen. Zudem steht für die Einschätzung des Implementationserfolgs des gesenkten Schuleintrittsalters in Berlin die schulische Integration der nun früher eingeschulten Schüler*innen im Mittelpunkt. Es liegt also eine Gegenüberstellung der Schüler*innen, die nur nach der neuen Einschulungsregelung eingeschult wurden, mit denjenigen, die nach beiden Regelungen eingeschult worden wären, und solchen, die nur nach der alten Regelung eingeschult wurden, nahe, was am ehesten der Beschreibung relativer Alterseffekte entspricht (Abschnitt 5.2.3.2).

Der nationale und internationale Forschungsstand zur Untersuchung relativer Alterseffekte ist spärlich und uneinheitlich (Black et al., 2011). So stellt ein Großteil der Studien Effekte zuungunsten der jüngeren Schüler*innen im Vergleich zu ihren älteren Mitschüler*innen fest. Gold und Kolleg*innen (2012) weisen solche Effekte unter Erstklässler*innen im Lesen nach, stellen aber weiterhin bereits ein Schuljahr später eine Angleichung der Leseleistung zwischen den Altersgruppen fest. Zu ähnlichen Ergebnissen kommen Crone und Whithurst (1999) in ihrer Untersuchung der Leseleistung US-amerikanischer Schüler*innen: zwar scheinen die Jüngeren zu Beginn der Schulzeit bezüglich ihrer Leseleistung benachteiligt, doch bereits zum Ende der

ersten Jahrgangsstufe sind keine relativen Alterseffekte mehr festzustellen. Dennoch konstatiert Kawaguchi (2011) für japanische Schüler*innen relative Alterseffekte zugunsten der Älteren in Mathematik noch in der vierten und achten Klasse. Eine frühere Untersuchung auf Grundlage der Orientierungs- und Vergleichsarbeiten in Berlin ergab keine bedeutsamen Unterschiede zwischen früh und spät eingeschulten Schülerinnen und Schülern der zweiten und dritten Jahrgangsstufe (Kuhl & Harych, 2007, 2008). Auch mit Blick auf weiterführende Schulen stellt eine Reihe von Studien Benachteiligungen vergleichsweise jüngerer Schüler*innen fest: Cobley und Kolleg*innen (2009) weisen für die Klassen 7, 8 und 9 in England in Mathematik, nicht aber im Fach Englisch relative Alterseffekte nach. Strom identifiziert solche Effekte in der Leseleistung von 15-Jährigen in Norwegen und Fredriksson und Öckert (2005) in Mathematik, Schwedisch und weiteren Fächern bei schwedischen Neuntklässler*innen. Doch auch hier zeigen wiederum anderen Studien, wie die von Lincove und Painter (2006) auf Grundlage der Lese- und Mathematikleistung US-amerikanischer Schüler*innen, dass sich relative Alterseffekte zuungunsten der vergleichsweise Jüngeren auch zu ihrem Vorteil umkehren können.

Wo aber können Gründe für Unterschiede zwischen relativ jüngeren und älteren Schüler*innen liegen? Unterschiede in der kognitiven Entwicklung zeigen sich bereits weit vor der Einschulung (Musch & Grondin, 2001). Dies kann auf unterschiedliche Ursachen zurückgeführt werden: (a) ältere Kinder, sind eben schlicht älter und damit meist biologisch reifer (Bedard & Dhuey, 2008; Stipek, 2002) und (b) hatten mehr Zeit, in einer vorschulischen Einrichtung gefördert zu werden. Nach der Einschulung können sich die Unterschiede verfestigen oder gar vergrößern, da (c) sich der Lehrplan in der Regel am Durchschnitt orientiert, was sich für die Jüngsten, aber auch die Ältesten in einer Klasse nachteilig auswirken kann (Elder & Lubotsky, 2009). Weiterhin kann vermutet werden, dass (d) Lehrkräfte unterschiedliche Erwartungen an Kinder unterschiedlichen Alters innerhalb ihrer Klasse stellen und sich somit auch unterschiedlich ihnen gegenüber verhalten (Hannover & Kessels, 2011; Stipek, 2002). Im Laufe der Schulzeit könnten sich diese Unterschiede in langfristige und stabile Benachteiligungen (z. B. niedrigere Schulabschlüsse oder ein geringeres Einkommen) für die vergleichsweise Jüngeren einer Klasse kumulieren.

Über die Untersuchung verschiedener Arten von Alterseffekten hinaus scheint die Untersuchung von Stichtagsverschiebungen erst seit Kurzem in den Blick bildungswissenschaftlicher und -ökonomischer Forschung genommen zu werden. Bedard und Dhuey (2012) stellen auf Grundlage US-amerikanischer Zensusdaten fest, dass die Erhöhung

des mittleren Einschulungsalters um einen Monat keine Auswirkungen auf die Wahrscheinlichkeiten zum Erreichen eines High-School-Abschlusses hat. Fletcher und Kim (2016) hingegen zeigen ebenfalls für die USA mit NAEP-Daten (National Assessment of Educational Progress), dass bei einer Erhöhung des mittleren Einschulungsalters um einen Monat Viertklässler*innen die Leistung um $SD = ,22$ (Lesen) bzw. $SD = ,14$ (Mathematik) besser ist, im Vergleich zum ursprünglichen mittleren Einschulungsalter. Bei Achtklässler*innen zeigt sich eine Verschlechterung der Leseleistung ($SD = -,23$) jedoch eine deutliche Verbesserung der mathematischen ($SD = ,35$) und der naturwissenschaftlichen ($SD = ,47$) Leistungen. Studien, die die Auswirkungen des Herabsetzens des Einschulungsalters auf die Gesamtleistung aller Schüler*innen untersucht haben, sind der Autorin nicht bekannt.

5.2.2 Forschungsfragen

In Studie IV soll der Blick auf die Auswirkungen der Einschulungsreform in Berlin gelenkt werden. Dabei wird die Leistungsentwicklung vor, während und nach der Reform untersucht. Im Rahmen dieser Studie stellt sich zunächst die Frage, (IV.1) ob und, wenn ja, wie sich die Anteile der verschiedenen Altersgruppen (Abschnitt 5.2.3.2) innerhalb eines Jahrgangs über die Zeit verändert haben. Verschiebungen in den Anteilen können Aufschluss darüber geben, ob in den Altersgruppen unterschiedlich viele Schüler*innen aus der Stichprobe herausgefallen sind (z. B. Wegzug, Wiederholen oder vorzeitiges Aufrücken) und somit Verzerrungen in der Leistung der Schüler*innen befürchtet werden müssen. Dabei wird auch untersucht, ob Schüler*innen bestimmten Alters häufiger ein Gymnasium oder eine Integrierte Sekundarschule besuchen. Unter Berücksichtigung der gewonnenen Erkenntnisse werden anschließend die Leistungen der Schüler*innen unterschiedlichen Alters einander gegenübergestellt (relative Alterseffekte; IV.2) und geprüft, ob sich Unterschiede feststellen lassen. Bei beiden Fragestellungen werden jeweils die Gesamtstichprobe, unterschiedliche Schüler*innengruppen (Mädchen und Jungen, Schüler*innen mit deutscher bzw. nichtdeutscher Herkunftssprache), unterschiedliche Geburtsjahrgänge (die vor der Reform, mit der Reform oder danach eingeschult wurden), sowie unterschiedliche Messzeitpunkte (z. B. zweite, dritte und achte Jahrgangsstufe) in den Blick genommen. Abschließend wird mittels der Leistungsverteilung geprüft, inwieweit sich Leistungen der Schüler*innen entsprechend ihrer Einschulungsmonate überlappen (IV.3).

5.2.3 Methode

Der folgende Abschnitt beschreibt das methodische Vorgehen in Studie IV. Da sich das Studiendesign und die Instrumente von denen der Studie III nicht unterscheiden, wird auf die Darstellungen in den Abschnitten 5.4.2.1 (S. 104) und 5.4.2.2 (S. 105) verwiesen. Die Stichprobe unterscheidet sich ebenfalls nicht von jener in Studie III. Da sie allerdings die Grundlage für die in dieser Studie notwendigen Altersgruppen (Abschnitt 5.3.4.2) darstellt, wird sie nachfolgend nochmals kurz erläutert (Abschnitt 5.3.4.1). Abschließend wird das analytische Vorgehen für die Fragestellungen in Studie IV beschrieben (Abschnitt 5.3.4.3).

5.2.3.1 Stichprobe

Für die Untersuchungen in Studie IV wurde die in Studie III bereits beschriebene Altersstichprobe verwendet (Abschnitt 5.2.4.3). Zur Erinnerung: die Altersstichprobe umfasst alle Schüler*innen, die laut Einschulungsverordnung (SenBWF, 2004, SenBWF, 2005) zur Einschulung verpflichtet waren und weder vorzeitig eingeschult noch von der Einschulung zurückgestellt wurden. Für die Hauptkohorte ergibt sich aufgrund der Stichtagsverschiebung daher eine Begrenzung auf die Schüler*innen, die zum Zeitpunkt der Einschulung zwischen 68 Monaten (fünf Jahre und acht Monate) und 85 Monaten (sieben Jahre und einen Monat) alt waren. Eine Übersicht der Daten ist in Tabelle 18 (S. 109) dargestellt.

5.2.3.2 Altersgruppen

Im Anschluss an die Definition der Alterstichprobe wurden zur Gegenüberstellung von vergleichsweise jüngeren und älteren Schüler*innen auf Grundlage dieser Altersstichprobe drei Gruppen gebildet: (a) Kinder, die im Alter von 68 bis 73 Monaten (max. sechs Jahre und einen Monat alt) eingeschult wurden, zählen zu der Altersgruppe der jüngeren Schüler*innen eines Jahrgangs. Das heißt, diese Kinder wurden aufgrund der neuen Stichtagsregelung im Schuljahr 2005/06 eingeschult, wären jedoch nach der alten Stichtagsregelung erst im kommenden Schuljahr (2006/07) in die Schule gekommen. (b) Kinder, die im Alter von 74 bis 79 Monaten (max. sechs Jahre und sieben Monate alt) eingeschult wurden, zählen zu der Mittelgruppe eines Jahrgangs. Diese Kinder wurden aufgrund der neuen Stichtagsregelung im Schuljahr 2005/06 eingeschult; sie wären aber auch nach der alten Stichtagsregelung in diesem Schuljahr in die Schule gekommen. (c) Kinder, die im Alter von 80 bis 85 Monaten (max. sieben Jahre und einen Monat alt) eingeschult wurden, zählen zu der Altersgruppe der älteren Schüler*innen eines

Jahrgangs. Diese Kinder kamen aufgrund der alten Stichtagsregelung im Schuljahr 2005/06 in die Schule; nach der neuen Stichtagsregelung wären diese Kinder aber ein Jahr früher eingeschult worden. Das Zusammenfassen von Schüler*innen in Altersgruppen stellt in der Forschung zu relativen Alterseffekten ein akzeptiertes Vorgehen dar (Cobley et al., 2009; z. B. Stipek & Byler, 2001). Weiterhin birgt die Definition der Altersgruppen, mit einer Altersspanne von sechs Monaten, für diese Studien den Vorteil, dass die Fallzahlen auch im Hinblick auf Schüler*innengruppen (z. B. Herkunftssprache) ausreichend groß sind, um Effekte präzise schätzen zu können. Tabelle 23 veranschaulicht die Zuordnung zu diesen Altersgruppen für die Hauptkohorte (Einschulung 2005/06), sowie die Vergleichskohorten I (Einschulung 2004/05) und II (Einschulung 2011/12).

Tabelle 23: Zuordnung der Schüler*innen zu den Altersgruppen entsprechend ihrem Einschulungsalter (in Monaten)

Altersgruppe	Alter in Monaten	Geburtsmonate		
		Vergleichskohorte I	Hauptkohorte	Vergleichskohorte II
Jüngere Schüler*innen	68	-	Dez 99	Dez 04
	69	-	Nov 99	Nov 04
	70	-	Okt 99	Okt 04
	71	-	Sep 99	Sep 04
	72	-	Aug 99	Aug 04
	73	-	Jul 99	Jul 04
Mittelgruppe	74	Jun 98	Jun 99	Jun 04
	75	Mai 98	Mai 99	Mai 04
	76	Apr 98	Apr 99	Apr 04
	77	Mrz 98	Mrz 99	Mrz 04
	78	Feb 98	Feb 99	Feb 04
	79	Jan 98	Jan 99	Jan 04
Ältere Schüler*innen	80	Dez 97	Dez 98	-
	81	Nov 97	Nov 98	-
	82	Okt 97	Okt 98	-
	83	Sep 97	Sep 98	-
	84	Aug 97	Aug 98	-
	85	Jul 97	Jul 98	-

Die Gruppenbildung für die Vergleichskohorten erfolgte analog. Dabei ist jedoch zu beachten, dass einerseits Vergleichskohorte I, deren Schüler*innen vor der Stichtagsverschiebung eingeschult wurden, keine jüngeren Schüler*innen umfasst. Andererseits wurden in der Vergleichskohorte II, die nach der Stichtagsverschiebung eingeschult wurde, planmäßig keine Schüler*innen mehr eingeschult, die der Gruppe der Älteren zuzuordnen wären. Diese

Schüler*innen sollten durch die Stichtagsverschiebung im Schuljahr davor mit der Schule begonnen haben.

5.2.3.3 Analysestrategie

Aufgrund der Tatsache, dass für die Hauptkohorte die Kriterien eines natürlichen Experiments gegeben sind, müssen bei den Analysen keine systematischen Selektionseffekte aufgrund von freiwilligen vorzeitigen oder verzögerten Einschulungen berücksichtigt werden. Für diese Effekte muss in Studien mit anderer Datengrundlage aufwendig kontrolliert werden (für eine Übersicht: Lüdtke, Robitzsch, Köller & Winkelmann, 2010). Die vorliegenden Studien hatten demnach den Vorteil, dass sie auf die Anwendung komplexer Verfahren, wie des Instrumentalvariablen-Ansatzes, verzichten und stattdessen auf einfachere beschreibende Verfahren und Effektgrößen sowie Konfidenzintervalle zurückgreifen können. Die Verwendung von Effektgrößen und Konfidenzintervallen erfreut sich in der pädagogisch-psychologischen Forschung wachsender Beliebtheit und ihre Vorzüge und auch Schwachstellen werden umfassend diskutiert (Cumming, 2012)). Auch die American Psychological Association (APA) fordert Wissenschaftler*innen zur verbreiteten Nutzung dieser Methoden auf (Wilkinson, 1999).

Um möglichen Verzerrungen der Stichproben in den Erweiterungsstudien entgegenzuwirken und die Ausprägungen der Population bestmöglich abzubilden, wurden die Schüler*innendaten vor der Durchführung der Analysen mithilfe des Pakets *survey* (Version 3.30-3; Lumley, 2014) in R (R Core Team, 2016) Gewichte berechnet. Dabei wurden die vorhandenen Informationen über die Verteilungen nach Bezirken, Geschlecht, Herkunftssprache und Schulart aus der Population verwendet. Die errechneten Gewichte wurden anschließend in allen multivariaten Verfahren berücksichtigt.

Zur Beantwortung der Forschungsfrage IV.1, wie sich der Anteil der früh eingeschulten Schüler*innen im Vergleich zu den Schüler*innen, die noch nach der alten Stichtagsregelung eingeschult wurden, in der Schullaufbahn ändert, wurde zunächst für die Hauptkohorte (2005/06 eingeschult) die prozentuale Verteilung der drei Gruppen in der Altersstichprobe (jüngere, mittlere, ältere) berechnet. Die Anteile ermöglichen es, einen allgemeinen Eindruck davon zu bekommen, wie viel Prozent der Schüler*innen in den untersuchten Jahrgängen in welchem Alter eingeschult wurden. Ob sich die Anteile der Altersgruppen über verschiedene Schüler*innengruppen (Jungen und Mädchen; Schüler*innen mit und ohne

Migrationshintergrund), Jahrgangsstufen (2, 3 und 8) und Schularten (Integrierte Sekundarschulen und Gymnasien) statistisch und praktisch bedeutsam unterscheiden, wurde mit 95%-Konfidenzintervallen (Cumming & Finch, 2005; Cumming, 2012) und Cohens *h* (Cohen, 1988) überprüft. Im nächsten Schritt konnte dann untersucht werden, wie sich diese Verteilungen für die Vergleichskohorten I und II (2004/05 bzw. 2010/11 eingeschult) darstellten, um so mögliche Unterschiede über die Kohorten hinweg zu identifizieren.

Zur Beantwortung der Forschungsfrage IV.2 zu den relativen Alterseffekten wurden einerseits mittlere Leistungsunterschiede der Hauptkohorte in Deutsch, Lernbereich Leseverstehen, und Mathematik zwischen den Altersgruppen untersucht, und andererseits beispielhaft die altersbezogenen Leistungsverteilungen miteinander verglichen. Um die Vergleichbarkeit der Leistungen über verschiedene Messzeitpunkte, Kohorten und Lernbereiche (z. B. Mathematik) zu ermöglichen, wurde die Leistung jeweils für Deutsch / Lesen und Mathematik an der Alterstichprobe z-standardisiert ($M = 0$, $SD = 1$). Diese Analysen geben Auskunft darüber, ob und, wenn ja, in welchem Ausmaß sich die Altersgruppen in ihren mittleren Leistungen in den einzelnen Jahrgangsstufen (2, 3, und 8) unterscheiden und wie stark sich die Leistungsverteilungen überlappen. Um etwaige Unterschiede in den Verteilungen aufzudecken, wurden diese Analysen zudem getrennt für (a) Jungen, (b) Mädchen sowie Schüler*innen mit (c) deutscher und (d) nichtdeutscher Herkunftssprache durchgeführt.

Sowohl für die Verteilungen als auch für die Leistungsmittelwerte wurden zudem 95%-Konfidenzintervalle berechnet, um die Genauigkeit der ermittelten Anteile und Mittelwerte zu prüfen (Cumming & Finch, 2001; Eid, Gollwitzer & Schmitt, 2013, S. 223; Wilkinson, 1999). Bei der Berechnung der Konfidenzintervalle wurden die hierarchische Datenstruktur berücksichtigt und die Standardfehler korrigiert.

Für die dritte Forschungsfrage in Studie IV, Überlappung der Leistungsverteilungen entsprechend dem Einschulungsalter, werden die Leistungen in Deutsch / Lesen und Mathematik nach Einschulungsmonaten in Boxplot abgebildet.

5.2.4 Ergebnisse

Im Anschluss an die Betrachtung der Umsetzung der Einschulungsreform wurde in der folgenden Studie überprüft, welche Auswirkungen die Reform auf die Verteilung der drei Altersgruppen (Frage IV.1) und die schulischen Leistungen der Altersgruppen (Frage IV.2)

sowie die Verteilung der Leistungen bei Schüler*innen mit unterschiedlichem Einschulungsalter (Frage IV.3) hat.

IV.1 Sind bezogen auf Verteilungen der Altersgruppen über die Zeit und Schüler*innengruppen Veränderungen festzustellen?

Abbildung 16 zeigt (a) die Verteilung der drei Altersgruppen für die drei Messzeitpunkte der Hauptkohorte insgesamt, sowie (b) für die Verteilung von Mädchen und (c) die Verteilung von Schüler*innen mit nichtdeutscher Herkunftssprache. In Abschnitt (a) der Abbildung 16 ist zu erkennen, dass die Gruppe der älteren Schüler*innen in den Jahrgangsstufen 2 und 3 mit $p_{\text{alt-JG2}} = 28,0 \%$, CI [27,4; 28,7] bis $p_{\text{alt-JG3}} = 28,8 \%$, CI [28,2; 29,5] geringer vertreten ist als die anderen beiden Altersgruppen. Die Prüfung der praktischen Bedeutsamkeit mittels Cohens h weist auf einen Effekt von $h \leq ,18$ hin (Tab. A- 5 bis Tab. A- 11). In der Kategorisierung nach Cohen (1988) entsprechen diese Effektstärken nicht einmal einem kleinen Effekt²⁷ und sind daher als praktisch nicht bedeutsam einzustufen. Die 95%-Konfidenzintervalle der Älteren überlappen sich nicht mit denen der anderen beiden Gruppen, was darauf hinweist, dass dieser Unterschied statistisch bedeutsam ist. Weiterhin fällt auf, dass die Gruppe der jüngeren Schüler*innen über die drei Messzeitpunkte hinweg vergleichsweise abnimmt ($p_{\text{jung-JG2}} = 35,5 \%$, CI [34,8; 36,1], $p_{\text{jung-JG8}} = 31 \%$, CI [29,5; 33,3]). Die Prüfung der praktischen Bedeutsamkeit mittels Cohens h weist auf einen Effekt von $h = ,04$ für die Jahrgangsstufen 2 und 3 und $h = ,09$ für die Jahrgangsstufen 2 und 8 hin.

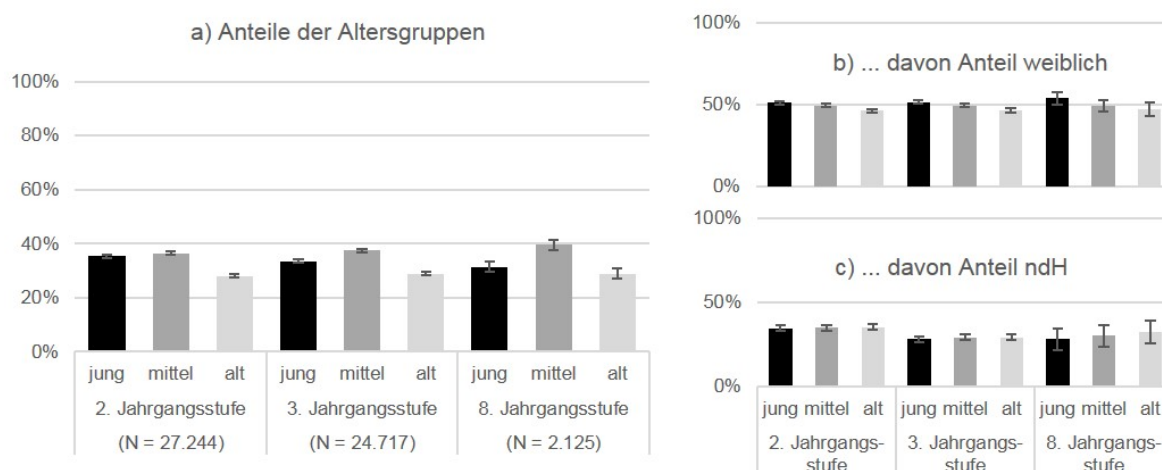


Abbildung 16: Altersgruppenverteilungen der Hauptkohorte

²⁷ $h \geq ,2$ kleiner, $h \geq ,5$ mittlerer, $h \geq ,8$ großer Effekt (vgl. Cohen, 1988, S. 198)

Abschnitt (b) der Abbildung 16 stellt die Anteile der Schülerinnen an der Gesamtanzahl der Schüler*innen in der jeweiligen Altersgruppe dar. Insgesamt sind keine bedeutsamen Unterschiede festzustellen. In der Gruppe der älteren Schüler*innen zeigt sich der Anteil der Mädchen als etwas geringer gegenüber den anderen Altersgruppen. Dieser Unterschied ist für die Jahrgangsstufen 2 und 3 statistisch aber nicht praktisch bedeutsam, was sich darin zeigt, dass die Konfidenzintervalle zwischen den Altersgruppen nicht überlappen und die Effektgröße einen Wert von $h = ,13$ nicht übersteigt. In der Gruppe der Jüngeren zeigt der Anteil der Mädchen eine leicht steigende, wenn auch nicht bedeutsame Tendenz. Die Anteile der Schüler*innen mit nichtdeutscher Herkunftssprache (c) verteilen sich zu allen drei Messzeitpunkten in etwa gleichmäßig über die Altersgruppen. Auffällig ist, dass die Anteile in der dritten Jahrgangsstufe in allen drei Altersgruppen niedriger ausfallen als noch zum vorherigen Messzeitpunkt (2. Jahrgangsstufe). Dieser Unterschied ist ebenfalls statistisch, aber nicht praktisch bedeutsam ($h = ,12 - ,14$).

Um sicherzustellen, dass sich die Verteilungen der Anteile vor und nach der Senkung des Schuleintrittsalters nicht unterscheiden, um folglich die Leistungen der einzelnen Altersgruppen vergleichen zu können, wurden die Altersgruppenverteilungen der Hauptkohorte denen der beiden Vergleichskohorten gegenübergestellt. Die Gegenüberstellung von Hauptkohorte und Vergleichskohorte I, also dem letzten Jahrgang vor der Senkung des Schuleintrittsalters, ist in Abbildung 17 dargestellt. Für beide Jahrgangsstufen zeigt sich, dass die vergleichsweise älteren Schüler*innen in der Hauptkohorte zahlenmäßig einen geringeren Anteil haben ($p_{JG2} = 43,4\%$, CI [42,6; 44,2]) als die älteren Schüler*innen der Vergleichskohorte I (z. B. $p_{JG2} = 47,3\%$, CI [46,5; 48,2]). Mit $h = -,08$ (JG 2) bzw. $h = -,13$ (JG 8) sind diese Unterschiede aber für beide Jahrgangsstufen nicht von praktischer Relevanz. Für die Mädchen zeigen sich zwischen den beiden Kohorten keine Unterschiede: hier machen die älteren insgesamt einen etwas geringen Anteil aus, als die vergleichsweise jüngeren Schülerinnen (hier: „mittel“). Und auch für Schüler*innen mit nichtdeutscher Herkunftssprache sind keine Verteilungsunterschiede zwischen den Kohorten erkennbar.

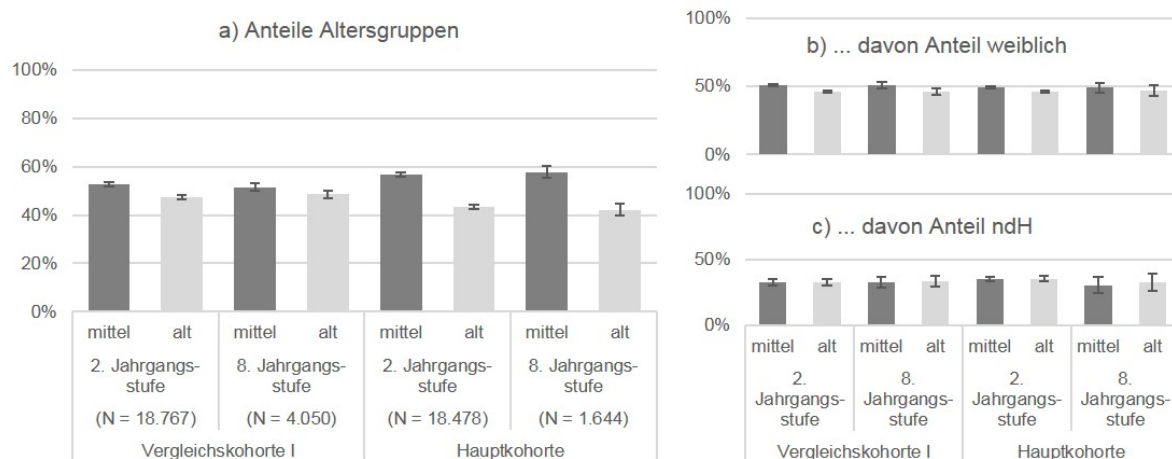


Abbildung 17: Altersgruppenverteilungen der Vergleichskohorte I im Vergleich zur Hauptkohorte

Für den Vergleich der Ergebnisse der Hauptkohorte mit einem Jahrgang, der bereits ausschließlich nach der neuen Stichtagsregelung eingeschult wurde, wird die Vergleichskohorte II herangezogen. Die Gegenüberstellung dieser beiden Kohorten im Hinblick auf die Verteilung der Altersgruppen zeigt keine statistisch bedeutsamen Unterschiede (Abbildung 18): die Jüngeren machen in beiden Kohorten einen etwa gleich großen Anteil aus (z. B. $p_{jung-HK} = 47,2\%$, CI [46,4; 48,0]; $p_{jung-VKII} = 47,4\%$, CI [44,9; 49,9]). Bei den Mädchen und den Schüler*innen mit nichtdeutscher Herkunftssprache zeigen sich ebenfalls keine Unterschiede in den Altersgruppenverteilungen über die beiden Kohorten hinweg.

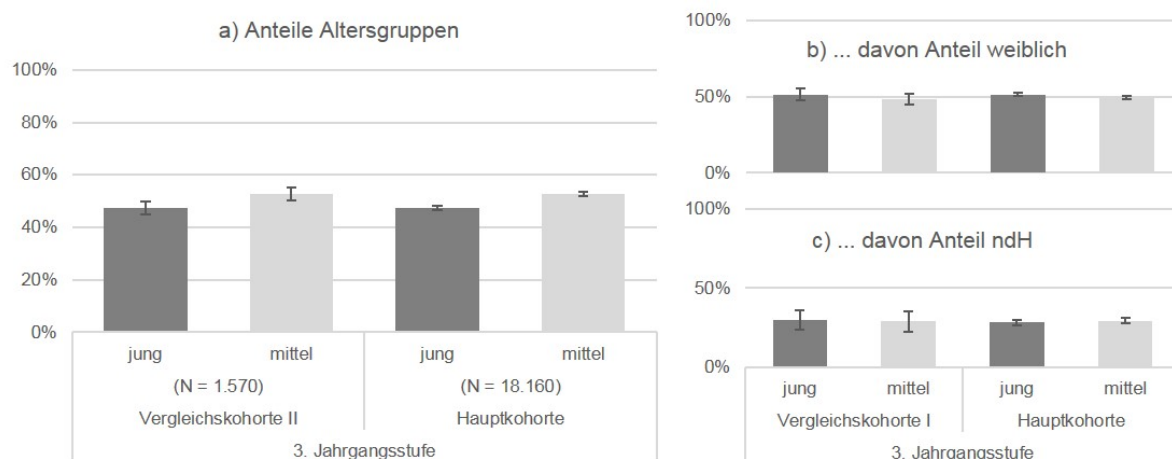


Abbildung 18: Altersgruppenverteilungen der Vergleichskohorte II im Vergleich zur Hauptkohorte

Insgesamt kann zusammenfassend festgehalten werden, dass, obwohl sich für einige Messzeitpunkte zwischen den Altersgruppen statistisch bedeutsame Verschiebungen feststellen lassen, keine dieser Verschiebungen nach Cohen (1988) praktisch bedeutsam ist. Eine

systematische Verzerrung der Leistung durch die Verschiebungen ist demzufolge nicht zu erwarten.

Im Anschluss an die Verteilung der Altersgruppen über die verschiedenen Messzeitpunkte hinweg wurde überprüft, wie sich die Altersgruppen in der achten Jahrgangsstufe der Hauptkohorte an den weiterführenden Schulen verteilen. In Berlin wechseln Schüler*innen in der Regel nach der sechsten Jahrgangsstufe auf ein Gymnasium oder eine Integrierte Sekundarschule. Die Ergebnisse sind in Abbildung 19 dargestellt. Es zeigt sich, dass die Altersgruppen in den beiden Schularten ähnlich verteilt sind (a). Dabei liegt der größte Anteil jeweils bei der Mittelgruppe (z. B. $p_{\text{mittel-GYM}} = 40,9\%$, CI [37,9; 44,0]; $p_{\text{mittel-ISS}} = 38,3\%$, CI [35,5; 41,1]). Jüngere bzw. ältere Schüler*innen verteilen sich in etwa gleich (z. B. $p_{\text{jung-ISS}} = 31,1\%$, CI [28,4; 33,7]; $p_{\text{alt-ISS}} = 30,5\%$, CI [27,4; 33,6]). Der Blick auf die 95%-Konfidenzintervalle zeigt Überlappungen der jeweiligen Altersgruppen für Gymnasium und ISS (z. B. Jüngere am Gymnasium und Jüngere an der ISS). Die geringen Anteilsunterschiede zwischen den Schularten (z. B. $p_{\text{jung-GYM}} = 31,7\%$, $p_{\text{jung-ISS}} = 31,1\%$) und die relativ kleine Größe der Konfidenzintervalle einerseits (z. B. CI_{GYM} [28,6; 34,5], CI_{ISS} [28,4; 33,7]) sowie die Interpretation der Überlappungen (vgl. Cumming & Finch, 2005) andererseits lassen den Schluss zu, dass es sicher weder um empirisch noch praktisch bedeutsame Unterschiede handelt. Für Mädchen (b) und Schüler*innen mit nichtdeutscher Herkunftssprache (c) kann dieser Befund repliziert werden.

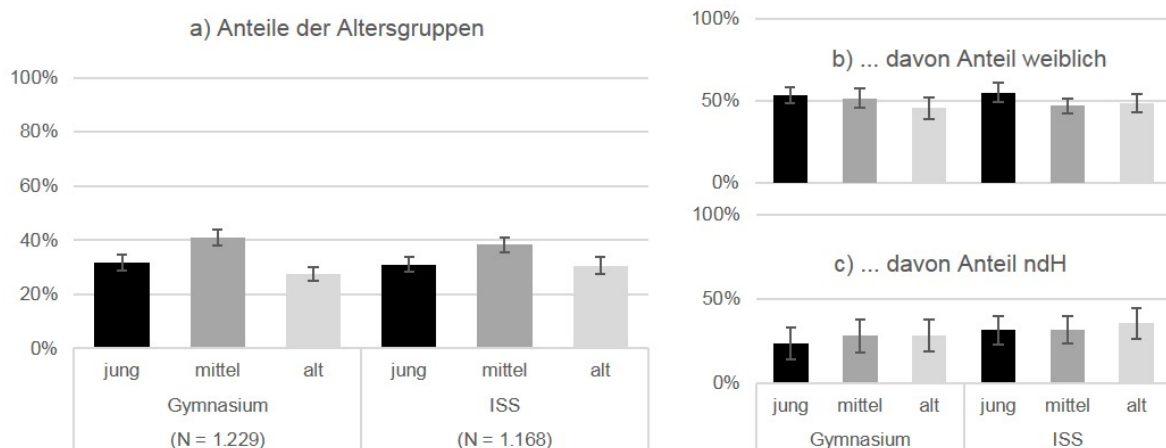


Abbildung 19: Altersgruppenverteilungen der Hauptkohorte, nach Schularten unterteilt

IV.2 Wie unterscheiden sich die Leistungen zwischen den einzelnen Altersgruppen?

In den folgenden Analysen wurden mögliche leistungsbezogenen Unterschiede, sogenannte relative Alterseffekte, zwischen den Altersgruppen in den Blick genommen. Dabei wurden zunächst die Leistungen der einzelnen Altersgruppen für die Hauptkohorte betrachtet. Abbildung 20 stellt die Ergebnisse für Deutsch / Lesen bzw. Leseverstehen und Abbildung 21 für Mathematik dar. Tabellarische Übersichten einschließlich aller Leistungsmittelwerte und den entsprechenden Konfidenzintervallen finden sich im Anhang dieser Arbeit (Tab. A- 12).

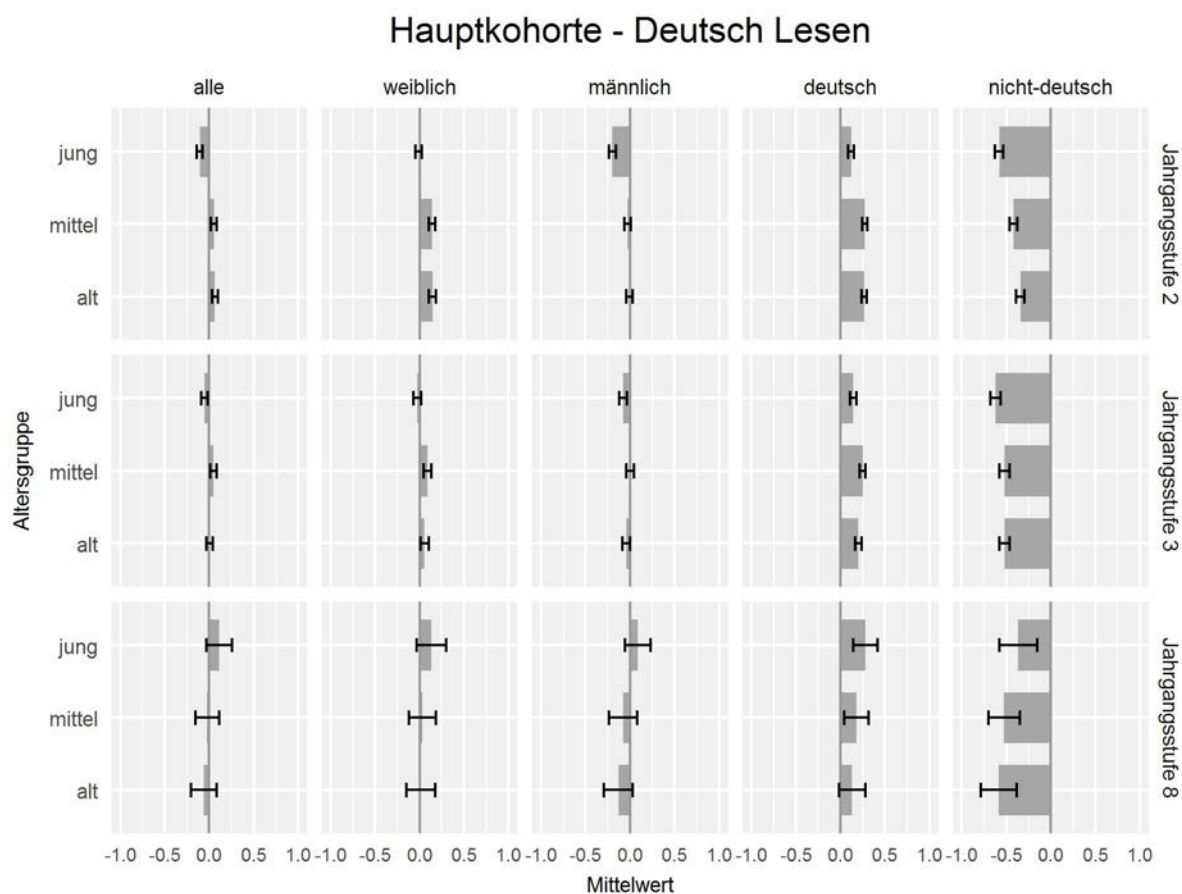


Abbildung 20: Mittlere Leistung (z-standardisiert) der Altersgruppen der Hauptkohorte in Deutsch / Lesen

In beiden Fächern ist in Jahrgangsstufe 2 eine Benachteiligung der jüngeren relativ zu den mittleren und älteren Schüler*innen festzustellen: die Jüngeren erreichten eine mittlere Leistung von $z_{\text{jung-DE}} = -0,10$, CI $[-0,14; -0,07]$ in Deutsch / Lesen und $z_{\text{jung-MA}} = -0,12$, CI $[-0,16; -0,08]$ in Mathematik. Im Vergleich dazu wiesen die Mittleren eine Leistung von $z_{\text{mittel-DE}} = 0,05$, CI $[0,02; 0,08]$ bzw. $z_{\text{mittel-MA}} = 0,06$, CI $[0,02; 0,09]$ und die Älteren von $z_{\text{alt-DE}} = 0,06$, CI $[0,03; 0,10]$ bzw. $z_{\text{alt-MA}} = 0,08$, CI $[0,04; 0,12]$ auf. Die Benachteiligung lässt sich mithilfe der Konfidenzintervalle in zwei Richtungen feststellen: (a) das Konfidenzintervall für die Gruppe der Jüngeren umfasst

nicht den Wert 0, der durch die z -Transformation den Leistungsmittelwert aller Schüler*innen der Altersstichprobe darstellt. Das bedeutet, die Gruppe der Jüngeren zeigt im Mittel signifikant schlechtere Leistungen als der Durchschnitt der Altersstichprobe. (b) Die Grenzen des Konfidenzintervalls der Jüngeren überschneiden sich mit keiner der Grenzen der Konfidenzintervalle der mittleren und älteren Schüler*innen. Das bedeutet, dass die Jüngeren zusätzlich im Mittel signifikant schlechtere Leistungen zeigen als die Schüler*innen der beiden übrigen Altersgruppen. Zwischen den Altersgruppen der mittleren und älteren Schüler*innen sind auf diese Weise keine bedeutsamen Leistungsunterschiede festzustellen.

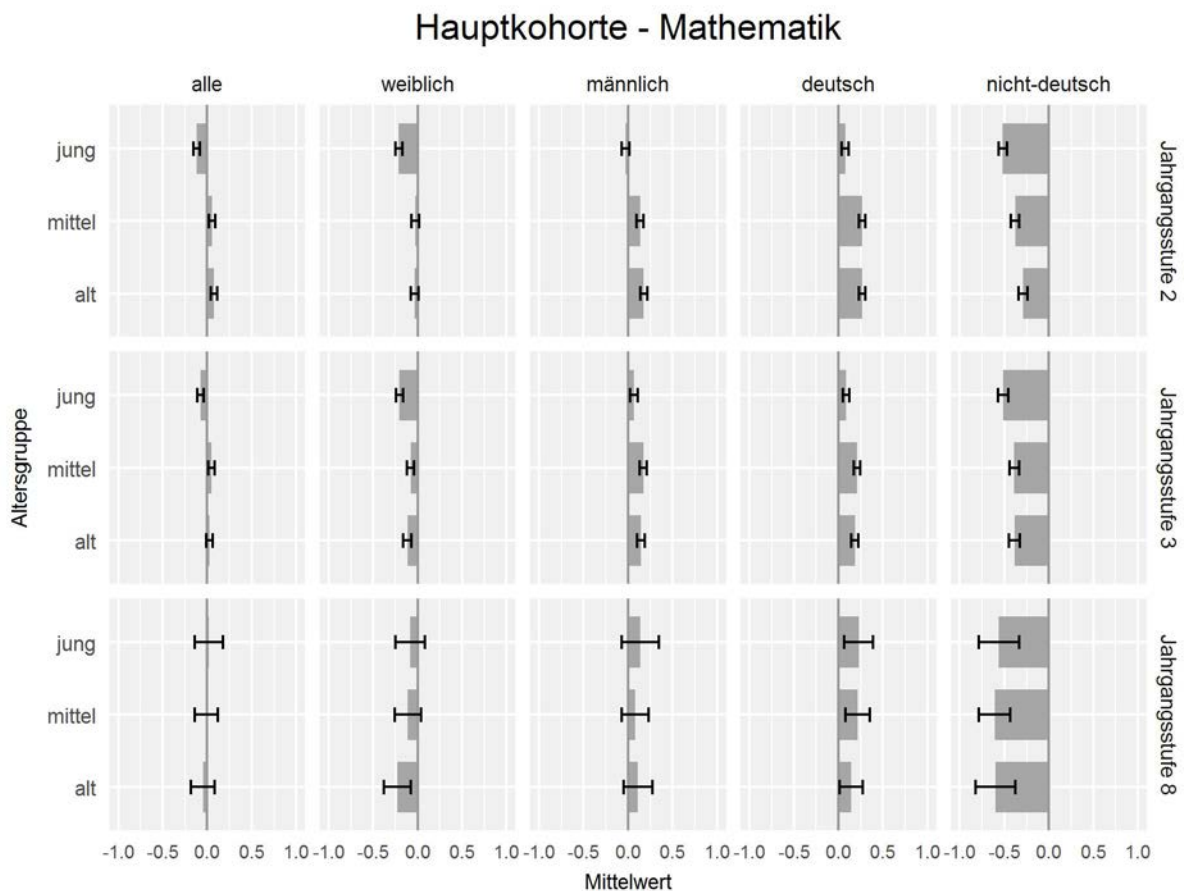


Abbildung 21: Mittlere Leistung (z-standardisiert) der Altersgruppen der Hauptkohorte in Mathematik

Die Benachteiligung hat sich bereits ein Jahr später, in der dritten Jahrgangsstufe, deutlich abgeschwächt (z. B. $z_{\text{jung-DE}} = -0,05$, CI $[-0,09; -0,02]$ und $z_{\text{mittel-DE}} = 0,04$, CI $[0,01; 0,08]$). Jedoch lässt sich hier für das Fach Deutsch auch ein kleiner, aber signifikanter Unterschied zwischen den Mittleren und Älteren feststellen ($z_{\text{mittel-DE}} = 0,04$, CI $[0,01; 0,08]$, $z_{\text{alt-DE}} = 0,00$, CI $[-0,03; 0,04]$). In Jahrgangsstufe 8 sind keine bedeutsamen Leistungsunterschiede zwischen den Altersgruppen in beiden Testdomänen feststellbar (z. B. $z_{\text{jung-MA}} = 0,02$, CI $[-0,14; 0,18]$ und $z_{\text{alt-MA}} = -0,05$, CI $[-0,18; 0,09]$).

Mit Blick auf die Geschlechterverteilung zeigen sich für Deutsch / Lesen signifikante Vorteile der mittleren und älteren Mädchen in den Jahrgangsstufen 2 und 3 gegenüber ihren jüngeren Mitschülerinnen und dem Durchschnitt der Altersstichprobe (z. B. in Jahrgang 3: $z_{\text{jung-DE}} = -,03$, CI [-,07; ,01]; $z_{\text{mittel-DE}} = ,09$, CI [,05; ,13]). In Jahrgangsstufe 8 sind keine bedeutsamen Leistungsunterschiede mehr festzustellen. In Mathematik ist eine generelle leistungsbezogene Benachteiligung der Mädchen zu erkennen, die in den Jahrgangsstufen 2 und 3 bei den Jüngeren signifikant stärker ausfällt (z. B. in Jahrgang 2: $z_{\text{jung-MA}} = -,21$, CI [-,25; ,17]). In Jahrgangsstufe 8 sind erneut keine bedeutsamen Leistungsunterschiede zwischen den Altersgruppen festzustellen. Bei den Jungen ist die mittlere Leistung in Deutsch / Lesen bei den Jüngeren der Jahrgangsstufe 2 signifikant schlechter als bei den mittleren und älteren Schülern (z. B. $z_{\text{jung-DE}} = -,20$, CI [-,24; -,16]; $z_{\text{alt-DE}} = -,01$, CI [-,04; ,03]). In Jahrgang 3 bestehen Unterschiede in der Leseleistung zwischen jüngeren und mittleren, aber nicht zwischen jüngeren bzw. mittleren und älteren Schülern (z. B. $z_{\text{jung-DE}} = -,08$, CI [-,12; -,04]; $z_{\text{alt-DE}} = -,04$, CI [-,08; ,00]). Keine bedeutsamen Unterschiede sind ebenfalls für die Jahrgangsstufe 8 festzustellen. In Mathematik schneiden die mittleren und älteren Schüler im Mittel signifikant besser ab als die jüngeren. Dies gilt für die Jahrgangsstufen 2 und 3 (z. B. in Jahrgang 2: $z_{\text{jung-MA}} = -,03$, CI [-,07; ,01]; $z_{\text{mittel-MA}} = ,13$, CI [,10; ,17]). Ab der dritten Jahrgangsstufe sind für alle Schüler der Altersgruppen bedeutsam bessere Ergebnisse im Vergleich zum Durchschnitt der Altersstichprobe festzustellen. In der achten Jahrgangsstufe haben sich die Leistungen erneut angeglichen (z. B. $z_{\text{jung-MA}} = ,14$, CI [-,07; ,35]; $z_{\text{mittel-MA}} = ,08$, CI [-,07; ,23]).

Für Schüler*innen mit deutscher Herkunftssprache zeigt sich sowohl für Deutsch / Lesen als auch für Mathematik eine signifikante leistungsbezogene Benachteiligung der jüngeren Schüler*innen zu den ersten beiden Messzeitpunkten (2. und 3. Jahrgangsstufe) im Vergleich zu den mittleren und älteren Schüler*innen. Dabei liegen die Leistungen aller Altersgruppen statistisch bedeutsam über dem Durchschnitt der Altersstichprobe (z. B. in Jahrgang 2: $z_{\text{jung-MA}} = ,08$, CI [,04; ,11]; $z_{\text{alt-MA}} = ,27$, CI [,23; ,31]). Bis zur achten Jahrgangsstufe sind diese Unterschiede ausgeglichen, sodass keine Leistungsunterschiede zwischen den Altersgruppen identifiziert werden können (z. B. $z_{\text{jung-DE}} = ,28$, CI [,14; ,42]; $z_{\text{alt-DE}} = ,13$, CI [-,02; ,28]). Bei Schüler*innen mit nichtdeutscher Herkunftssprache zeigt sich ein ähnliches Bild: hier sind die vergleichsweise Jüngeren, zusätzlich zu den deutlich schwächeren Leistungen der Schüler*innen nichtdeutscher Herkunftssprache im Allgemeinen, zunächst im Nachteil (z. B. in Jahrgang 2: $z_{\text{jung-MA}} = -,52$, CI [-,57; -,47]; $z_{\text{alt-MA}} = -,29$, CI [-,34; -,24]). Am Ende der Studie (Jahrgangsstufe 8) lassen sich, wie in den übrigen Schüler*innengruppen auch, keine bedeutsamen Unterschiede zwischen den Altersgruppen feststellen (z. B. $z_{\text{jung-DE}} = -,37$, CI [-

,58; -,15]; $z_{\text{alt-DE}} = -,59$, CI [-,79; -,39]). Die statistisch bedeutsame Leistungsbenachteiligung von Schüler*innen nichtdeutscher Herkunftssprache ist zu allen Messzeitpunkten und über alle Altersgruppen hinweg evident.

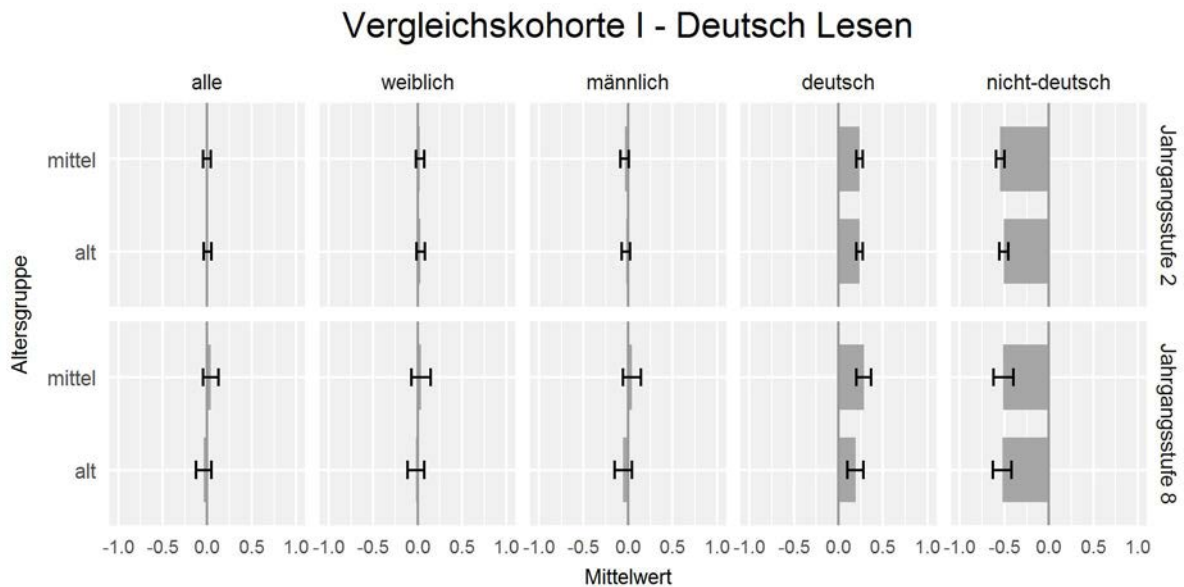


Abbildung 22: Mittlere Leistung (z-standardisiert) der Altersgruppen der Vergleichskohorte I in Deutsch / Lesen

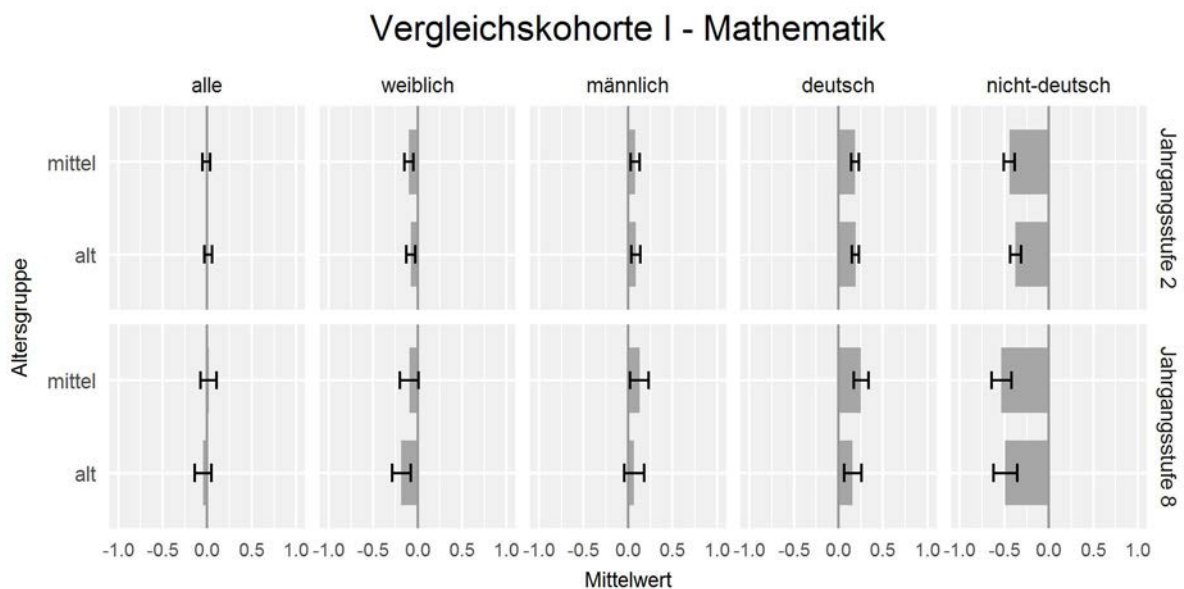


Abbildung 23: Mittlere Leistung (z-standardisiert) der Altersgruppen der Vergleichskohorte I in Mathematik

Für die Einordnung der gefundenen Gemeinsamkeiten und Unterschiede in den mittleren Leistungen für Deutsch / Lesen und Mathematik zwischen den Altersgruppen werden an dieser Stelle erneut die Vergleichskohorten herangezogen. Die Ergebnisse dieser Analysen sind in Abbildung 22 bis Abbildung 25 sowie in den Tabellen Tab. A- 13 und Tab. A- 14 dargestellt. Es

lässt sich hier für beide Kohorten und allen dargestellten Ergebnissen zusammenfassend festhalten, dass keine bedeutsamen Unterschiede in der mittleren Leistung zwischen den Altersgruppen feststellbar sind.

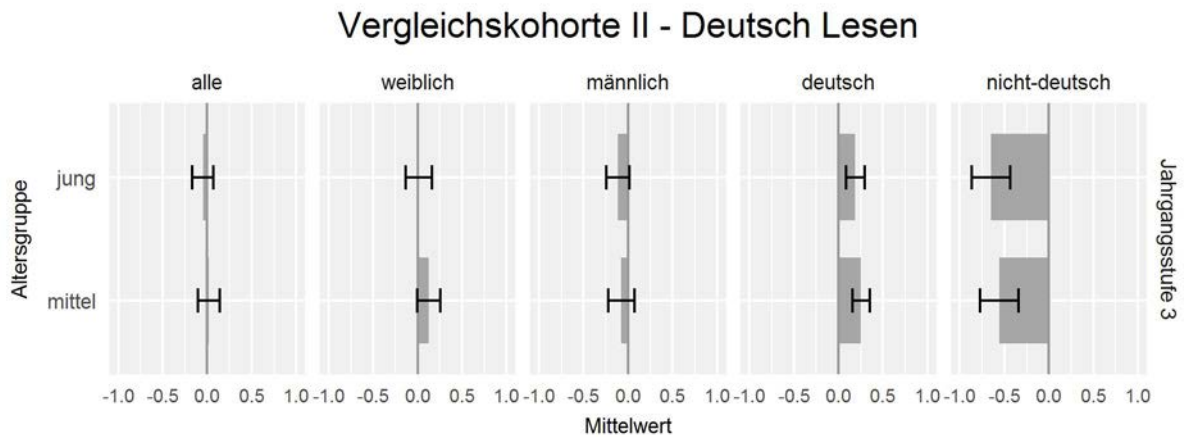


Abbildung 24: Mittlere Leistung (z-standardisiert) der Altersgruppen der Vergleichskohorte II in Deutsch / Lesen

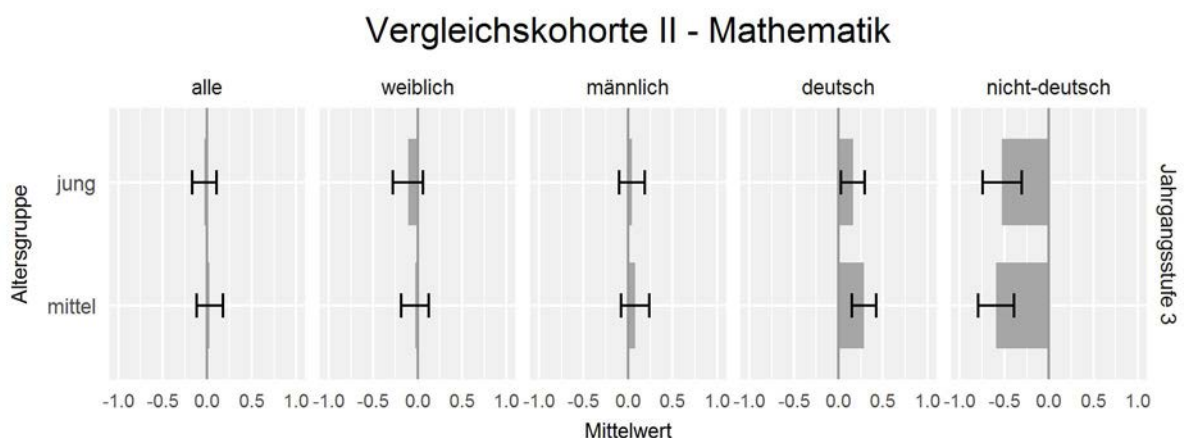


Abbildung 25: Mittlere Leistung (z-standardisiert) der Altersgruppen der Vergleichskohorte II in Mathematik

IV.3 Sind Überlappungen der altersspezifischen Leistungen festzustellen und, wenn ja, wie lassen sich diese beschreiben?

Abschließend zu den Untersuchungen zur Wirkung der Senkung des Schuleintrittsalters wurde mithilfe von Boxplots überprüft, inwieweit sich die Leistungen in Deutsch / Lesen und Mathematik je nach Geburtsmonat in ihrer Verteilung überlappen. Aus platzökonomischen Gründen und weil keine systematischen Unterschiede zwischen den Kohorten festgestellt wurden, werden im Folgenden nur die Ergebnisse für die Hauptkohorte präsentiert. Die Boxplots der beiden Vergleichskohorten finden sich im Anhang der Arbeit (Abb. A- 1 und Abb. A- 2). Die Ergebnisse für die Hauptkohorte sind in Abbildung 26 dargestellt.

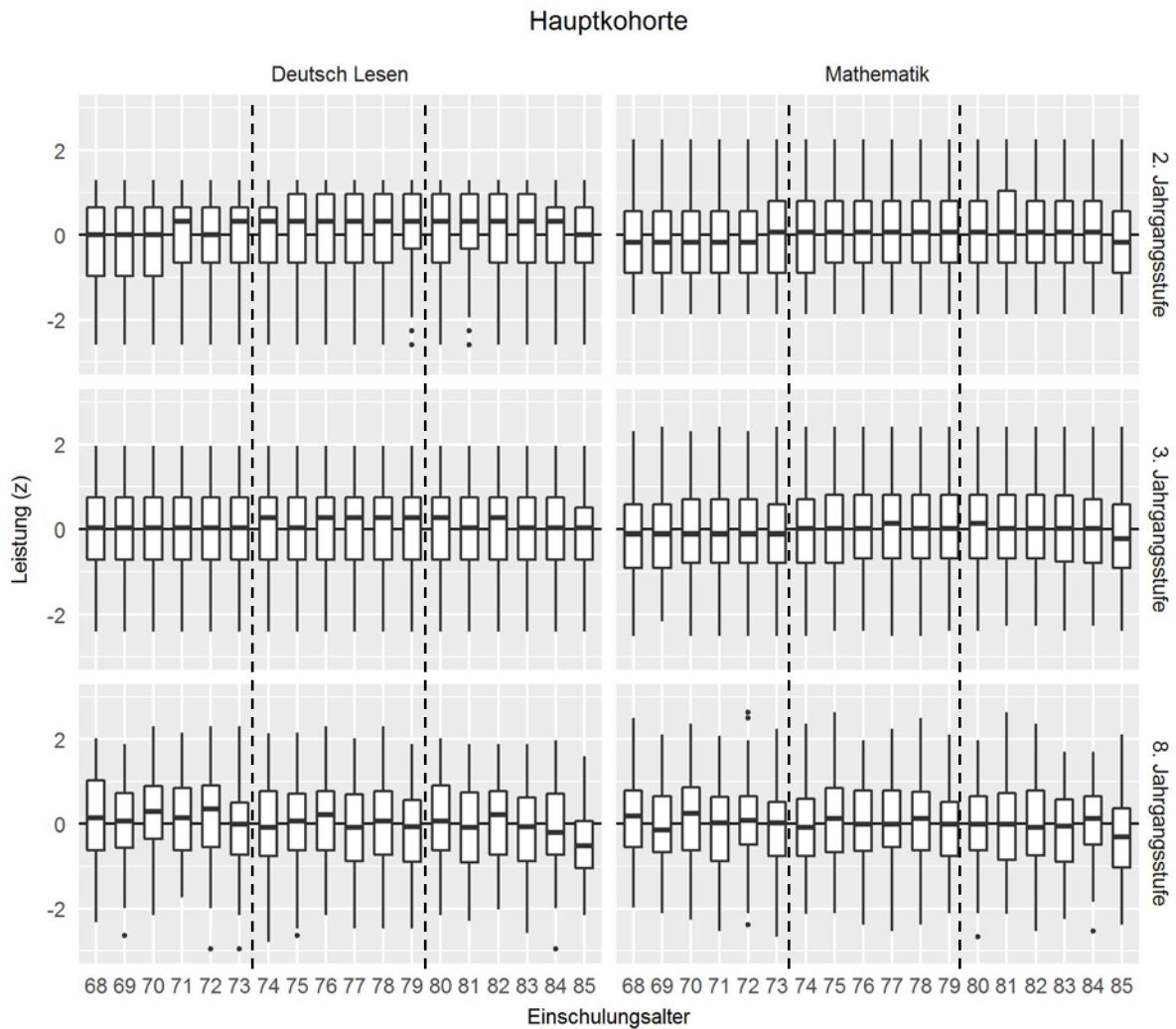


Abbildung 26: Leistungsverteilung (Boxplots) der Hauptkohorte nach Einschulungsalter (in Monaten)

Die sogenannte Box der Boxplots enthält, als fettmarkierter schwarzer Querstrich hervorgehoben, den Median. Die unteren und oberen Ränder der Box kennzeichnen das 25%- bzw. 75%-Quartil. Dieser Bereich wird auch Interquartilsabstand genannt und kennzeichnet die Leistungswerte für 50 % der Stichprobe. Die nach oben bzw. unten verlaufenden Linien oder Antennen, die sogenannten Whisker, bilden die Leistungswerte ab, die in dem Bereich liegen, der 1,5-mal so groß ist, wie der Interquartilsabstand. Alle darüber- oder darunterliegenden Leistungswerte einzelner Schüler*innen werden als Ausreißerpunkte dargestellt.

Sowohl für Deutsch/Lesen als auch für Mathematik zeigt sich insgesamt eine große Überlappung der Leistungen in den einzelnen Altersmonaten. Interessant ist, dass, obwohl die vergleichsweise Jüngeren in Deutsch/Lesen in der zweiten Jahrgangsstufe insgesamt benachteiligt sind, innerhalb dieser Altersgruppe Schüler*innen mit einem Einschulungsalter

von 71 bis 73 Monaten größtenteils ähnliche Leseleistungen zeigen, wie ihre vergleichsweise älteren Mitschüler*innen. Weiterhin fällt auf, dass in beiden Tests und über alle Jahrgangsstufen hinweg die Leistung der Schüler*innen mit einem Einschulungsalter von 85 Monaten vergleichsweise niedriger ausfallen. Abschließend ist festzustellen, dass sich die Leistungen in der achten Jahrgangsstufe als deutlich heterogener zwischen den Einschulungsmonaten darstellen, was auf die geringe Stichprobengröße der Erweiterungsstudien zurückgeführt werden kann.

Insgesamt lassen die Boxplots bereits erkennen, dass im Rahmen dieser Studie das Alter eines Kindes nicht systematisch mit einer besseren oder schlechteren Leistung zusammenhängt. Mithilfe des Determinationskoeffizienten R^2 kann diese Annahme bestätigt werden. Tabelle 24 zeigt, dass maximal 1,1 % der Varianz der Schüler*innenleistung auf das Einschulungsalter zurückgeführt werden können.

Tabelle 24: Anteil der durch das Alter eines Kindes aufgeklärten Varianz an der Varianz schulischer Leistungen

Varianz- aufklärung in %	Deutsch / Lesen					Mathematik				
	alle	weiblich	männlich	deutsch	nicht- deutsch	alle	weiblich	männlich	deutsch	nicht- deutsch
Vergleichskohorte I										
Jahrgang 2	0	0	0	0	0	0	0	0	0	0
Jahrgang 8	0,3	0,1	0,3	0,3	0	0,2	0,2	0,2	0,3	0
Hauptkohorte										
Jahrgang 2	0,5	0,5	0,7	0,5	1,1	0,7	0,5	0,8	0,8	0,7
Jahrgang 3	0,1	0,1	0	0,1	0,2	0,1	0,1	0,1	0,2	0,2
Jahrgang 8	0,6	0,6	0,4	0,5	0,6	0,1	0,4	0	0,1	0
Vergleichskohorte II										
Jahrgang 3	0	0	0	0	0	0	0,1	0	0,3	0

5.2.5 Zusammenfassung

Zur Prüfung der Wirkung der Stichtagsverschiebung wurden zunächst die Anteile der drei Altersgruppen über die Messzeitpunkte und Kohorten hinweg betrachtet. Dabei wurden keine praktisch bedeutsamen Verschiebungen festgestellt. Weder für die gesamten Altersstichproben noch für einzelne Schüler*innengruppen (Mädchen und Schüler*innen nichtdeutscher Herkunftssprache). Es gibt demnach keine Hinweise darauf, dass Schüler*innen bestimmter Altersgruppen besonders häufig von Klassenwiederholungen oder vorzeitigem Vorrücken betroffen waren oder aus sonstigen Gründen ihren Klassenverband verlassen haben. Um dies zweifelsfrei zu prüfen, müssten jedoch die Daten auf Individualebene vorliegen.

Auf diese Untersuchung aufbauend, wurde im Folgenden geprüft, ob es Hinweise auf Ungleichverteilung der einzelnen Altersgruppen auf die weiterführenden Schulen gibt. Es wurden Gymnasien und Integrierte Sekundarschulen gegenübergestellt. Dabei zeigten sich erneut keine Ungleichverteilungen für die gesamten Altersstichproben oder für einzelne Schüler*innengruppen (Mädchen und Schüler*innen nichtdeutscher Herkunftssprache).

Im Anschluss an die Betrachtung der Altersgruppenanteile, wurden die Leistungen der einzelnen Altersgruppen relativ zueinander betrachtet (relative Alterseffekte). Zunächst wurden für die Hauptkohorte auf der Ebene aller Schüler*innen in den Jahrgangsstufen 2 und 3 sowohl für Deutsch / Lesen als auch für Mathematik leistungsbezogene Nachteile der jüngsten Altersgruppen festgestellt. In der achten Jahrgangsstufe haben sich die Leistungen der drei Altersgruppen in beiden Lernbereichen angeglichen. In den Schüler*innengruppen zeigten sich die Stärken der Mädchen im Leseverstehen sowie die Stärken der Jungen in Mathematik (vgl. u.a. Hannover & Kessels, 2011) besonders in den mittleren und älteren Altersgruppen. Auf der anderen Seite sind es im Lesen vor allem die jüngsten Jungen und in Mathematik die jüngsten Mädchen, die vergleichsweise schlechtere Leistungen zeigen. Doch haben sich die Leistungen in der achten Jahrgangsstufe soweit angeglichen, dass keine bedeutsamen Unterschiede zwischen den Altersgruppen festgestellt werden können. Bei der Gegenüberstellung von Schüler*innen mit deutscher und nichtdeutscher Herkunftssprache zeigen sich die Ergebnisse analog: dort, wo leistungsbezogene Stärken auszumachen sind (Schüler*innen mit deutscher Herkunftssprache), zeigen sich die Vorteile zu den ersten Messzeitpunkten vor allem bei den mittleren und älteren Schüler*innen, wo Schwächen festgestellt werden (Schüler*innen mit nichtdeutscher Herkunftssprache), sind es vor allem die Jüngeren, die stärker benachteiligt sind. Auch hier sind für die achte Jahrgangsstufe keine Unterschiede zwischen den Altersgruppen festzustellen. Zwischen den Altersgruppen der Vergleichskohorten konnten keine statistisch bedeutsamen Unterschiede in der Leistung ausgemacht werden.

Die abschließende Betrachtung der Leistungsverteilungen entsprechend dem Einschulungsalter (in Monaten) zeigt darüber hinaus, dass es selbst zwischen den Jüngsten und Ältesten der Hauptkohorte große Überschneidungen in der Leistung gibt. Hierbei ist jedoch zu beachten, dass vor allem die Gruppe der 83–85 Monate alten Schüler*innen leicht negativ verzerrt ist, da Schüler*innen, die vorzeitig eingeschult wurden (und in der Regel leistungsstark sind), in diesen Gruppen fehlen. Da im Schuljahr 2004/05, vor der Einschulung der Hauptkohorte, 13 % der Schüler*innen vorzeitig eingeschult wurden, ist der Einfluss auf die altersbezogenen Leistungen an dieser Stelle nicht zu unterschätzen. Dennoch zeigt auch die Zerlegung der

Varianz mithilfe von R^2 , dass das Einschulungsalter nicht sehr viel zur Aufklärung der Leistungsvarianz beiträgt.

Insgesamt bestätigt die Studie IV, zumindest teilweise, die leistungsbezogene Benachteiligung vergleichsweise jüngerer Schüler*innen zu Beginn der Schullaufbahn. Auf der anderen Seite gibt es keine Anzeichen dafür, dass diese Unterschiede sich langfristig auf (a) die Art der besuchten weiterführenden Schule auswirken oder (b) sich diese Unterschiede über die Schulzeit manifestieren. Weiterhin liefern die Ergebnisse, bezogen die Reform der Einschulung in Berlin, Hinweise darauf, dass es gelungen ist, einen Großteil der (jüngeren) Schüler*innen zu integrieren und zu fördern.

5.3 Studie V: Relative Age Effects in Mathematics and Reading: Investigating the Generalizability across Students, Time, and Classes

Anmerkung: Die Ergebnisse dieser Studie wurden in Form eines Fachartikels in der Zeitschrift *Frontiers in Psychology* veröffentlicht. Der Artikel ist im Folgenden in Form der finalen Version der Erstautorin eingefügt.

Thoren, K., Heinig, E. & Brunner, M. (2016). Relative Age Effects in Mathematics and Reading: Investigating the Generalizability across Students, Time, and Classes. *Front. Psychol.* 7:679. doi: 10.3389/fpsyg.2016.00679

5.3.1 Introduction

Effects of age on various outcomes such as academic achievement (Fertig & Kluve, 2005; Puhani & Weber, 2005), educational attainment (e.g. Angrist & Krueger, 1992; Black et al., 2011), or earned wages (e.g. Bedard & Dhuey, 2006; Mayer & Knutson, 1999) have been addressed in many studies in education and economic research. One widely-recognized way of investigating this question is to study how the relative position of a child with regard to her or his age has an effect on achievement when the average age of a class or learning group is higher or lower (Gold et al., 2012). A major advantage in focusing on these so-called relative age effects is that here, one makes use of the naturally occurring variation of birth months among students who are enrolled in school at the same time. In most OECD countries, education administrations define one cut-off date for school enrollment, stipulating that children who have reached a defined age (usually 6 or 7 years) by that date are eligible for school. This procedure usually results in an age range of up to 12 months within a school entrance cohort that coincides with differences in maturity as well as different learning experiences prior to school enrollment in preschool or in the family. In fact, studies have shown that relatively older students might show between .45 and .49 of a standard deviation better results in literacy at the beginning of formal schooling (Crone & Whitehurst, 1999; Gold et al., 2012). These findings illustrate the large heterogeneity among students in one class, particularly in the early elementary school grades. Accordingly, the importance of studying the relationship of relative age on educational outcomes also results from the fact that it affects children and teachers in their everyday lives.

The overarching goal of the present study is to address the generalizability of relative age effects on students' achievement. To this end, we study relative age effects (a) across two important academic domains (reading and mathematics), (b) in different subgroups of students by contrasting general effects with effects among students with and without immigrant backgrounds, (c) across time by investigating relative age effects for different school entrance cohorts (2004, 2005, and 2011), and (d) across classes by analyzing whether our results are applicable to different learning groups or classes. To this end, we capitalize on a unique set of population and large-scale data from students in the capital of Germany: Berlin.

5.3.2 Literature Review

5.3.2.1 Relative Age Effects

Most of the studies analyzing achievement related relative age effects at the beginning of formal schooling have found that relatively older students perform better academically than relatively younger students: For German students, Gold and colleagues (2012) find effects in favor of older students for reading in Grade 1 but similar achievement already by Grade 2. A study by Kawaguchi (2011) states effects in favor of older students in Grade 4 mathematics in Japan. However, a study by Crone and Whitehurst (1999) for the United States finds relative age effects in favor of the relatively older in literacy skills at the beginning of formal schooling, but no relative age effects for reading skills by the time students reach the end of first grade. Analyzing relative age effects in secondary school, most studies have found effects favoring the older: Cogley and colleagues (2009) report positive effects for the relatively older for mathematics but not for English in Grades 7, 8, and 9 in England. Strom (2004) finds these effects for reading among 15-year-olds in Norway, Fredriksson & Öckert (2005) for mathematics, Swedish and other domains in Grade 9 in Sweden, and Kawaguchi (2011) for mathematics in Grade 8 in Japan). There is, however, also evidence from the United States that relative age effects might vanish or even reverse in favor of the relatively younger in the course of students' school careers (Lincove & Painter, 2006- using a composite score of reading and mathematics). In summary, findings with relative age effects in favor of relatively older students dominate at the beginning of formal schooling. Over the course of primary and secondary education, evidence for the persistence of relative age effects seems to be mixed. Nonetheless, most studies that focus on long-term outcomes state at least decreasing effects (Stipek, 2002).

What mechanisms may explain differences in achievement between relatively younger and older children? Differences in cognitive development between younger and older children already occur before the beginning of formal schooling (Musch & Grondin, 2001) because older children (a) are simply older and on average more mature (Bedard & Dhuey, 2008; Stipek, 2002), and (b) will have had more time to receive institutionalized support while attending preschool facilities (Gold, et al. 2012). After school entry these differences can persist or even increase because (c) the curriculum typically targets the average developmental state of students in one grade, which might be to the disadvantage for the youngest of the class (Elder & Lubotsky, 2009). Moreover (d) teachers' behavior in the classroom might differ for older compared to younger children. For instance, they might have higher expectations of older

children, thus challenging and motivating them more, which results in more positive learning experiences and a steeper learning curve (Hannover & Kessels, 2011; Stipek, 2002). Throughout an academic career these differences might accumulate into long-lasting and stable disadvantages (e.g. lower educational attainment or labor wages) for those who started school when they were the relatively youngest in a class.

5.3.2.2 Generalizability of Relative Age Effects across ...

The previous section dealt with relative age effects that refer to general student populations. An important issue when studying relative age effects is to examine their external validity by addressing the question of how far findings can be generalized. In doing so, it is possible to draw a much more differentiated picture and gain a deeper understanding of the factors that influence the relationship between students' relative age and their achievement. In the present study, we investigate the generalizability of relative age effects by focusing on the extent to which relative age effects hold over different student groups, time, and classes.

5.3.2.2.1 Students

Given the importance of relative age effects for educational research and policy, it is surprising that only a few studies have investigated the generalizability of relative age effects across different subgroups. Furthermore, these previous studies merely focused on differences between females and males (Black et al., 2011; Kawaguchi, 2011). However, to the best of our knowledge there are no studies on differences in relative age effects for students with and without immigrant backgrounds. Yet, this subgroup is particularly interesting and relevant since achievement and education related differences between students with and without immigrant backgrounds are well-documented internationally (OECD, 2006). Specifically, substantial achievement gaps were found throughout formal schooling: in primary (Bos, Tarelli et al., 2012; Bos, Wendt et al., 2012; Kristen, 2006; Reardon & Galindo, 2009; Schnepf, 2004) as well as in secondary education (OECD, 2006, OECD, 2010a; 2010b; Prenzel, Blum & Klieme, 2015) and at the end of formal schooling (e.g. Diefenbach, 2009; Kristen & Granato, 2007). Possible explanations for these differences might be that children from immigrant families attend preschool facilities (such as kindergarten or formal and informal day care) less often and start institutional daycare at an older age compared to non-immigrant children (Schober & Spieß, 2012; Schober & Stahl, 2014). Furthermore, immigrant children are more likely to attend lower-quality facilities (Stahl, 2015) and, even among those not attending a preschool facility, more

likely to experience lower home learning quality (OECD, 2011). The same principles apply to formal schooling facilities: Students with immigrant backgrounds are more likely to attend schools with fewer resources and a high proportion of immigrant students, often in tandem with a high percentage of students from families of lower socioeconomic status. They therefore experience disadvantages with regard to learning opportunities (OECD, 2006, OECD, 2007, OECD, 2010b, OECD, 2010c, OECD, 2015). To conclude, relative age effects may be compensated in high quality learning environments. As students with immigrant backgrounds are more likely to experience lower quality learning environments in preschool and at school, relative age effects among students with immigrant backgrounds may be more pronounced compared to students with nonimmigrant backgrounds. This prediction, however, has so far not been empirically examined.

5.3.2.2.2 Time

The question of the generalizability of relative age effects across time can be addressed (a) by longitudinally following the same cohort and studying possible changes in relative age effects as the cohort progresses through different grades, or (b) by studying possible differences among different cohorts who experience differing environmental conditions. So far, most relative age research addressing the aspect of time has taken the former approach (see Section 2.1; e.g., Crone & Whitehurst, 1999; Gold et al., 2012; Lincove & Painter, 2006). Research on the generalizability across time focusing on differences between cohorts, however, is sparse. In a study by Elder and Lubotsky (2009), a postponement of the cut-off date for enrollment by three months (increasing cohorts' average age) simultaneously increased academic achievement and the likelihood of repeating a grade. In this article we analyze relative age effects across different cohorts because students in each cohort were affected by different school enrollment laws. Thus, comparing relative age effects across student cohorts provides important information for policy makers and educational administrators on how school enrollment regulations may influence these effects.

5.3.2.2.3 Classes

Investigating the generalizability of relationships across schools and classes has become a popular and well-accepted empirical approach in current education research (e.g. Cools, Fraine, van den Noortgate & Onghena, 2009; Hill & Rowe, 1996; Maag Merki, Emmerich & Holmeier, 2015). In addition, a great deal of research has shown the effects of academic learning environments, such as the school or class a student attends, on important educational outcomes

(e.g. for achievement: OECD, 2010b, OECD, 2015). Especially the class has been found to be a particularly important factor in research on students' outcomes (e.g. Martínez, 2012) because it is at this level that the differences in teaching quality, the implementation of school curricula, and thus, the handling of heterogeneous student characteristics (e.g. age) is most evident. Thus, it seems plausible that relative age effects may vary between classes, with some classes showing stronger and some classes showing weaker relative age effects.

5.3.3 Research Objectives and Hypotheses

Relative age effects have often been investigated for the general student population in various grade levels. The generalizability of relative age effects, however, has not been fully explored. The aim of the current study is therefore to investigate achievement-related relative age effects and their generalizability across student subgroups, time, and classes. To this end, we pursue the following four research objectives. (1) We examine whether relative age effects in the general student population (as found in several previous studies) can be replicated with our data. Following the previously summarized literature, we expect that substantial relative age effects in favor of relatively older students will be found in the early grades of elementary school. Further, we expect that relative age effects will get smaller as children progress through elementary and secondary school and as relatively younger students get the chance to catch up with their relatively older peers (see Stipek, 2002). Moreover, we address three key questions on the generalizability of relative age effects. (2) In examining the generalizability across different subgroups of students, we compare general relative age effects and specific relative age effects for students with and without immigrant backgrounds. Here, we expect to find substantial differences in relative age effects between students with and without immigrant backgrounds, with relatively young immigrant students being particularly disadvantaged. In addition to the fact that they, like all children their age, are less mature and have had less time to attend preschool facilities, this group of students is further characterized by their differences with regard to education and achievement (see 5.3.2.2.1). (3) In order to answer the question of the generalizability of effects across time, we survey differences in relative age effects across three different student cohorts who entered school before and after state-wide reforms on school enrollment age were implemented. In doing so, we can explore whether relative age effects change (a) when the age range of a school entrance cohort is 18 in comparison to 12 months and (b) when the average age of a cohort is higher or lower. (4) Finally, given that teachers differ in their teaching quality, the implementation of school curricula, and the handling of heterogeneity amongst their students, we explore whether relative age effects are applicable to

all school classes or whether students in some classes show different relative age effects than students in other classes.

5.3.4 Materials and Methods

5.3.4.1 Sample

In our analyses, we first capitalize on representative large-scale data sets from one cohort of students who were obliged to enter public elementary school on August 1, 2005 according to the Berlin school law (see “TC” in Figure 1). In the present analyses, we have deliberately excluded any students from our analyses who entered school early or late because with the available archive data we could not control for possible selection bias among these students (e.g., influences as a result of unobserved heterogeneity due to voluntarily early enrollment or retention upon parental request). We paid special attention to this cohort, which we call the Target Cohort, because this cohort was the first to be affected by a revision of the Berlin law of school enrollment. Here, the Berlin Senate preponed the cut-off date, the date until which a child must reach 6 years of age in order to be eligible for school entry, from June, 30 to December, 31. This cohort is exceptional because it comprises three clearly distinguishable student age groups. (1) Children who were enrolled only according to the reformed (new) school entry law. These children were born between July and December 1999, and would have otherwise been enrolled a year later. We refer to this age group as young: children in this group were between 68 and 73 months old by the time they started school. (2) Children who would have been enrolled under the old and the new law of school enrollment. These children were born between January and June 1999. We refer to this age group as intermediate: children in this age group were between 74 and 79 months old at the time they started school. (3) Children who would have been enrolled according to the expiring (old) school entry law. These children were born between July and December 1998. We refer to this age group as old: students in this group were between 80 and 85 months of age at school entrance.

Taken together, the Target Cohort includes students with an average age of 6.4 years and an age range of 18 months. (Under the old law, a cohort to be enrolled at school had a mean age of 6.7 years and an age range of 12 months). Importantly, in studying the Target Cohort we have a unique opportunity to compare those students who would not have entered school yet if the reform had not been implemented with those students who would have entered school even if the reform had not been implemented, both being taught in the same learning environment. For

the Target Cohort, we were able to access archive data from Grade 2 (in 2007), Grade 3 (in 2008), and Grade 8 (in 2013).

In order to investigate the generalizability of effects across time for different student cohorts, we examined two further cohorts (see “CI” and “CII” in Figure 1). We call them comparison cohorts. Comparison Cohort I includes students enrolled one year before the cut-off date was changed and thus comprises students with an average age of 6.7 years, with students being 74 to 85 months old when entering school (i.e., with an age range of 12 months between the youngest and the oldest children). Hence, students in Comparison Cohort I belong either to the intermediate or the old age group (as defined for the Target Cohort). For Comparison Cohort I, archive data from Grades 2 (in 2006) and 8 (in 2012) were available. Comparisons between the Target Cohort and Comparison Cohort I give insights as to whether the reform in the age of school enrollment influenced the size of relative age effects.

Comparison Cohort II, in turn, consists of students enrolled in 2011, 6 years after the new law of school enrollment became effective. Accordingly, students in this cohort were between 68 and 79 months old when they entered school, with an average age of 6.1 years and an age range of 12 months between the youngest and the oldest children. Hence, students in Comparison Cohort II belong either to the young or intermediate age group (as defined for the Target Cohort). For Comparison Cohort II archive data from Grade 3 (in 2013) was available. Comparing Cohort II with the Target Cohort allows us to test the long-term stability of relative age effects in Grade 3. For all three cohorts, we draw on archive data, which were collected in order to allow cross-sectional analyses of each cohort. Thus, we could not track individual students across time.

To determine a student’s relative age, we were able to access data on age for all students of the Target Cohort in Grade 2 and Grade 3, and for students in Comparison Cohort I in Grade 2. However, information on age was not included in the archive data for some cohorts at some grade levels. We therefore drew random subsamples of students to collect these data for students in the Target Cohort and Comparison Cohort I at Grade 8, and for students in Comparison Cohort II in Grade 3 (see Table 1).

5.3.4.2 Instruments

Data for all three student cohorts stem from state-wide proficiency assessment programs where student participation was compulsory. In all programs, students worked on tests assessing proficiency in mathematics and reading comprehension. The mathematics and reading tests were based on extensive pilot studies and professional development by practitioners and psychometricians. Reliability estimates for students' proficiency scores were acceptable to good: Cronbach's Alpha for Grade 2 (Cronbach, 1951) ranged from $\alpha = .80$ to $\alpha = .82$. Because the tests in Grades 3 and 8 were scaled using item response theory (IRT; Wu, Adams & Wilson, 1997), we use IRT-based estimates of reliability (Rost, 1996). These estimates ranged between .60 and .89. Note that the Berlin Ministry of Education also changed the grade levels when students take part in the assessment programs. This implies that students belonging to the Target Cohort participated in Grades 2, 3, and 8; students in Comparison Cohort I participated in Grades 2 and 8; students in Comparison Cohort II participated in Grade 3.

To assess students' immigrant backgrounds we referred to student's language background of origin or the family's lingua franca. Unfortunately, as we used archive data it was not possible to apply the same criterion for all cohorts in all grade levels. However, in data sets where both attributes were recorded at the same time the correlation coefficient exceeded $r = .80$, pointing to a large overlap between these attributes. Proportions of students with immigrant backgrounds ranged from 26.6 % in Grade 3 of the Target Cohort up to 32.5% in Grade 2 of Comparison Cohort I and the Target Cohort.

5.3.4.3 Analytic Strategy

To study relative age effects, we used age groups (as defined above) as independent variables. The intermediate group served as the reference group because children in this group would have been enrolled under both the old and the new school law. Defining these age groups in all three student cohorts makes it possible to (a) study relative age effects within each student cohort and (b) to examine how these effects vary across cohorts that (b) entered school before (Comparison Cohort I) or (c) after the new law became effective (Comparison Cohort II). We defined groups with an age range of six months to balance substantive, theoretical and methodological requirements: (a) the data that we used for the present analyses included students who were enrolled only according to the old (old) or the new (young) law of enrollment and students who would have enrolled according to either law (intermediate). Thus, the definition of three age

groups corresponds perfectly to the grouping of students as implied by the effect of the new enrollment law. (b) Pooling students in larger age groups is an accepted procedure in the research on relative age effects (e.g. Cobley et al., 2009; Stipek & Byler, 2001): A grouping procedure allows for the study of the potential non-linearity of relative age effects, whereas correlation-based estimates, for example, assume a linear relation between age and achievement across the whole age range. (c) We also estimated relative age effects in subsamples of students with and without immigrant backgrounds. Thus, age groups with an age range of six months comprise a sufficient number of students to estimate relative age effects with high precision (i.e., with small standard errors).

For the analyses contained in this paper, achievement scores were z-standardized (with $M = 0$, $SD = 1$) for mathematics and reading and for each cohort in each grade level, respectively. Doing so allowed us to compare relative age effects across cohorts, grade levels, and domains. To account for possible sampling error within the subsamples of schools for which additional information on age was collected, we (a) applied formal stratified random sampling (Shadish et al., 2002, S. 342) including achievement, district, and school type as strata and (b) weighted the collected data including population information on district, gender, migration background, and school type using the survey package (Version 3.30-3; Lumley, 2014) in R (R Core Team, 2014). We chose class (and not school) as the clustering variable because relative age effects for an individual may depend on the instructional quality he or she experiences in his or her class (e.g., individual learning support).

To study relative age effects, we ran random intercept models for each time point and outcome variable to estimate standardized mean differences (SMDs) between age groups. We chose SMDs because they are conceptually very similar to standardized effect sizes such as Cohens d (Cohen, 1992). Yet, in contrast to standardized effect sizes like Cohens d SMDs can be directly estimated within a multilevel framework that we used in the present study. Importantly, the advantages of standardized effects sizes also apply to SMDs. More specifically, standardized effects sizes can be used to compare relative age effects across samples, study designs, and statistical methods (even when different measures were used); they can also be used to evaluate the practical relevance of effects, and they provide key information for power analyses and meta-analyses (see Wilkinson, 1999). We used the multilevel framework as implemented in the lme4-Package (Version 1.1-8; Bates et al., 2015) in R (R Core Team, 2014) to estimate SMDs and their standard errors. This procedure accounts for the clustering in the data (students within classes), which is necessary to compute correct standard errors. In order to estimate the

precision of the results (Cumming & Finch, 2001; Wilkinson, 1999), 95% confidence intervals (CI) were computed using the `confint`-function in `lme4`. Accordingly, we specified class as the cluster variable for each class j , achievement in mathematics or reading as the dependent variable for each student i (γ_{ij}), and membership in the young ($young_{ij}$) or old (old_{ij}) age group (dummy-coded with the intermediate group as reference group) as independent variables (1); γ_{00} defines the intercept of the regression line and r_{ij} and u_{0j} represent random error terms on Levels 1 and 2, respectively. Doing so yielded Model Set 1a:

$$\gamma_{ij} = \gamma_{00} + \gamma_{10}old_{ij} + \gamma_0young_{ij} + u_{0j} + r_{ij} \quad (1a)$$

To compute the SMD between age groups young and old, and intermediate ($\gamma_0intermediate_{ij}$) and old, a corresponding Model Set 1b was specified (with the old group as reference group):

$$\gamma_{ij} = \gamma_{00} + \gamma_{10}young_{ij} + \gamma_0intermediate_{ij} + u_{0j} + r_{ij} \quad (1b)$$

We applied Model Set 1a and 1b to various samples in order to investigate our research objectives. For Research Objective 1, we drew on the overall sample of the Target Cohort. For Research Objective 2 we used the information on students' language spoken at home to subdivide the overall sample of the Target Cohort into a subsample of students with immigrant backgrounds and a subsample without immigrant backgrounds. We applied Models 1 and 2 to both subsamples separately. For Research Objective 3, we drew on data from Comparison Cohorts I and II, respectively. In contrast to the Target Cohort, however, for Comparison Cohorts I and II we did not introduce a dummy variable for the young group and old group, respectively, because these groups did not exist in these cohorts.

To address Research Objective 4 on the generalizability of relative age effects across classes, we first ran random-intercept models on the basis of the Target Cohort as specified in Equation 1 (see above):

$$\gamma_{ij} = \gamma_{00} + \gamma_{10}old_{ij} + \gamma_0young_{ij} + u_{0j} + r_{ij} \quad (2.1)$$

Subsequently, we allowed the age effect for the old group to vary across classes in Model Set 2 and estimated a variance component for the regression coefficient u_{1j} :

$$\gamma_{ij} = \gamma_{00} + \gamma_{10}old_{ij} + \gamma_0young_{ij} + u_{0j} + u_{1j}old_{ij} + r_{ij} \quad (2.2)$$

In Model Set 3, we additionally allowed the age effect for the young group to vary across classes and estimated a variance component for the regression coefficient u_{2j} :

$$\gamma_{ij} = \gamma_{00} + \gamma_{10}old_{ij} + \gamma_{01}young_{ij} + u_{0j} + u_{1j}old_{ij} + u_{2j}young_{ij} + r_{ij} \quad (2.3)$$

By executing deviance tests that inspected whether a model fits the data better than its preceding model, we examined whether the relative age effects applied for all classes in a similar way.

5.3.5 Results

5.3.5.1 Descriptive Statistics

Before proceeding to the results for our research questions, we are going to take a short look at the distribution of the three age groups. Although we aimed to keep selection bias at a minimum by referring to only those students who were obliged to enter school in a certain year (see Section 4.1), we found that the group of relatively older students includes a smaller number of students on average in comparison to the other age groups. This is due to the fact that some parents voluntarily enrolled their child a year earlier. In the year before the Target Cohort entered school, approximately 5.5 % of the children were enrolled early (Autorengruppe Bildungsberichterstattung, 2012), meaning that the group of relatively older students in the Target Cohort is to some extent selected (e.g., because these students might have shown exceptionally good precursory verbal or mathematical skills or appeared to be quite mature for their age). For the year before Comparison Cohort I started school, approximately 3% of children were enrolled early. In Comparison Cohort II, around 1% of students were enrolled early (Autorengruppe Bildungsberichterstattung, 2012).

For students with immigrant backgrounds (column MigB in Table 2), the proportions are distributed similarly, with the proportion of the old within the students with immigrant backgrounds being a little bit bigger than in the group of all students. When focusing on changes in the proportions over time, we found that it is the group of the young and especially the young with immigrant backgrounds whose percentages decrease the most over the grades (from 35.5% in Grade 2 to 30.2% in Grade 8 for the Target Cohort, Table 2).

5.3.5.2 Relative Age Effects

Table 3 (“*Target Cohort*”) displays the relative age effects for the Target Cohort. The results for reading in Grade 2 show effects to the advantage of the relatively older students. The effect is largest between young and old students ($SMD_{\text{Reading}} = -.19$) and smallest between intermediate and old ($SMD_{\text{Reading}} = -.04$). For mathematics, the effects are identical, albeit more distinctive (young vs. old: $SMD_{\text{Mathematics}} = -.23$; intermediate vs. old: $SMD_{\text{Mathematics}} = -.07$).

In Grade 3, effects in reading and mathematics between young and intermediate as well as between young and old are somewhat smaller than in Grade 2, but still in favor of relatively older students (except for the comparison of the intermediate vs. old group).

Finally, in Grade 8 relative age effects have vanished in the general sample in reading and mathematics or have even been reversed. Specifically, in reading, students in the young group now outperform students in the old group ($SMD_{\text{Reading}} = .10$).

In sum, the present results show that relative age effects on reading achievement and on mathematics achievement in favor of older students are most pronounced in Grade 2 and diminish (or even reverse in favor of younger students) with increasing grade levels.

5.3.5.3 Generalizability of Relative Age Effects across ...

The following subsections tackle questions on the extent to which the results as observed for the overall sample in the Target Cohort can be generalized across students with and without immigrant backgrounds, time (i.e., different cohorts), and classes.

5.3.5.3.1 Students

Table 4 presents the results for the immigrant and the non-immigrant samples respectively. The only substantial difference between students with and without immigrant backgrounds, which we detected by way of non-overlapping confidence intervals, is for young vs. old in Grade 2 reading. Besides that, the present results suggest that the results as observed for the overall sample can also be generalized to both students with immigrant backgrounds and students without immigrant backgrounds.

5.3.5.3.2 Time

The results from the study on the generalizability of relative age effects across different cohorts are displayed in Table 3 (“*Comparison Cohort I*”, “*Comparison Cohort II*”). For Comparison Cohort II, the results from the Target Cohort were replicated in both reading and mathematics with no substantial differences in size. For Comparison Cohort I, the results were also replicated for mathematics, but not for reading. Here, there is already a zero effect in Grade 2 ($SMD_{\text{Reading}} = -.02$) and a positive effect for the relatively younger (here: intermediate) in Grade 8 ($SMD_{\text{Reading}} = .06$). However, none of the standardized mean differences in Comparison Cohorts I and II are substantially different from the general relative age effects found for the Target Cohort.

5.3.5.3.3 Classes

In our last set of analyses, we studied the generalizability of age effects across classes. The results for reading in Grade 2 (Table 5) show that a model including a random coefficient for the age group old (Model Set 2) fits the data significantly better than a model where there are only fixed effects specified (Model Set 1). However, these significant results are likely due to the large sample size, which may render even trivial differences statistically significant. Moreover, when a random coefficient is further added for the young (Model Set 3), no significant changes in model fit are found. Further, focusing on the size of effects, the close to zero variance of the random coefficients (e.g. for Grade 2: $\tau_{\text{young}}^2 = .02$ and $\tau_{\text{old}}^2 = .01$) indicates that there is no substantial variation in relative age effects across classes. These results apply for reading in all grades studied. For mathematics, the results are similar, implying no systematic variation in general relative age effects across classes.

5.3.6 Discussion

The overarching goal of the present study was to tackle key questions regarding the replicability and generalizability of relative age effects on students’ achievement with respect to (a) students with and without immigration backgrounds, (b) time in terms of reforms to school enrollment regulations, and (c) across classes. To this end, we capitalized on a quasi-experimental research situation that resulted from an educational reform as well as a unique set of population and large-scale data from students in Berlin, the capital of Germany.

5.3.6.1 Relative Age Effects

We examined whether previous empirical evidence on general relative age effects could be replicated with the overall sample for the Target Cohort. Based on previous research, we hypothesized that substantial relative age effects in favor of relatively older students would be found in the early grades of elementary school (Grade 2), as well as that these effects will get smaller (Grade 3) and eventually vanish (Grade 8) over the course of academic education. Our results generally confirm these predictions for reading and mathematics. We even found a significant reversed relative age effect favoring the young over the old in reading in Grade 8.

To further evaluate the relative age effects as found in the present study, we refer to the seminal paper by Hill, et al. (2008; see Table 1), who meta-analyzed annual achievement gains in reading and mathematics for the K-12 age range. In Grade 2, the differences between the youngest and the oldest students of that cohort can be quantified as a fifth of the average annual gain in reading and mathematics, which means that the oldest were on average ten weeks ahead in their achievement compared to the youngest. As the cohorts proceeded through Grade 3, the differences decreased to an advantage of a little more than four weeks in reading and approximately seven weeks in mathematics. In Grade 8 the only substantial relative age effect pointed to a 20-week advantage among the youngest in comparison to the oldest in reading.

5.3.6.2 Generalizability of Relative Age Effects across ...

After studying relative age effects in general we proceeded to examine their generalizability.

5.3.6.2.1 Students

We examined the generalizability of general effects across students with and without immigrant backgrounds. We hypothesized stronger relative age effects to the disadvantage of young immigrants (e.g., because of inequalities in preschool experiences). The findings provided some support for this prediction for reading in Grade 2, where the differences represent an eight-week advantage for the oldest in comparison to the youngest Germans and a 12 week advantage for the oldest in comparison to the youngest students with immigrant backgrounds. However, substantial differences in relative age effects between students with and without immigrant backgrounds were not observed with the available data for either Grade 3 or Grade 8. The general trend of the relative age effects decreasing over time is similar for all subsamples. However, particularly interesting here is that in contrast to their German peers, relative age

effects to the disadvantage of young students with immigrant backgrounds did not yet show a substantial decrease by Grade 3. This finding might indicate that younger immigrants need more time to catch up with their older peers than the younger Germans do.

5.3.6.2.2 Time

With regard to the question on the generalizability of relative age effects across time, we surveyed differences in relative age effects between three different school entrance cohorts (2004, 2005, and 2010). In doing so, we were able to explore whether relative age effects change (a) when the age range of a school entrance cohort is 18 months rather than 12 and (b) when the average age of a cohort is higher or lower. Specifically, in Comparison Cohort I, the age average was 6.7 years and the age range was 12 months; in the Target Cohort, the average age was 6.4 years with a range of 18 months; and in Comparison Cohort II, the average age was 6.1 years with a range of 12 months.

Importantly, relative age effects could be generalized across the three cohorts in mathematics. In reading, similar effects were found for the Target Cohort and Comparison Cohort II, a cohort that started school 6 years after the Target Cohort. This points to the sustainability of the findings as observed for the Target Cohort. In contrast to the Target Cohort, however, students in Comparison Cohort I showed no relative age effect in Grade 2 and a positive effect for the relatively young (here: intermediate) in Grade 8, as opposed to a negative effect for the young in Grade 2 and zero effect in Grade 8 for students in the Target Cohort. A potential explanation for the “missing” relative age effects in Grade 2 might be that the relatively younger in Comparison Cohort I are still older (and hence more mature) than the young in the Target Cohort and thus able to achieve at a similar level to their older peers.

In summary, relative age effects in mathematics and reading did not differ substantially between student cohorts and can thus be generalized to student cohorts with a higher average age and a larger age range. Moreover, the reversal of relative age effects (which was observed for both the Target Cohort and Comparison Cohort I in reading) is particularly interesting because it shows that the youngest children’s disadvantage in reading achievement in Grade 2 might turn into an advantage in Grade 8 (but see also limitations section).

5.3.6.2.3 Classes

We investigated whether relative age effects are applicable to all school classes or whether students in some classes show different relative age effects than students in other classes. Here, we expected to find substantial variation in effects between classes. Our results indicate that relative age effects did not vary substantially across classes in reading and mathematics. In other words: Even though attending a certain class or learning group is known to be an important indicator in education research (e.g. Martínez, 2012), the different learning environment within each class did not seem to substantially affect the effects of a child's relative position according to her or his age on achievement in public schools in Berlin.

5.3.6.3 Limitations

In our analyses, we examined a unique set of data that covered the whole student population or representative samples of student achievement data in reading and mathematics. However, the present data did not allow us to track individual students across time, which limited our opportunities for studying the reasons why students dropped out of the sample. It is possible that the vanishing or even the reversal of relative age effects results from younger students with a lower level of achievement being retained at lower grade levels. However, the proportion of students with each age group changed only slightly as students proceeded through their educational careers: For example, the proportion of students in the young group in the Target Cohort dropped from 35.5% in Grade 2 to 33.4% in Grade 8 (see Table 1). Thus, it seems unlikely that this explanation fully accounts for the pattern of results on relative age effects as observed in the present study. To overcome this limitation, future studies on relative age effects may benefit considerably from (a) tracking all students who enter school in a certain school year irrespective of whether they are retained at a certain grade level or not and (b) collecting data on a broad range of students' individual (e.g., early number or verbal competencies, affective-motivational variables) and family characteristics (e.g., parents' educational aspirations for their children) that are known to predict whether students linearly progress through their educational career or not. Finally, to get a fuller picture of relative age effects, it would be beneficial to include further outcome variables, particularly non-cognitive ones (e.g., social skills), because these outcomes receive a great deal of attention in the public and scientific discussion of school starting age (see Stipek, 2002).

5.3.6.4 Concluding Remarks

In sum, the results of our study confirm the pattern of findings on relative age effects as identified by the majority of previous studies: Relative age effects are relatively small at the beginning of formal schooling and largely vanish as students proceed through their educational careers. Moreover, our study has provided new evidence on the wide generalizability of these findings. Although we found a few differences, overall the present study has empirically underscored that the pattern of results as observed for the overall sample in the Target Cohort holds (a) for students with and without immigrant backgrounds; (b) for student cohorts who entered school under differing enrollment regulations, resulting in variations in the average age and age range; and (c) across school classes. Importantly, the size of these relative effects, especially in earlier grades, adds to the empirical body of knowledge on the heterogeneity that exists among students, not only but especially in elementary schools. These relative age effects are particularly relevant for teachers and policy makers as they present a challenge with regard to age-appropriate instruction and curriculum development. On the other hand, findings from secondary schools support the school's role as the "great equalizer" (e.g. Downey, Hippel & Broh, 2004; Kerstan & Spiewak, 2008), as school minimizes the effect of children's relative age in their academic future.

5.3.7 Figures

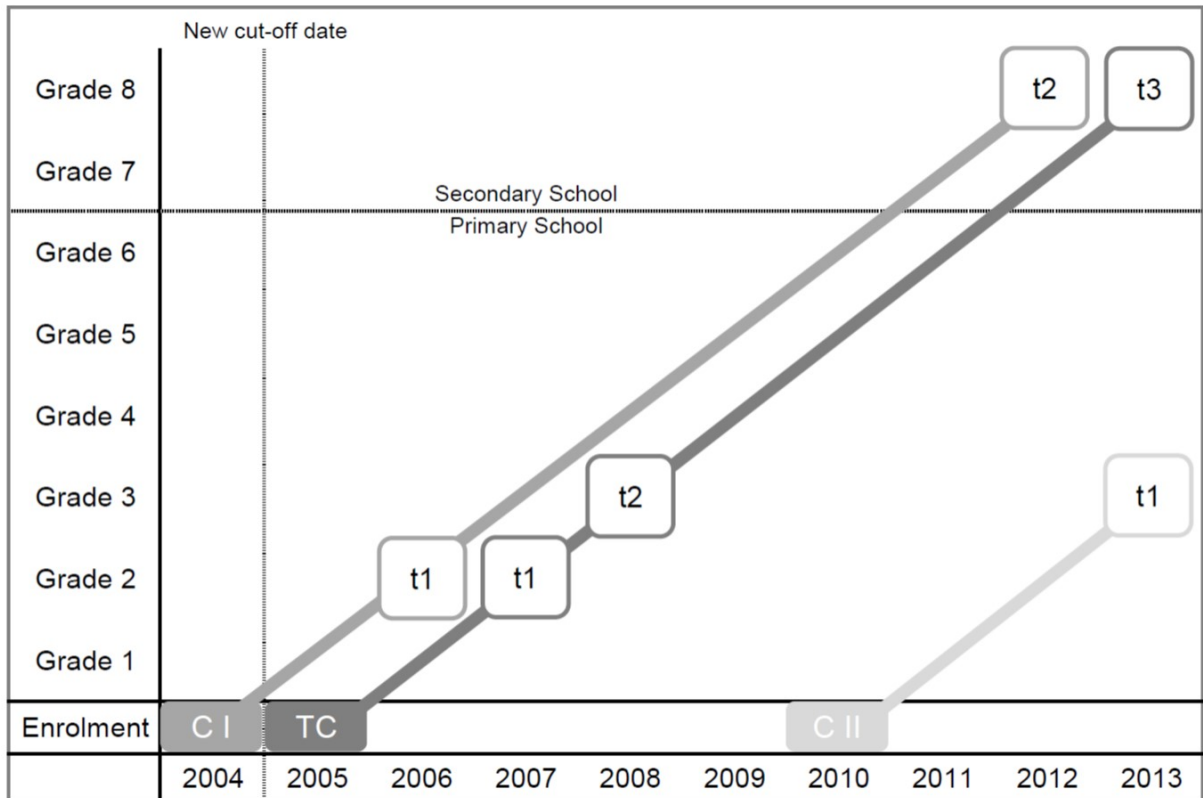


Figure 1: multi sequence cohort design

5.3.8 Tables

Table 1: Sample description by student cohort

	<i>Comparison Cohort I</i>				<i>Target Cohort</i>				<i>Comparison Cohort II</i>			
	N		Male	MB	N		Male	MB	N		Male	MB
	students	classes			students	classes			students	classes		
Grade 2	18,767	786	51.4%	32.1%	28,662	1,553	51.0%	33.0%	—	—	—	—
Grade 3	—	—	—	—	25,500	1,433	50.8%	27.3%	1,570	112	50.3%	26.6%
Grade 8	4,050	281	51.1%	32.0%	2,397	122	49.8%	28.4%	—	—	—	—

Note: % MigB = percentage of students with immigrant backgrounds.

Table 2: Proportions of the three age groups with respect to the overall student sample and within subgroups of students with and without immigrant backgrounds by student cohort

Age groups	<i>Comparison Cohort I</i>			<i>Target Cohort</i>			<i>Comparison Cohort II</i>		
	Total	no MigB	MigB	Total	no MigB	MigB	Total	no MigB	MigB
grade 2									
young	—	—	—	35.5% [34.9, 36.1]	35.6% [34.8, 36.3]	35.4% [34.4, 36.5]	—	—	—
intermediate	52.6% [51.8, 53.5]	52.9% [51.9, 53.9]	52.1% [50.8, 53.3]	36.5% [35.9, 37.1]	36.8% [36, 37.5]	35.9% [34.9, 36.9]	—	—	—
old	47.4% [46.5, 48.2]	47.1% [46.1, 48.1]	47.9% [46.7, 49.2]	28.1% [27.4, 28.7]	27.7% [27, 28.4]	28.7% [27.7, 29.6]	—	—	—
Total N	18,767	12,745	6,022	28,662	19,177	9,458			
grade 3									
young	—	—	—	33.6% [33, 34.3]	34% [33.2, 34.7]	32.9% [31.6, 34.1]	47.4% [44.9, 53.2]	47.5% [44.6, 54.2]	47.1% [42.3, 52.9]
intermediate	—	—	—	37.6% [36.9, 38.2]	37.5% [36.7, 38.2]	37.7% [36.6, 38.9]	52.6% [50.1, 50.2]	52.5% [49.6, 49.9]	52.9% [48, 52.9]
old	—	—	—	28.8% [28.2, 29.5]	28.6% [27.8, 29.3]	29.5% [28.3, 30.7]	—	—	—
Total N				25,500	18,550	6,950	1,570	1,152	418
grade 8									
young	—	—	—	31.4% [29.5, 33.4]	31.9% [29.6, 34.2]	30.3% [26.5, 33.9]	—	—	—
intermediate	51.5% [49.8, 53.2]	52.1% [50.1, 54.2]	50% [47.1, 52.9]	39.6% [37.6, 41.6]	39.3% [36.8, 41.7]	40.4% [36.7, 44.1]	—	—	—
old	48.5% [46.8, 50.2]	47.9% [45.8, 49.9]	50% [47.1, 52.9]	29% [26.9, 31]	28.7% [26.2, 31.2]	29.2% [25.7, 32.6]	—	—	—
Total N	4,050	2,755	1,295	2,397	1,717	680			

Note: Total = general sample of students; no MigB = students without immigrant backgrounds; MigB = students with immigrant backgrounds; 95% confidence intervals are displayed in square brackets

Table 3: Relative age effects (standardized mean differences) as observed for the overall student samples by student cohort

Age groups	<i>Target Cohort</i>			<i>Comparison Cohort I</i>		<i>Comparison Cohort II</i>
	grade 2	grade 3	grade 8	grade 2	grade 8	grade 3
Reading Achievement						
young vs. intermediate	-.14 [-.17, -.12]	-.09 [-.12, -.07]	.06 [-.01, .13]			-.14 [-.23, -.05]
intermediate vs. old	-.04 [-.07, -.02]	.01 [-.02, .04]	.04 [-.03, .11]	-.02 [-.05, .01]	.06 [.01, .11]	
young vs. old	-.19 [-.21, -.16]	-.09 [-.12, -.07]	.10 [.03, .17]			
Mathematics Achievement						
young vs. intermediate	-.16 [-.19, -.14]	-.12 [-.14, -.09]	.00 [-.07, .07]			-.15 [-.24, -.05]
intermediate vs. old	-.07 [-.09, -.04]	-.02 [-.04, .01]	.05 [-.02, .12]	-.04 [-.07, -.02]	-0.03 [-.13, .07]	
young vs. old	-.23 [-.25, -.2]	-.13 [-.16, -.11]	.05 [-.03, .12]			

Note: 95%-confidence intervals are displayed in square brackets. Negative values indicate that the mean achievement level of the group comprising younger students (e.g., young) is lower than that mean achievement level of the group comprising older students (e.g., intermediate).

Table 4: Relative Age Effects (Standardized Mean Differences) for students with and without immigrant backgrounds

Age groups	<i>no migration backgrounds</i>			<i>with migration backgrounds</i>		
	grade 2	grade 3	grade 8	grade 2	grade 3	grade 8
Reading Achievement						
young vs. intermediate	-.14 [-.17, -.11]	-.09 [-.12, -.06]	.04 [-.04, .12]	-.16 [-.2, -.11]	-.11 [-.16, -.06]	.09 [-.04, .21]
intermediate vs. old	-.02 [-.05, .01]	.01 [-.02, .04]	.03 [-.06, .11]	-.09 [-.14, -.04]	.00 [-.05, .05]	.08 [-.05, .21]
young vs. old	-.16 [-.19, -.13]	-.08 [-.11, -.05]	.06 [-.02, .15]	-.25 [-.3, -.2]	-.11 [-.16, -.06]	.16 [.02, .3]
Mathematics Achievement						
young vs. intermediate	-.18 [-.2, -.15]	-.12 [-.14, -.09]	-.02 [-.1, .06]	-.14 [-.17, -.1]	-.14 [-.19, -.1]	.03 [-.09, .16]
intermediate vs. old	-.04 [-.08, -.01]	-.01 [-.04, .02]	.04 [-.04, .12]	-.11 [-.15, -.07]	-.03 [-.08, .02]	.08 [-.05, .21]
young vs. old	-.22 [-.25, -.19]	-.13 [-.16, -.09]	.02 [-.07, .11]	-.24 [-.29, -.2]	-.17 [-.22, -.12]	.12 [-.02, .26]

Note: Results refer to the Target Cohort. 95% confidence intervals are displayed in square brackets. Negative values indicate that the mean achievement level of the group comprising younger students (e.g., young) is lower than that mean achievement level of the group comprising older students (e.g., intermediate).

Table 5: Relative age effects (random coefficients models) across classes

	Parameter estimates						Variance components			Model comparison (deviance tests)					
	Intercept		young		old		Intercept	young	old	compare	$\Delta\chi^2$	<i>p</i>	df	AIC	BIC
	estimate	SE	estimate	SE	estimate	SE	estimate	estimate	estimate						
Reading Achievement															
Grade 2															
Model 1	.01	.02	-.14	.01	.04	.01	.25	–	–			5	72689	72731	
Model 2	.01	.02	-.14	.01	.04	.01	.26	–	.00	M2 vs. M1	12.75	.002	7	72681	72738
Model 3	.01	.02	-.14	.01	.05	.01	.26	.02	.01	M3 vs. M2	3.86	.277	10	72683	72765
Grade 3															
Model 1	.01	.02	-.09	.01	-.01	.01	.34	–	–			5	63854	63894	
Model 2	.01	.02	-.09	.01	-.01	.01	.34	–	.02	M2 vs. M1	8.78	.012	7	63849	63906
Model 3	.01	.02	-.09	.01	-.01	.01	.34	.00	.02	M3 vs. M2	0.36	.948	10	63855	63936
Grade 8															
Model 1	-.14	.07	.11	.03	.02	.03	.53	–	–			5	6098.2	6127.7	
Model 2	-.14	.07	.11	.03	.02	.03	.56	–	.04	M2 vs. M1	10.42	.005	7	6091.8	6133.1
Model 3	-.14	.07	.12	.03	.02	.04	.59	.00	.04	M3 vs. M2	4.42	.219	10	6093.3	6152.4
Mathematics Achievement															
Grade 2															
Model 1	.00	.02	-.16	.01	.07	.01	.36	–	–			5	69244	69285	
Model 2	.00	.02	-.16	.01	.07	.01	.36	–	.00	M2 vs. M1	1.77	.412	7	69246	69303
Model 3	.00	.02	-.16	.01	.07	.01	.37	.00	.00	M3 vs. M2	0.15	.985	10	69252	69334
Grade 3															
Model 1	.01	.02	-.12	.01	.02	.01	.34	–	–			5	64056	64097	
Model 2	.01	.02	-.12	.01	.02	.01	.34	–	.01	M2 vs. M1	2.43	.297	7	64058	64115
Model 3	.01	.02	-.12	.01	.02	.01	.34	.00	.01	M3 vs. M2	0.89	.827	10	64063	64144
Grade 8															
Model 1	-.15	.07	.04	.03	-.01	.03	.55	–	–			5	5850.9	5880.3	
Model 2	-.15	.07	.04	.03	-.01	.03	.57	–	.00	M2 vs. M1	2.80	.246	7	5852.1	5893.3
Model 3	-.15	.07	.04	.04	.00	.03	.59	.02	.01	M3 vs. M2	2.11	.550	10	5856	5914.8

Note: Results refer to the Target Cohort. SE = standard error; $\Delta\chi^2$ = change in chi-square; *p* = probability of the observed deviance difference given that the deviance between models is zero in the population; df = degrees of freedom; AIC = Akaike information criterion; BIC = Bayesian information criterion

5.4 Fazit der Studien zum Schuleintrittsalter

Die Reform der Einschulungsverordnung hatte vordergründig zum Ziel, das mittlere Einschulungsalter Berliner Schüler*innen zu senken. Studie I bestätigt, dass dies zum letzten Messzeitpunkt der Studie erreicht wurde. Darüber hinaus wurde gezeigt, dass das mittlere Einschulungsalter sich nicht systematisch zwischen einzelnen Klassen unterscheidet, sodass diesbezüglich keine Verzerrungen für die folgenden Analysen zu erwarten waren. Als weiteres Kriterium zur Bestimmung des Implementationserfolgs und als Antwort auf die häufig geäußerte Sorge um die schulische Integration der nun früher eingeschulten Schüler*innen nahm Studie IV die Verteilung sowie die relative Leistung der Altersgruppen in den Blick. Dabei konnten keine bedeutsamen Verschiebungen über die Zeit festgestellt werden, sodass nicht davon auszugehen war, dass Schüler*innen entsprechend ihrem Alter häufiger oder weniger häufig ihre Jahrgangsstufe verlassen haben. Auch konnten, bezogen auf den Besuch der weiterführenden Schulen, keine Unterschiede der Altersgruppenverteilungen festgestellt werden. Weiterhin zeigten sich zum ersten Messzeitpunkt (Jahrgangsstufe 2) leistungsbezogene Benachteiligungen der relativ Jüngsten, die sich bereits ein Jahr später verringert hatten. In der achten Jahrgangsstufe konnten weder für Deutsch / Lesen noch für Mathematik relative Alterseffekte festgestellt werden. Studie V erweiterte die Erkenntnisse aus Studie IV, indem sie die Größe der relativen Alterseffekte mithilfe von Effektstärken betrachtete und deren Generalisierbarkeit über unterschiedliche Schüler*innengruppen, Messzeitpunkte und Kohorten sowie Klassen untersuchte. Dabei zeigte sich, dass die relativen Alterseffekte in Jahrgangsstufe 2 zwischen den vergleichweisen Jüngsten und Ältesten einen Unterschied von circa 10 Schulwochen (Hill et al., 2008) in Deutsch / Lesen und Mathematik ausmachten. Ein Jahr später war dieser Effekt bereits auf vier Wochen geschrumpft. Mit Blick auf die Generalisierbarkeit der relativen Alterseffekte zeigt sich zwischen Schüler*innen mit deutscher und nichtdeutscher Herkunftssprache ein ähnliches Bild mit einer etwas stärkeren Benachteiligung der relativ Jüngeren mit nichtdeutscher Herkunftssprache bis zur Jahrgangsstufe 3. Bezogen auf die drei untersuchten Kohorten, replizieren sich die Effekte über alle Kohorten in Mathematik und für Hauptkohorte und Vergleichskohorte II in Deutsch / Lesen. In Vergleichskohorte I ist in der zweiten Jahrgangsstufe 2 kein relativer Alterseffekt festzustellen. Die Ergebnisse lassen sich über alle Klassen replizieren und somit generalisieren.

Insgesamt zeigen die Ergebnisse, dass es gelungen ist, die vergleichsweise jung eingeschulerten Schüler*innen bezüglich ihrer Leistungen über die Zeit in die Schulen zu integrieren. Nichtsdestotrotz wurden, im Einklang mit bisherigen Erkenntnissen zu relativen Alterseffekten, leistungsbezogene Benachteiligungen für relativ Jüngere der Hauptkohorte und Vergleichskohorte II zu Beginn der Schullaufbahn festgestellt. Dies zeigt, dass gerade am Schulanfang, zumindest für die hier untersuchten Kohorten, die Förderung der Jüngsten noch nicht optimal gelungen ist.

6. Gesamtdiskussion

Bildungsreformen sind ein wesentlicher Bestandteil der Qualitätssicherung und -entwicklung eines Bildungssystems. Doch stellt ihre Umsetzung Schulen, Eltern und die Bildungspolitik immer wieder vor große Herausforderungen. Ziel dieser Dissertation war es, den Implementationserfolg zweier wichtiger Reformen, Jahrgangsübergreifendes Lernen und Senkung des Schuleintrittsalters, in Berliner Grundschulen erstmals umfassend zu untersuchen. Dabei wurden einerseits die Bedingungen rund um den Schulanfang sowie die Implementation sowohl theoretisch als auch anhand der Erfahrungen Berlins ausführlich dargestellt (Kapitel 1 und 2). Gerade vor dem Hintergrund der kontroversen öffentlichen und politischen Debatte um diese beiden Reformen ist die Berücksichtigung solcher Bedingungen elementar für ihre Einschätzung. Den Kern dieser Arbeit stellt die multidimensionale Evaluation des Implementationserfolgs der beiden Reformen dar, wobei die Definition von Erfolg stets an den Zielen und Erwartungen der jeweiligen Reform bestimmt wird. Die Ziele wurden hier von der Senatsverwaltung für Bildung, Jugend und Wissenschaft (SenBJW) festgelegt, Erwartungen und Ängste hingegen häufig von Schulen und Eltern formuliert.

Zunächst sollen die Ergebnisse aus den fünf Studien zusammenfassend dargestellt und vor dem Hintergrund der Kenntnisse rund um den Schulanfang und der Implementationsforschung diskutiert werden (Abschnitt 6.1). Daran anschließend werden die Limitationen der vorliegenden Arbeit betrachtet (Abschnitt 6.2) und abschließend Implikationen für pädagogische Praxis (Abschnitt 6.3) und die weitere Forschung (Abschnitt 6.4) formuliert. Das abschließende Fazit (Abschnitt 6.5) fasst die wesentlichen Erkenntnisse der Arbeit insgesamt zusammen.

6.1 Zusammenfassende Darstellung und Diskussion der Ergebnisse

6.1.1 Studie I: Wie wurde das Jahrgangsübergreifende Lernen an Berliner Grundschulen umgesetzt?

Die erste Studie der Dissertation nahm, entsprechend den Überlegungen zum Implementationserfolg von Coburn (2003), die Verbreitung des Jahrgangsübergreifenden Lernens in der Berliner Schulanfangsphase in den Blick. Auf Grundlage von Daten der amtlichen Schulstatistik wurde dabei untersucht, wie JÜL vor der verpflichtenden Einführung,

währenddessen und danach in den Schulen und Klassen umgesetzt wurde. Es zeigte sich, dass in einem Großteil der Schulen JÜL erwartungsgemäß verbreitet war: einige (12,4 %) setzten JÜL bereits vor der Verpflichtung und ein Großteil der Schulen (max. 87 %) organisierten während des Verpflichtungszeitraums in den Schuljahren 2008/09 bis 2012/13 ihren Schulanfang, zumindest teilweise, jahrgangsübergreifend. Seit der Verpflichtungsrücknahme nimmt die Zahl der Schulen mit JÜL stetig ab. Hervorzuheben ist hierbei, dass in jedem Schuljahr mindestens 13 % der Schulen, trotz der gesetzlichen Verankerung von JÜL keine ihrer Klassen jahrgangsübergreifend organisierten. Darüber hinaus wurde aufgezeigt, dass einerseits die meisten Schulen, die JÜL umsetzten, dies auch in allen ihren Klassen der Schulanfangsphase taten. Andererseits gab es zu jedem Zeitpunkt Schulen, die JÜL in nicht allen SAPH-Klassen umsetzten. Das Ziel der Senatsbildungsverwaltung, JÜL in der SAPH vollständig zu implementieren, wurde demnach in keinem Schuljahr in allen Schulen erreicht. Bei dieser rein beschreibenden Erfassung, ohne Berücksichtigung der pädagogischen Umsetzung von JÜL innerhalb der Klassen, ist dies ein bemerkenswerter Befund. Dennoch erscheint es vor dem Hintergrund der großen öffentlichen und medialen Kontroverse um JÜL, wie z. B. der Brandbrief der Schulleitungen zeigt (Vieth-Entus, 2011), fast überraschend, dass immerhin 90 % der Schulen in mindestens einem Schuljahr in mindestens einer Klasse JÜL umgesetzt haben und dass es im Schuljahr 2015/16, also bereits drei Schuljahre nach Aufhebung der Verpflichtung, noch über die Hälfte der Schulen sind, die ihren Schulanfang, zumindest teilweise, jahrgangsübergreifend organisieren. Die Untersuchung der Anzahl der zusammengefassten Jahrgänge innerhalb einer JÜL-Klasse zeigte, dass vor allem seit dem Ende der Verpflichtung die Anzahl der – (reform-) pädagogischen favorisierten – JÜL-3-Klassen gestiegen ist und sich somit vermuten lässt, dass vor allem Schulen mit diesen Klassen JÜL auch nach der Verpflichtung umsetzen. Der Anteil der JÜL-2-Klassen liegt im Schuljahr 2015/16 jedoch immer noch bei 60 %. Weiterhin wurde im Anschluss an die Diffusionstheorie von Rogers (2003) gezeigt, dass sich Berliner Schulen entsprechend ihrer JüL-Umsetzung mittels latenter Profilanalysen in sechs Typen unterscheiden lassen. Die Unterschiede zeigten sich dabei vor allem darin, dass (1) Schulen bereits vor der verpflichtenden Einführung („Pioniere“) bzw. (2) mit der Verpflichtung mit dem Jahrgangsübergreifenden Lernen begonnen haben („Überzeugte“) und auch nach der Aufhebung der Verpflichtung bei JüL geblieben sind. Diese Typen umfassen 50 % der Schulen. (3) Drei weitere Typen bezeichnen Schulen, die JüL zwar in mindestens zwei Schuljahren während der Verpflichtung umgesetzt haben, jedoch spätestens im Schuljahr nach der Aufhebung der Verpflichtung zur jahrgangsbezogenen Organisation zurückgekehrt sind („Pflichtbewusste“, „nichtüberzeugte Skeptiker“, „Nicht-

Überzeugte“; insgesamt 35 %). (4) Fast jede sechste Schule in Berlin hat zu keinem Zeitpunkt ihre gesamte Schulanfangsphase jahrgangsübergreifend organisiert („Verweigerer“).

Um einen Einblick in die Intensität der JÜL-Umsetzung innerhalb einer Schule zu erhalten, wurde im letzten Abschnitt von Studie I geprüft, welcher Anteil der Gesamtjahresunterrichtsstunden in den beiden Hauptfächern Deutsch und Mathematik jahrgangsübergreifend organisiert wurde und ob sich diese Anteile zwischen den beiden Fächern unterscheidet. Die Ergebnisse zeigten, dass die Schulen, die für eine Klasse angeben, dass sie jahrgangsübergreifend organisiert wird, in dieser Klasse den größten Teil der Unterrichtsstunden jahrgangsgemischt unterrichten. Nur eine vergleichsweise geringe Anzahl an Schulen organisiert den Unterricht in den deklarierten JÜL-Klassen in drei von fünf Wochenstunden (oder weniger) jahrgangsgemischt. Fächerunterschiede konnten nicht festgestellt werden. Dies zeigten auch die sehr hohen Korrelationskoeffizienten bei der Gegenüberstellung der beiden Fächer. Die Sorge, dass Schulen JÜL lediglich in Form von Projektwochen und in den Nebenfächern umsetzen, konnte, zumindest für die Schuljahre 2011/12 und 2012/13, nicht bestätigt werden.

6.1.2 Studie II: Unterscheiden sich Schulen entsprechend ihrer Zuordnung zu einem Implementationstypen in ihrer Schul- und Unterrichtsqualität?

Anschließend an die quantitativ-beschreibenden Erkenntnisse aus Studie I, hatte Studie II zum Ziel, die Qualität der JÜL-Implementation in den Blick zu nehmen und somit einen Eindruck der Implementationstiefe (Coburn, 2003) zu erhalten. Dabei sollte vor allem die in der Literatur häufig hervorgehobene Schul- und Unterrichtsqualität (Ditton, 2002), insbesondere im Zusammenhang mit JÜL (Carle & Metzen, 2014), untersucht werden. Mithilfe der Daten der zweiten Runde der Berliner Schulinspektion wurden zunächst die in Studie I identifizierten Typen der JÜL-Implementation bezüglich einzelner Merkmale der Schul- und Unterrichtsqualität gegenübergestellt. Die Ergebnisse der Varianzanalysen zeigen für die „Pioniere“ auf Ebene der Schule bedeutsam bessere Bewertungen der Schulleitungsaktivitäten, der Kommunikation innerhalb des Kollegiums und des Personalmanagements. Auch erhalten die „Pioniere“ eine bessere Einschätzung für die Evaluation schulischer Entwicklungsvorhaben. Jedoch sind in allen genannten Bereichen die Unterschiede nicht sehr groß und ihre praktische Bedeutsamkeit wird nach Cohen (1988) als gering eingeschätzt. In der Beteiligung von Schüler*innen und Eltern sowie in der Fortschreibung des Schulprogramms zeigen sich keine Unterschiede zwischen den einzelnen Implementationstypen.

Für die Unterrichtsebene sind die für die pädagogisch tiefgreifende Umsetzung von JÜL notwendigen Bereiche Innere Differenzierung und Selbstständiges Lernen bei den „Pionieren“ höher bewertet als bei den „Verweigerern“, jedoch sind die Effektgrößen auch hier als gering zu bewerten. Zudem werden diese beiden Merkmale an Berliner Schulen als insgesamt schwach bewertet. In weiteren, zentralen Merkmalen von JÜL (Förderung von Selbstvertrauen und Selbsteinschätzung, pädagogisches Klima im Unterricht, Kooperation des pädagogischen Personals, Kooperatives Lernen, u. a. m.) lassen sich ebenfalls keine systematischen Unterschiede zwischen den Implementationstypen feststellen. Insgesamt muss demnach festgehalten werden, dass selbst die „Pioniere“, die JÜL nach Erkenntnissen aus Studie I erfolgreich in ihren Schulen umsetzen, wesentliche Merkmale des jahrgangsübergreifenden Lernens auf der Ebene des Unterrichts nicht berücksichtigen.

Um die Schulen weiter zu differenzieren und zu prüfen, ob Schulen, die JÜL einerseits langfristig umsetzen und andererseits die pädagogisch favorisierte Mischung dreier Jahrgänge in ihren JÜL-Klassen praktizieren, sich in ihrer Qualität von anderen Schulen unterscheiden, wurden im zweiten Teil von Studie II „Pioniere“ und „Überzeugte“ entsprechend ihrer Zuordnung zu JÜL-2 und JÜL-3 einander gegenübergestellt. Dabei zeigten sich auf Ebene der Schule zugunsten der JÜL-3-Schulen erneut bessere Bewertungen für die Bereiche Kommunikationskultur im Kollegium und Personalmanagement, und auch hier haben die Effekte eine geringe praktische Bedeutsamkeit. Für die Qualität auf der Ebene des Unterrichts wurden für den größten Teil der untersuchten Merkmale (13 von 17) bessere Bewertungen zugunsten der JÜL-3-Schulen festgestellt. Darüber hinaus kann für einige der für JÜL besonders relevanten Merkmale zumindest eine mittlere praktische Bedeutsamkeit der Unterschiede konstatiert werden (Kooperatives Lernen, Systematische Förderung, Selbstständiges Lernen). Bei der inneren Differenzierung unterscheiden sich beide Formen der JÜL-Umsetzung sogar mit einem nahezu großen Effekt zugunsten von JÜL-3. Jedoch gibt es auch für diese Schulen noch reichlich Entwicklungspotenzial in der Umsetzung von Binnendifferenzierung. Die Ergebnisse lassen darauf schließen, dass es einem Teil (14,8 %) der Schulen, nämlich solchen, die JÜL-3 langfristig umsetzen, gelungen ist, wesentliche Merkmale des Jahrgangsübergreifenden Lernens in ihren Unterricht zu implementieren. Weiterhin kann angenommen werden, dass diese Schulen JÜL vergleichsweise tiefgreifend (Coburn, 2003) und pädagogisch-didaktisch angemessen (Carle & Metzen, 2014) umsetzen. Jedoch lässt sich aus diesen Untersuchungen keine Antwort ableiten auf die Frage, ob diese Schulen bereits vor der JÜL-Reform eine hohe Schul- und Unterrichtsqualität zeigten oder ob diese Qualität in Zusammenhang mit der Implementation von JÜL erarbeitet wurde.

Insgesamt kann aus den Ergebnissen der ersten beiden Studien dieser Arbeit geschlossen werden, dass eine Bildungsreform mit dem Ziel einer flächendeckenden und dauerhaften Umsetzung des Jahrgangsübergreifenden Lernens in der Schulanfangsphase nicht als erfolgreich bewertet werden kann, wenn es für lediglich 15 % der Schulen Hinweise darauf gibt, dass sie die Reform entsprechend den bildungspolitischen und pädagogischen Vorgaben umgesetzt haben.

6.1.3 Studie III: Wie wurde die Senkung des Schuleintrittsalters an Berliner Grundschulen umgesetzt?

In Studie III, der ersten von drei Studien zur Senkung des Schuleintrittsalters, wurde zunächst untersucht, inwieweit das Schuleintrittsalter in Berlin tatsächlich gesenkt wurde und ob dies auch in allen Klassen gleichermaßen geschehen ist. Die Ergebnisse bestätigen, dass das mittlere Einschulungsalter, welches vor der Reform sechs Jahre und sieben Monate betrug, mit der Einschulung der Hauptkohorte, also dem Jahrgang, der sowohl nach dem alten als auch nach dem neuen Stichtag eingeschult wurde, gesenkt wurde. Allerdings ist die Senkung um zwei Monate auf sechs Jahre und fünf Monate geringer als die unter Annahme der Gleichverteilung erwarteten drei Monate. Zurückgeführt werden kann dieser Umstand auf die Tatsache, dass im Schuljahr vor der Reform (2004/05) bis zu 6 % der Eltern ihr Kind von der Einschulung zurückgestellt haben (AfS, 2016). In der Hauptkohorte bestand für Eltern offiziell keine Möglichkeit, ihr Kind zurückzustellen, sodass der Anteil der vergleichsweise älteren Schüler*innen höher ist als in den Jahren zuvor. Hinzukommt, dass im selben Schuljahr 13 % der Eltern ihr Kind vorzeitig eingeschult haben (ebd.). Im Schuljahr 2005/06, dem Jahr der Reform, sank dieser Anteil auf 2,1 %, was ebenfalls dazu beiträgt, dass das mittlere Einschulungsalter im direkten Vergleich der beiden Kohorten nicht wie gewünscht gesenkt wurde. Vier Jahre nach der Einschulungsreform betrug das mittlere Einschulungsalter dann sechs Jahre und zwei Monate, was dem bildungspolitischen Ziel von sechs Jahren und einem Monat sehr nahekommt. Bei der Prüfung der Verteilung des mittleren Einschulungsalters in den einzelnen Klassen der Schulanfangsphase konnten mithilfe von Intraklassenkorrelationen keine systematischen Unterschiede festgestellt werden, sodass angenommen werden kann, dass die Reform in allen Klassen der Berliner Schulanfangsphase umgesetzt wurde und sich Schüler*innen eines bestimmten Alters (z.B. ältere Schüler*innen) nicht in einzelnen Klassen konzentrieren.

Insgesamt liefern die Ergebnisse, vor allem aus Vergleichskohorte II, Hinweise darauf, dass das mittlere Einschulungsalter, zumindest solange der Stichtag nicht erneut verändert wurde, gesenkt werden konnte. Da Eltern jedoch in den vergangenen Jahren mehr und mehr die Möglichkeit erhielten, ihr Kind zurückzustellen, und davon auch wieder verstärkt Gebrauch machen (Autorengruppe Bildungsberichterstattung, 2016), kann davon ausgegangen werden, dass aktuelle mittlere Einschulungsalter über sechs Jahren und zwei Monaten liegt.

6.1.4 Studie IV: Wie wirkt sich die Senkung des Schuleintrittsalters auf die Schüler*innenleistung aus?

Nach der Prüfung der Umsetzung der Einschulungsreform widmete sich Studie IV der Untersuchung der Auswirkungen auf die schulischen Leistungen und Bildungswege. Hierbei wurden die einzelnen Schüler*innen entsprechend ihrem Alter einer Altersgruppe zugeordnet (z. B. in der Hauptkohorte: jung, mittel, alt). Anschließend wurden die Leistungen der Altersgruppen innerhalb der drei Einschulungskohorten und über die Kohorten hinweg vergleichend betrachtet (relative Alterseffekte). Die Ergebnisse zeigten, dass sich die Verteilungen der Altersgruppen innerhalb der Kohorten (über die Jahrgangsstufen) und über die Kohorten hinweg in einem nicht relevanten Ausmaß veränderten. So nahm die Gruppe der vergleichsweise jüngeren Schüler*innen in der Hauptkohorte über die Zeit zwar ab, nach Cohen (1988) jedoch nicht in einem praktisch bedeutsamen Ausmaß, sodass keine systematischen Auswirkungen zu erwarten sind. Bezogen auf die Verteilung auf die weiterführenden Schulen (in Berlin Gymnasium und Integrierte Sekundarschule) konnten für die Hauptkohorte keine Unterschiede festgestellt werden. Sowohl jüngere als auch mittlere und ältere Schüler*innen finden sich in beiden Schularten zu etwa gleichen Anteilen. Die altersgruppenspezifischen Leistungen (relative Alterseffekte) zeichnen ein ähnliches Bild: ist zu Beginn der Schullaufbahn (Jahrgangsstufe 2) noch eine (eher geringe) leistungsbezogene Benachteiligung der jüngeren Schüler*innen in allen Schüler*innengruppen (gesamt, weiblich, männlich, deutscher bzw. nichtdeutscher Herkunftssprache) der Hauptkohorte für die Lernbereiche Deutsch / Lesen und Mathematik festzustellen, so ist diese Benachteiligung in den meisten Gruppen für beide Lernbereiche bereits am Ende der dritten Jahrgangsstufe ausgeglichen, was auch die Ergebnisse der Vergleichskohorte II bestätigen und mit den Erkenntnissen von Gold und Kolleg*innen (2012) und Crone und Whitehurst (1999) einhergeht. In der achten Jahrgangsstufe lassen sich sowohl für die Hauptkohorte als auch für die Vergleichskohorte I keine Leistungsunterschiede zwischen den drei Altersgruppen mehr feststellen.

6.1.5 Studie V: Relative Age Effects in Mathematics and Reading: Investigating the Generalizability across Students, Time, and Classes

Studie V stellte im Anschluss an Studie IV relative Alterseffekte sowie deren Generalisierbarkeit über verschiedene Schüler*innengruppen, Kohorten und Klassen in den Mittelpunkt. Die Ergebnisse bestätigten zunächst die Erkenntnisse der Mehrzahl bisheriger Studien sowie der Studie IV dieser Arbeit: zu Beginn der Schullaufbahn (hier: Jahrgangsstufe 2) lassen sich altersgruppenbezogene Leistungsunterschiede zuungunsten der vergleichsweise jüngeren Schüler*innen feststellen. Weiterhin wurden mit Blick auf die praktische Bedeutsamkeit die Richtlinien von Hill und Kolleg*innen (2008) herangezogen. Dabei zeigt sich, dass die Älteren im Vergleich zu den Jüngeren im Mittel einen Leistungsvorsprung von zehn Wochen eines Schuljahres haben, was als kleiner Effekt eingestuft wird. Im weiteren Verlauf der Studie (Jahrgangsstufe 8) zeigt sich auch hier die Leistungsangleichung der Altersgruppen. Darüber hinaus konnte in den Analysen gezeigt werden, dass sich dieses Befundmuster über weite Teile verallgemeinern lässt. Konkret bedeutet dies, dass die Ergebnisse aller Schüler*innen in der Alterstichprobe der Hauptkohorte (a) auf Schüler*innen mit deutscher und nichtdeutscher Herkunftssprache, (b) auf Schüler*innen, die unter unterschiedlichen Stichtagsregelungen und daraus resultierenden Unterschieden im Durchschnittsalter und der Altersverteilung eingeschult wurden, sowie (c) über Schulklassen hinweg übertragen werden können.

Insgesamt zeigen die Ergebnisse der Studie V, mit welcher Leistungsheterogenität Lehrkräfte vor allem in der Grundschule umzugehen haben, und heben die Notwendigkeit einer auf individuelle Förderung ausgelegten Lernumgebung hervor. Weiterhin bestätigen die Befunde, vor allem aus den achten Jahrgangsstufen, das homogenisierende Potenzial der Schule (z. B. Downey, Von Hippel & Broh, 2004; Kerstan & Spiewak, 2008), das sich in der Angleichung der Leistungen zwischen den Altersgruppen zeigt.

Im Hinblick auf den Implementationserfolg zeigen die Ergebnisse der Studien zur Senkung des Schuleintrittsalters, dass es den Schulen in weiten Teilen gelungen ist, die Einschulungsreform, einerseits flächendeckend umzusetzen und andererseits, zumindest bezogen auf die Verteilungen der Schüler*innen und ihre Leistungen, erfolgreich zu bewältigen und die vergleichsweise jüngeren Kinder größtenteils zu integrieren. Eine systematische und substanzielle Benachteiligung der Jüngsten, wie sie vor allem in der Öffentlichkeit häufig befürchtet wurde (u. a. Hecht, 2010; Vieth-Entus, 2013a), kann, zumindest leistungsbezogen,

über die Zeit nicht festgestellt werden. Zwar sind in zweiten Jahrgangsstufe Leistungsbenachteiligungen zuungunsten der Jüngeren auszumachen, mit Blick auf die Größe der Effekte, sind diese jedoch eher als klein einzuschätzen.

6.2 Diskussion vor dem Hintergrund implementationstheoretischer Erkenntnisse

Im Hinblick auf die in Kapitel 1 dargestellten vier Phasen des Implementationsprozesses – Initiation, Umsetzung, Evaluation, Institutionalisierung (Tabelle 1, S. 26) – wurden in den Untersuchungen im Rahmen dieser Arbeit zunächst die Phasen der Initiation und der Umsetzung der Reformen JÜL und Einschulung in Berlin dargestellt (Kapitel 2). Anschließend bildete die Evaluation in den Studien zum Implementationserfolg den eigentlichen Kern der Dissertation. Da JÜL als Standardpraktik in der SAPH bereits zurückgenommen wurde und ab dem Schuljahr 2017/18 der Stichtag zur Einschulung auf den 30. September verschoben und somit das Schuleintrittsalter um drei Monate angehoben wird, kann von einer Institutionalisierung der beiden Reformen, zumindest auf der Ebene des Systems, nicht ausgegangen werden. Bezogen auf einzelne Schulen ist eine Verankerung von JÜL jedoch möglich und nach den Erkenntnissen dieser Arbeit auch in einigen Schulen Berlins erfolgt. Weiterhin werden in der deutschen bildungswissenschaftlichen Implementationsliteratur wesentliche Einflussfaktoren für den Implementationsprozess und -erfolg auf vier Ebenen zusammengefasst (Tabelle 2, S. 31): die Ebenen der Lehrkräfte, der Innovation, der Einzelschule sowie des Schulsystems. In den Untersuchungen stand dabei die Ebene des Schulsystems im Vordergrund, sowohl in der Beschreibung der Umsetzung und Initiierung der Reformen JÜL und Früheinschulung als auch in ihrer Bewertung. Der Einbezug der Ebene der Innovation erfolgte in der Beschreibung der Berliner Situation. Lehrkräfte und Einzelschule konnten für Analyse des Implementationserfolgs nicht systematisch einbezogen werden.

6.2.1 Studien zur Einführung des Jahrgangsübergreifenden Lernens

Im Hinblick auf die Erfassung des Implementationserfolgs wurde dieser für die Untersuchungen zum Jahrgangsübergreifenden Lernen entlang der (1) vier Erfolgsindikatoren nach Coburn und (2) der Innovationsbereitschaft nach Rogers (2003) auf der Basis schulstatistischer Daten sowie einer repräsentativen Erweiterungsstudie operationalisiert. Im Hinblick auf die Verbreitung ist JÜL zum Schuljahr 2015/16 in der Hälfte der Schulen mit Schulanfangsphase erfolgreich implementiert. Die von Rogers postulierte idealtypische Beschreibung lässt sich in den identifizierten Implementationstypen wiederfinden. Abbildung

27 stellt die Anteile der einzelnen Gruppen gegenüber. Im JÜL-Modell bilden Innovative und Vorreiter*innen eine gemeinsame Gruppe („Pioniere“). Diese sind mit einem Anteil von 30 % deutlich stärker vertreten als in dem Modell von Rogers (2,5 % + 13,5 %). Der höhere Anteil kann damit begründet werden, dass im Land Berlin im Rahmen von Modellversuchen („JÜL“; „Schuleingangsphase“) Schulen die Jahrgangsmischung bereits erprobten. Die Gruppe der „Überzeugten“ (frühes Hauptfeld bei Rogers) hat einen etwas geringeren Anteil als idealtypisch angenommen. Insgesamt beschreiben die Gruppen der Innovativen, Vorreiter*innen und das frühe Hauptfeld bzw. die „Pioniere“ und „Überzeugten“ in beiden Modellen beinahe exakt die Hälfte der Schulen, sodass lediglich eine Verschiebung der Kategorien festgestellt wird, die grundlegende Beschreibung von Rogers für die JÜL-Implementation in Berlin jedoch als zutreffend bezeichnet werden kann. Auch die zweite Hälfte der Verteilung von Rogers spiegelt sich in der JÜL-Implementation wieder. So können die Gruppen „Pflichtbewusste“, „nichtüberzeugte Skeptiker“ und „Nicht-Überzeugte“ als spätes Hauptfeld zusammengefasst werden. Sie machen etwa ein Drittel der Schulen aus. Die Gruppe der Nachzügler*innen, also die Traditionalist*innen, entspricht in Berlin den „Verweigerern“, die JÜL zwar kaum oder gar nicht umgesetzt haben, bei Beibehaltung der Verpflichtung möglicherweise jedoch auch noch überzeugt hätten werden können.

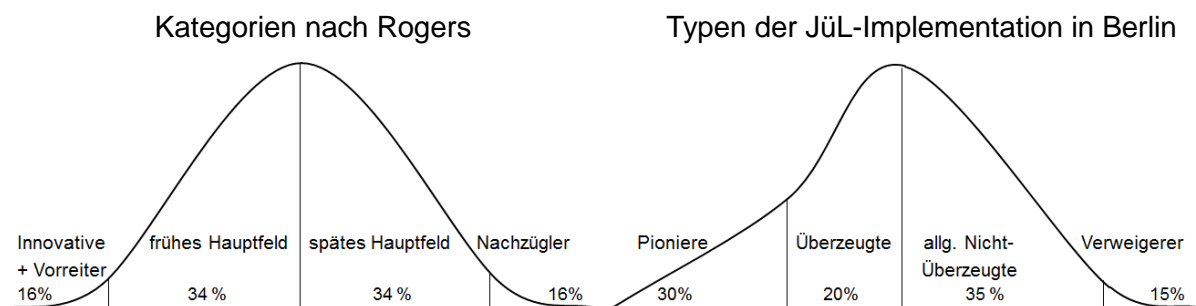


Abbildung 27: Gegenüberstellung der Anwendungskategorien mit den JÜL- Implementationstypen

In der Zusammenfassung der Erkenntnisse aus beiden Studien zu JÜL lässt sich der Implementationserfolg von JÜL in Berlin entlang der vier Erfolgsindikatoren nach Coburn (2003) wie folgt beschreiben: (1) Bezogen auf die Nachhaltigkeit der Reform wurde festgestellt, dass trotz der Verpflichtungsrücknahme etwa die Hälfte der Schulen in Berlin ihre Schulanfangsphase zumindest teilweise jahrgangsbezogen organisiert. Auch wenn dies von dem ursprünglichen bildungspolitischen Ziel einer flächendeckenden, dauerhaften Umsetzung abweicht, kann es doch als zumindest oberflächlicher Erfolg gewertet werden. (2) Mit Blick auf die Reformübernahme kann auf Grundlage der Studienergebnisse angenommen werden, dass Schulen der Gruppen „Pioniere“ und „Überzeugte“ JÜL weitestgehend übernommen

haben und sich mit der Reform, also mit dem Jahrgangsübergreifenden Lernen als pädagogischem Konzept, identifizieren. Inwieweit sich innerhalb einer Schule alle beteiligten Akteur*innen mit der Reform identifizieren, konnte dabei nicht untersucht werden. (3) Die Tiefe der Implementation betreffend, liefern die Untersuchungen in Studie II, vor allem bezogen auf die tiefgreifende und konsequente Veränderung des Unterrichtsgeschehens, Hinweise, dass dies lediglich den Schulen gelungen ist, die JÜL nachhaltig und im Zusammenhang mit relativ hoher Schul- und Unterrichtsqualität umzusetzen. (4) Die Verbreitung der Reform betreffend, gibt es zur Umsetzung innerhalb der Schulen (Anzahl der Klassen) und innerhalb der Klassen (Anzahl der Unterrichtsstunden) quantitative Hinweise, dass es gelungen ist, JÜL an etwa der Hälfte der Schulen zu implementieren.

6.2.2 Studien zur Senkung des Schuleintrittsalters

Auch für die Studien zur Senkung des Schuleintrittsalters wurden die vier Indikatoren von Coburn (2003) herangezogen: (1) Ziel der Reform war es, das mittlere Einschulungsalter der Schüler*innen systematisch zu verringern und somit, langfristig betrachtet, in Kombination mit anderen Reformen wie der Verkürzung des Abiturs (SenBJW, 2013a), das Berufseintrittsalter Berliner Absolvent*innen ebenfalls abzusenken. Die Studien III bis V bestätigen, dass dies zum Teil gelungen ist: das Einschulungsalter war fünf Jahre nach Stichtagsverschiebung (Vergleichskohorte II) um fünf Monate auf sechs Jahre und zwei Monate gesenkt. Auch wurde die Reform in allen Schulen und Klassen, also in der Breite, umgesetzt. (2) Die Betrachtung der altersgruppenspezifischen Leistungen liefert Hinweise, dass die Reform auch tiefgreifend umgesetzt wurde, indem die vergleichsweise Jüngsten bezogen auf ihrer Leistung größtenteils integriert werden konnten. Ob die Reform auch Einfluss auf die konkrete Lehr-Lern-Situation im Unterricht hatte, konnte im Rahmen dieser Studie nicht untersucht werden. (3) Im Hinblick auf die Nachhaltigkeit der Reform ist einerseits festzuhalten, dass die erneute Stichtagsverschiebung zum Schuljahr 2017/18 keine dauerhafte Umsetzung des Vorhabens im ursprünglichen Sinne, mit dem 31. Dezember als Stichtag, ermöglicht. Andererseits liefert die erneut steigende Zahl der Rückstellungen (Studie III) Belege dafür, dass (a) einige schulpflichtige Kinder von ihren Eltern, oder auch ihren Erzieher*innen, als nicht schulfähig eingeschätzt werden oder aber (b) Eltern das bildungspolitische Ziel eines niedrigeren Berufseintrittsalters für ihr Kind nicht verfolgen. (4) Über die Identifikation der Schulen bzw. ihrer Akteur*innen mit der Reform liefern die Untersuchungen keine Hinweise. Neben den Erfolgsindikatoren nach Coburn (2003) wurden zur Bewertung des Implementationserfolgs der Einschulungsreform die Leistungen der Schüler*innen herangezogen.

Insgesamt erwiesen sich die für alle hier durchgeführten Studien zugrunde gelegten theoretischen Modelle von Coburn (2003) und Rogers (2003) als geeignet, den Implementationserfolg zu beschreiben. Besonders die multidimensionale Betrachtung, wie sie die Indikatoren von Coburn anregen, erscheint hier sinnvoll und notwendig, wie auch die Ergebnisse der fünf Studien zeigen. Die Anwendungskategorien, die Rogers zur Beschreibung der Innovationsbereitschaft von Individuen definiert, lassen sich im Rahmen dieser Arbeit auch für die Kategorisierung von Schulen verwenden und bilden eine geeignete Grundlage für die weitere Charakterisierung der Schulen. Insbesondere für die Untersuchung der qualitativen Erfolgsindikatoren nach Coburn (Identifikation, Nachhaltigkeit, Tiefe) erwiesen sich die Implementationstypen als gewinnbringend und dimensionsreduzierend. Mit Blick auf Reformprozesse im Bildungssystem erinnern die Typen an die Trägheit des Bildungssystems (Rogers & Shoemaker, 1971) und die damit einhergehenden Notwendigkeit zur Berücksichtigung der Größe Zeit in Reformprozessen. Eine Schwäche, insbesondere der Roger'schen Anwendungskategorien, stellt aber auch Trägheit des Modells selbst dar. So erschwert die Abbildung der Innovationsbereitschaft auf der Zeitachse die Berücksichtigung von Adaption- und Veränderungsprozessen innerhalb der Reform. Weiterhin werden mit der alleinigen Anwendung dieser Modelle die Phasen der Initiation und Umsetzung von Reformen nicht berücksichtigt, sodass für die Einschätzung des Erfolgs einer Reform der Einbezug der Implementationsphasen und Einflussfaktoren, wie sie vor allem von Goldenbaum (2012) umfassen und übersichtlich zusammengetragen wurden, notwendig erscheint. Abschließend werden in den für diese Arbeit herangezogenen implementationstheoretischen Überlegungen sogenannte Nebenfolgen oder nicht intendierte Effekte (Bellmann, Schweizer & Thiel, 2016) von Reformimplementationen außer Acht gelassen. Doch gerade vor dem Hintergrund der hier festgestellten, oftmals nicht tiefgreifenden Umsetzung von JÜL stellt sich die Frage, welche Folgen sich aus einer eher oberflächlichen Umsetzung für die Lehr-Lern-Prozesse und somit auch für die Schüler*innen ergeben. Und auch im Zusammenhang mit der Einschulungsreform sind Nebenfolgen, die sich in Verhaltensweisen der Eltern oder Lehrkräfte oder im Umgang der Schüler*innen untereinander zeigen können, denkbar.

6.3 Limitationen der Arbeit und Ausblick

Trotz des umfangreichen Erkenntnisgewinns zum Implementationserfolg von JÜL und Früheinschulung in Berlin sind folgende Limitationen der Arbeit festzuhalten: (1) eine der wesentlichen Beschränkungen für die Studien stellt die Tatsache dar, dass die Bewertung des

Implementationserfolgs ex post erfolgte. Dies ist besonders für die Beschreibung der konkreten Prozesse in den Schulen, bei Fortbildungen und der Schulaufsicht eine Einschränkung. So konnten diese Prozesse mithilfe von Expert*innengesprächen sowie umfangreicher Recherche und Lektüre zwar zum Teil, aber eben nicht voll umfassend rekonstruiert werden. (2) Auf Grundlage der verwendeten Daten konnten die Indikatoren zur Prüfung des Implementationserfolgs nach Coburn (2003) größtenteils untersucht werden. Es lagen jedoch z. B. keine Informationen zu den Einstellungen der beteiligten Akteur*innen (Schüler*innen, Lehrkräfte, Schulleitung) vor, was eine Prüfung der Implementationstiefe (*depth*) im Rahmen der JÜL-Studien auf die Untersuchung der Qualität der Lehr-Lern-Prozesse beschränkte und es nicht zuließ, z. B. die Identifikation mit der Reform (*shift in reform ownership*) zu untersuchen. Die Studien zum Schuleintrittsalter ließen auf Grundlage der Daten keine Rückschlüsse auf das Unterrichtsgeschehen zu. Hinweise zur Implementationstiefe konnten einzig auf der Basis der Leistungsdaten untersucht werden. (3) Bezogen auf Studie II, folgt aus dem Prozedere der Schulinspektion, dass unter den untersuchten Schulen auch solche sind, die bereits vor einigen Jahren inspiziert wurden. Der Annahme von Borman und Kolleg*innen (2003) folgend, wonach sich Effekte von Implementationen auf Schüler*innenleistung erst nach fünf Jahren zeigen, liegt die Vermutung nahe, dass es in der Stichprobe Schulen gibt, deren JÜL-Umsetzung sich noch nicht bedeutsam auf die Schul- und Unterrichtsqualität auswirken konnte. (4) Weiterhin erfasst die Schulinspektion die Qualität aller Jahrgangsstufen einer Schule, sodass vor allem mit Blick auf die große Varianz der Qualität innerhalb von Schulen (z. B. Wurster & Gärtner, 2013) anzunehmen ist, dass Befunde aus der Schulanfangsphase mit denen der übrigen Jahrgangsstufen vermischt werden. Folglich kann nicht ausgeschlossen werden, dass eine konkrete Betrachtung insbesondere der Unterrichtsqualität in der Schulanfangsphase zu veränderten Erkenntnissen bezüglich des Implementationserfolgs von JÜL führen könnte. (5) Die Datenbasis der Studien zum Schuleintrittsalter ließ keine Betrachtung individueller Bildungsverläufe z. B. von relativ jüngeren Schüler*innen zu. Die Studienergebnisse liefern somit keine Hinweise für individuelle Einschulungsentscheidungen. Weiterhin kann trotz der repräsentativen Datengrundlage aus Auftreten von Selektionseffekten nicht kontrolliert und folglich nicht gänzlich ausgeschlossen werden.

Trotz der Limitationen ist es mit dieser Arbeit gelungen, wichtige Aspekte der Implementation von JÜL und Früheinschulung zu untersuchen. Besonders die Anwendung der Modelle von Coburn und Rogers ermöglichte eine multidimensionale Betrachtung des Implementationserfolgs dieser beiden Reformen und liefert wichtige Hinweise für künftige Reformprozesse in Berlin, aber auch darüber hinaus. Zudem können die Ergebnisse der Arbeit

zur Versachlichung der Diskussion um JÜL und Früheinschulung beitragen, da u. a. der Einfluss des Alters auch für Berliner Schüler*innen relativiert wurde. Ferner gilt: Auch, wenn die Datengrundlage nur wenige Hinweise auf die Lehr-Lern-Prozesse liefert, so kann sie doch auf einer beschreibenden Ebene die Implementation der beiden Reformen umfassend und repräsentativ darlegen. Abschließend können aus den Erkenntnissen der Arbeit bildungspolitische und forschungspraktische Implikationen abgeleitet werden, die die folgenden zwei Abschnitte beschreiben.

6.4 Implikationen für die Bildungspolitik

Eine wesentliche bildungspolitische Implikation ist die bereits vielerorts geäußerte Forderung nach wissenschaftlicher Begleitung und formativer Evaluation großer Reformvorhaben (Abgeordnetenhaus Berlin, 2005a, Abgeordnetenhaus Berlin, 2007c), wie es die Neuerungen der Schulanfangsphase darstellen. Auch wenn Kenntnisse aus Modellvorhaben Eingang in die Transferüberlegungen von SAPH und JÜL gefunden haben (Wilde, 2006), so benötigen Schulen doch umfangreiche personelle und materielle Begleitung bei der Umsetzung solch tiefgreifender Veränderungen. Gerade die Erkenntnisse aus der umfassenden begleitenden Evaluation der FLEX haben gezeigt, vor welchen Herausforderungen Brandenburger Schulen bei der Einführung einer flexiblen, integrativen Schulanfangsphase stehen (Liebers et al., 2008). Aus den Ergebnissen, insbesondere der JÜL-Studien, kann auf die Notwendigkeit umfassender Unterstützungssysteme auf allen Ebenen (Innovation, Lehrkräfte, Einzelschule, Schulsystem) geschlossen werden, da ein Mangel an selbigen eine tiefgreifende Implementation von Reformen erschwert. Das Beispiel der Implementation von JÜL verdeutlicht zudem die enge Verzahnung der vier Ebenen: so trug die Verschiebung der JÜL-Verpflichtung um ein Schuljahr dazu bei, dass die bereits fortgebildeten Lehrkräfte größtenteils in Jahrgangsklassen zum Einsatz kamen und bei der anschließenden Einführung von JÜL bedeutend weniger fortgebildete Lehrkräfte zur Verfügung standen (Wilde, 2014), was, so kann angenommen werden, eine professionelle Umsetzung, aber auch die Akzeptanz der Reform merklich einschränkte. Gerade auch die Einführung von JÜL zeigt beispielhaft die Bedeutung des Wissens um die Inhalte und Eigenschaften und in diesem Fall auch die pädagogischen Erkenntnisse und insbesondere auch intradisziplinären Kontroversen im Zusammenhang mit einem Reformvorhaben. Ebenso, was bildungspolitisch eine besondere Herausforderung darstellen dürfte, muss der Faktor Zeit verstärkte Berücksichtigung erfahren: jede Reform benötigt, entsprechend ihrem Umfang, eine vorangestellte Phase der Planung und Vorbereitung.

Diese Phase umfasst die Erarbeitung von Zeit- und Maßnahmenplänen, Entwicklung von Konzepten zur Sicherung der Qualifikation der beteiligten Akteur*innen (z. B. im Rahmen von Fortbildungen) und die Durchführung entsprechender Qualifikationsmaßnahmen. Aber auch der Austausch mit den Zielgruppen der Reform, das Werben um Akzeptanz und die Überzeugung von der Notwendigkeit einer Reform sind wesentliche Aktivitäten in der Vorbereitung von Reformvorhaben. Zudem verdeutlichen die Ergebnisse dieser Arbeit auch die Relevanz des Faktors Zeit für die eigentliche Umsetzung einer Reform. So erscheint es wenig sinnvoll, einen Reformprozess zu initiieren und bereits nach kurzer Zeit die notwendigen personellen, materiellen und weiteren Unterstützungssysteme einzuschränken oder gar die Reform selbst infrage zu stellen.

Bezüglich des „idealen“ Einschulungsalters bestätigen die Befunde der Studien, dass es zwar altersspezifische Unterschiede zu Beginn der Schullaufbahn gibt, diese aber als nicht sehr groß einzuschätzen sind und auf Grundlage des Alters nur maximal 1,1 % der Leistungsvarianz erklärt werden können. Leistungsunterschiede zwischen Schüler*innen werden demnach in viel größerem Ausmaß durch andere Merkmale (z. B. soziale Herkunft: OECD, 2010b) erklärt. Für bildungspolitische Entscheidungen und vor allem ihre Kommunikation können solche Befunde (noch) stärker in den Vordergrund gerückt werden und zu einer Versachlichung der Diskussion beitragen. An dieser Stelle erscheint eine starke Vernetzung von Bildungspolitik, Schulaufsicht und mindestens einzelnen Ansprechpersonen in den Schulen von Vorteil. Im Hinblick auf die für jegliche Reform notwendige Innovationsbereitschaft und Akzeptanz vonseiten der beteiligten Akteur*innen (Böse, 2015; Pant et al., 2008b), sollte geprüft werden, inwieweit Lehrkräfte bereits in ihrer Ausbildung und darüber hinaus auch berufsbegleitend sensibilisiert und darauf vorbereitet werden, als *change agents* (Fullan, 1993; Jäger, 2004), also „als eine vermittelnde Instanz zwischen Entwicklungs- und Adressatenkontext“ zur stetigen Entwicklung des Bildungswesens beizutragen.

6.5 Implikationen für die Forschung

Für die Einschätzung der Wirkung von JÜL werden meist Effekte auf die Leistung oder Kompetenzstände sowie die sozio-emotionale Entwicklung und Motivation von Schüler*innen untersucht. Die Auswirkungen des Jahrgangsübergreifenden Lernens auf die Leistungsentwicklung von Schüler*innen wurde in der Meta-Analyse von Hattie (2009) untersucht. Basierend auf den Meta-Analysen von Gutiérrez und Slavin (1992) und Veenman

(1995) wurden hier die Ergebnisse von über 94 Einzelstudien zusammengefasst und eine gemittelte standardisierte Cohen'sche Effektgröße d bei nahezu null ($d = ,04$) festgestellt. Das bedeutet, JÜL fördert die Leistungen von Schüler*innen nicht mehr und nicht weniger als der Besuch des Unterrichts in „traditionellen“ jahrgangshomogenen Klassen. Zu diesem Ergebnis kommen ebenfalls Kuhl und Kolleg*innen bei ihrer Untersuchung der Effekte von JÜL auf die Kompetenzstände von Viertklässlern in ganz Deutschland mithilfe von Daten aus dem IQB-Ländervergleich (Kuhl, Felbrich, Richter, Stanat & Pant, 2013). Auch für Berlin wurde in einer eigenen Untersuchung, ebenfalls mit Daten aus dem IQB-Ländervergleich, dieser Nullbefund bestätigt (Thoren, Bröder, Gelke & Brunner, 2016). Der aktuelle Forschungsstand zur Wirkung von JÜL auf die Lernmotivation ist gemischt. Die Meta-Analyse von Veenman (1995) findet auch hier Effekte um null. Eine eigene Untersuchung der Wirkung auf Lernmotivation und affektives Erleben mit Daten aus dem IQB-Ländervergleich bestätigt diesen Befund (Thoren, Bröder et al., 2016). Jedoch legen einige Studien nahe, dass JÜL kleine positive Effekte auf die sozio-emotionale Entwicklung (z. B. Roßbach, 2010), das psychosoziale Wohlbefinden (Hattie, 2009) sowie die Lernfreude und die Anstrengungsbereitschaft (Kuhl et al., 2013) haben kann.

Eine Limitation vieler dieser Studien ist häufig die unterschiedliche Definition von JÜL und die damit einhergehende Erfassung. So lässt beispielsweise ein Selbstbericht der Schule, ob und in welchen Klassen jahrgangsübergreifend unterrichtet wird, noch keinen Rückschluss darüber zu, ob in diesen Klassen wichtige Bestandteile von JÜL (z. B. Kooperatives und selbstständiges Lernen, Carle & Metzen, 2014) umgesetzt werden, also das Wesen von JÜL (Hanke, 2008; Wendt, 2007) erfasst wurde, oder ob sich zwei oder mehr Jahrgänge in einem Raum befinden, aber dennoch nicht gemeinsam unterrichtet werden. Die Ergebnisse der vorliegenden Studien bestätigen diese Kritik, indem sie aufzeigen, dass JÜL in Berliner Grundschulen einerseits verbreitet ist und im quantitativen Sinne zum Teil nachhaltig implementiert wurde, andererseits eine tiefgreifende Implementation für einen Großteil der Schulen jedoch nicht angenommen werden kann. Künftige Studien zur Umsetzung und Wirkung von JÜL sollten demnach stets die Qualität der Umsetzung berücksichtigen. Die Daten der Schulinspektion liefern dafür eine nicht perfekte (siehe Abschnitt 6.2), aber aus Sicht der Forschung doch ökonomisch günstige Datengrundlage, die zumindest einen Einblick in die Qualität von Schule und Unterricht ermöglicht. Für umfassende Einblicke in den Implementationsprozess einerseits, aber auch, um mögliche Veränderungen in Lehr-Lern-Prozessen durch Reformen identifizieren zu können, können Daten aus Unterrichtsbeobachtungen, Videografiestudien oder Expert*innengesprächen gewinnbringend sein. Weiterhin können Verfahren, wie die der sozialen Netzwerkanalyse (Kolleck, 2013, 2014; Zander, Kolleck & Hannover, 2014),

Aufschluss darüber geben, wie eine Verbreitung von Innovationen innerhalb der Einzelschule gelingen kann und dazu beitragen die Relevanz sozialer Prozesse und Kommunikationswege in Reformprozessen hervorzuheben.

Mit dem Ziel der Konkretisierung von Gelingensbedingungen für JÜL in Berlin kann der Blick auf die Schulen, die JÜL mit der Reform eingeführt haben und nach Aufhebung der Verpflichtung auch dabei geblieben sind (hier die „Überzeugten“), aufschlussreich sein. Mithilfe von Einzelfallstudien (Morgan & Morgan, 2009) könnten so Faktoren identifiziert werden, die die Schulen zu einer Weiterführung von JÜL bewogen bzw. ihnen die Weiterführung ermöglicht haben. Darüber hinaus wäre auch ein vertiefender Blick auf die „Verweigerer“ interessant, um zu erfahren, welche Gegebenheiten sie an einer Umsetzung vom JÜL hinderten.

7. Literaturverzeichnis

- Abgeordnetenhaus Berlin. (Mai 2000). *Kleine Anfrage: Auswirkungen der Individualisierung der Schulanfangsphase'in der Grundschulreform 2000 auf die Angebote der Tageseinrichtungen für Kinder*. Kleine Anfrage 14/582. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/14/KlAnfr/k140582.pdf>
- Abgeordnetenhaus Berlin. (Februar 2003). *Kleine Anfrage: Schulversuche*. Drucksache 15/10367. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/15/KlAnfr/ka15-10367.pdf>
- Abgeordnetenhaus Berlin. (Februar 2004). *Mitteilung: Konzeption für eine flexible Schulanfangsphase*. Drucksache 15/2504. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/15/DruckSachen/d15-2504.pdf>
- Abgeordnetenhaus Berlin. (Oktober 2005a). *Antrag: Was bringt die Reform wirklich? – Wissenschaftliche Begleitung und Evaluation der flexiblen Schulanfangsphase dringend erforderlich!* Drucksache 15/4322. Verfügbar unter www.parlament-berlin.de/ados/IIIPlen/vorgang/d15-4322.pdf
- Abgeordnetenhaus Berlin. (November 2005b). *Antrag: Schule mit Zukunft II – Entscheidungen in der Schule treffen, nicht in der Amtsstube!* Drucksache 15/4402. Verfügbar unter <http://www.parlament-berlin.de/ados/IIIPlen/vorgang/d15-4402.pdf>
- Abgeordnetenhaus Berlin. (Januar 2006a). *Kleine Anfrage: Grundschulreformen – Anspruch und Wirklichkeit I: Für bessere individuelle Förderung aller Kinder notwendige Entwicklung einer neuen Lern- und Schulkultur*. Drucksache 15/13109. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/15/KlAnfr/ka15-13109.pdf>
- Abgeordnetenhaus Berlin. (Januar 2006b). *Kleine Anfrage: Grundschulreformen – Anspruch und Wirklichkeit III: Für die bessere individuelle Förderung aller Kinder notwendige räumliche Ausstattung*. Drucksache 15/13111. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/15/KlAnfr/ka15-13111.pdf>
- Abgeordnetenhaus Berlin. (November 2006c). *Antrag: Gesetz zur Änderung des Schulgesetzes für das Land Berlin (Schulgesetz – SchulG) Flexible Schulanfangsphase gründlich vorbereiten*. Drucksache 16/0064. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/16/DruckSachen/d16-0064.pdf>
- Abgeordnetenhaus Berlin. (Dezember 2006d). *Antrag: Schulanmeldung und Sprachstandsfeststellung vorziehen – Früheinschulung ermöglichen*. Drucksache 16/0058 Neu. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/16/DruckSachen/d16-0058.pdf>
- Abgeordnetenhaus Berlin. (Januar 2007a). *Kleine Anfrage: Ende der Schulanfangsphase*. Drucksache 16/10259. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/16/KlAnfr/ka16-10259.pdf>

- Abgeordnetenhaus Berlin. (Februar 2007b). *Antrag: Wiedereinführung der Schulreifeuntersuchung für Schulanfänger*. Drucksache 16/0239. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/16/DruckSachen/d16-0239.pdf>
- Abgeordnetenhaus Berlin. (Februar 2007c). *Antrag: Wissenschaftliche Begleitung der flexiblen Schulanfangsphase unverzüglich einführen*. Drucksache 16/0238. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/16/DruckSachen/d16-0238.pdf>
- Abgeordnetenhaus Berlin. (Februar 2007d). *Kleine Anfrage: Sind die Berliner Schulen nicht ausreichend auf die Flexible Schulanfangsphase und die Einführung der Jahrgangsmischung vorbereitet?* Drucksache 16/10347. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/16/KlAnfr/ka16-10347.pdf>
- Abgeordnetenhaus Berlin. (März 2007e). *Antrag: Bessere Bildung: Individuelle Förderung der Kinder in der Schulanfangsphase sichern – verpflichtende Einrichtung der jahrgangsgemischten Lerngruppen gründlich vorbereiten*. Drucksache 16/0326. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/16/DruckSachen/d16-0326.pdf>
- Abgeordnetenhaus Berlin. (Dezember 2008). *Mitteilung: Verlässliche Rahmenbedingungen für die Schulanfangsphase schaffen – Einrichtung jahrgangsgemischter Lerngruppen an allen Grundschulen ab 2008/09 sichern*. Drucksache 16/2024. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/16/DruckSachen/d16-2024.pdf>
- Abgeordnetenhaus Berlin. (2010a). *Kleine Anfrage: Rückstellung vom Schulbesuch*. Drucksache 16/14450.
- Abgeordnetenhaus Berlin. (Februar 2010b). *Kleine Anfrage: Jahrgangsübergreifendes Lernen an Berliner Schulen*. Drucksache 16/14083. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/16/KlAnfr/ka16-14083.pdf>
- Abgeordnetenhaus Berlin. (Dezember 2010c). *Kleine Anfrage: Umgang der Schulverwaltung mit der Heterogenität von Schulen und Leistungsdaten*. Drucksache 16/15027. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/16/KlAnfr/ka16-15027.pdf>
- Abgeordnetenhaus Berlin. (2011, 1. Januar). *Plenarprotokoll des Abgeordnetenhauses von Berlin 16/86*. Verfügbar unter <http://www.parlament-berlin.de/ados/16/IIIPlen/protokoll/plen16-086-pp.pdf>
- Abgeordnetenhaus Berlin. (2012a, 12. Januar). *Plenarprotokoll 17/6*. Plenarprotokoll. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/17/PlenarPr/p17-006-wp.pdf#page=49>
- Abgeordnetenhaus Berlin. (April 2012b). *Kleine Anfrage: Schuleingangsphase Klasse 1 und 2 nach Änderung des Paragraphen 7 der Grundschulverordnung*. Drucksache 17/10429. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/17/KlAnfr/ka17-10429.pdf>
- Abgeordnetenhaus Berlin. (September 2012c). *Kleine Anfrage: JüL und JabL – Zur Schulanfangsphase in Berlin*. Drucksache 17/11037. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/17/KlAnfr/ka17-11037.pdf>

- Abgeordnetenhaus Berlin. (März 2013). *Kleine Anfrage: Schulversuche und Modellvorhaben an Berliner Schulen*. Drucksache 17/11832. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/17/KlAnfr/ka17-11832.pdf>
- Abgeordnetenhaus Berlin. (Februar 2014). *Wortprotokoll 17/34: Sitzung des Ausschusses für Bildung, Jugend und Familie vom 13.02.2014 (17/34)*. Berlin: Abgeordnetenhaus Berlin. Verfügbar unter <http://www.parlament-berlin.de/ados/17/BildJugFam/protokoll/bjf17-034-wp.pdf>
- Abgeordnetenhaus Berlin. (Januar 2015). *Schriftliche Anfrage: Schulanfangsphase in Berlin – aktueller Stand II*. Drucksache 17/15274. Verfügbar unter <http://pardok.parlament-berlin.de/starweb/adis/citat/VT/17/SchrAnfr/s17-15274.pdf>
- Altrichter, H. & Maag Merki, K. (Hrsg.). (2010). *Handbuch Neue Steuerung im Schulsystem* (Educational Governance, Bd. 7, 1. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-92245-4>
- Altrichter, H. & Wiesinger, S. (2005). Implementation von Schulinnovationen – aktuelle Hoffnungen und Forschungswissen. *Journal für Schulentwicklung*, 9 (4), 28–36.
- Angrist, J. D. & Krueger, A. B. (1992). The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples. *Journal of the American Statistical Association*, 87 (418), 32–336. <https://doi.org/10.1080/01621459.1992.10475212>
- Autorengruppe Bildungsberichterstattung. (2012). *Bildung in Deutschland 2012. Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf*. Bielefeld: W. Bertelsmann Verlag.
- Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration* (1. Auflage). Bielefeld: W. Bertelsmann Verlag.
- Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67 (1). <https://doi.org/10.18637/jss.v067.i01>
- Baumert, J., Maaz, K., Neumann, M., Becker, M. & Dumont, H. (2013). Die Berliner Schulstrukturreform: Hintergründe, Zielstellungen und theoretischer Rahmen. In K. Maaz, J. Baumert, M. Neumann, M. Becker & H. Dumont (Hrsg.), *Die Berliner Schulstrukturreform: Bewertung durch die beteiligten Akteure und Konsequenzen des neuen Übergangsverfahrens von der Grundschule in die weiterführenden Schulen* (S. 9–34). Münster u.a.: Waxmann.
- Bedard, K. & Dhuey, E. (2006). The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects. *The Quarterly Journal of Economics*, 121 (4), 1437–1472. <https://doi.org/10.1093/qje/121.4.1437>
- Bedard, K. & Dhuey, E. (2008). *Is September better than January? The effect of school entry age laws on skill accumulation*. Working Paper. Verfügbar unter http://homes.chass.utoronto.ca/~edhuey/SeptBetterJanuary_December2008.pdf

- Bedard, K. & Dhuey, E. (2012). School-Entry Policies and Skill Accumulation Across Directly and Indirectly Affected Individuals. *Journal of Human Resources*, 47 (3), 643–683. <https://doi.org/10.3368/jhr.47.3.643>
- Bellmann, J., Schweizer, S. & Thiel, C. (2016). Nebenfolgen Neuer Steuerung unter Bedingungen von „low-stakes“ und „no-stakes“ – Qualitative und quantitative Befunde einer Untersuchung in vier Bundesländern. In Bundesministerium für Bildung und Forschung (Hrsg.), *Steuerung im Bildungssystem. Implementation und Wirkung neuer Steuerungsinstrumente im Schulwesen* (Bildungsforschung, Bd. 43, S. 208–237).
- Berthold, B. (2008a). *Einschulungsregelungen und flexible Eingangsstufe. Recherche für den Nationalen Bildungsbericht 2008 im Auftrag des Deutschen Jugendinstituts*. (Wissenschaftliche Texte / Deutsches Jugendinstitut). München: Dt. Jugendinstitut.
- Berthold, B. (2008b). *Unterricht entwickeln in der Schuleingangsphase: Grundlagen – Ziele – Anregungen*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Bierhoff, H.-W. (1990). *Psychologie hilfreichen Verhaltens* (Urban-Taschenbücher, Bd. 418). Stuttgart u.a.: Kohlhammer.
- Black, S. E., Devereux, P. J. & Salvanes, K. G. (2011). Too Young to Leave the Nest? The Effects of School Starting Age. *The Review of Economics and Statistics*, 93 (2), 455–467.
- Borman, G. D., Hewes, G. M., Overman, L. T. & Brown, S. (2003). Comprehensive School Reform and Achievement: A Meta-Analysis. *Review of Educational Research*, 73 (2), 125–230.
- Bormann, I. (2011). *Zwischenräume der Veränderung. Innovationen und ihr Transfer im Feld von Bildung und Erziehung* (1. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bos, W., Hornberg, S. & Arnold, K.-H. (Hrsg.). (2008). *IGLU-E 2006. Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Schwippert, K. & Arnold, K.-H. (Hrsg.). (2010). *IGLU 2006 – die Grundschule auf dem Prüfstand. Vertiefende Analysen zu Rahmenbedingungen schulischen Lernens*. Münster: Waxmann.
- Bos, W., Tarelli, I., Bremerich-Vos, A. & Schwippert, K. (Hrsg.). (2012). *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Wendt, H., Köller, O. & Selter, C. (Hrsg.). (2012). *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Böse, S. (2015). *Implementation von Schulreformen. Zur Auseinandersetzung von Schulleiterinnen und Schulleitern mit Innovationen im Bildungswesen*. Dissertation, Universität Potsdam. Potsdam.
- Breidenstein, G. & Rademacher, S. (2013). Vom Nutzen der Zeit. Beobachtungen und Analysen zum individualisierten Unterricht. *Zeitschrift für Pädagogik*, 59 (3), 336–356.

- Breyvogel, W. (2012). Zur Entstehung von Bildungsarmut. In W. Breyvogel (Hrsg.), *Wie aus Kindern Risikoschüler werden. Fallstudien zu den Ursachen von Bildungsarmut* (Wissen & Praxis, Bd. 153, 1. Aufl., S. 15–31). Frankfurt: Brandes & Apsel Verlag.
- Burk, K. (Hrsg.). (1996). *Mehr gestalten als verwalten! 12. Jahrgangübergreifendes Lernen in der Grundschule* (Beiträge zur Reform der Grundschule). Frankfurt am Main: Arbeitskreis Grundschule.
- Cahan, S. & Cohen, N. (1989). Age versus Schooling Effects on Intelligence Development. *Child Development*, 60, 1239–1249.
- Carle, U. (2008). *Anfangsunterricht in der Grundschule – beste Lernchancen für alle Kinder. Expertise für die Enquetekommission 'Chancen für Kinder' des Landtages Nordrhein-Westfalen vorgestellt am 11. Januar 2008.*
- Carle, U. (2013). Jahrgangübergreifendes Lernen. Worauf kommt es an? *Die Grundschulzeitschrift*, 27 (270), 4–5.
- Carle, U. & Metzen, H. (2008a). *Begabung ist anders. Die Ziele der Integrativen, individualisierten und jahrgangsgemischten Schuleingangsphase – eine öffentlichkeitswirksame, überschaubare und schulnahe Beschreibung des Zielmodells (Inhalte, Funktionen und Entwicklungswege).*, Universität Bremen, Arbeitsgebiet Grundschulpädagogik.
- Carle, U. & Metzen, H. (2008b). Projektentwicklungsbeurteilung zur Unterrichtsqualität der Flex-Schulen auf der Basis exemplarischer Unterrichtsanalysen. In K. Liebers, A. Prengel & G. Bieber (Hrsg.), *Die flexible Schuleingangsphase. Evaluationen zur Neugestaltung des Anfangsunterrichts* (S. 97–137). Weinheim: Beltz.
- Carle, U. & Metzen, H. (2009). *Die Schuleingangsphase lohnt sich! Erfolgsmomente für die bestmögliche Entwicklung des landesweiten Schulentwicklungsvorhabens 'Begleitete Schuleingangsphase' in Thüringen. Bericht der wissenschaftlichen Begleitung nach zweieinhalb Jahren 'BeSTe' (2005-2008).*
- Carle, U. & Metzen, H. (2014). *Wie wirkt jahrgangübergreifendes Lernen? Internationale Literaturübersicht zum Stand der Forschung, der praktischen Expertise und der pädagogischen Theorie. Eine wissenschaftliche Expertise des Grundschulverbandes.* Frankfurt: Grundschulverband.
- Chan, S. C. H. & Ngai, E. W. T. (2012). Electronic Learning Systems in Hong Kong Business Organizations. A Study of Early and Late Adopters. *Journal of Education for Business*, 87 (3), 170–177. <https://doi.org/10.1080/08832323.2011.586005>
- Cobley, S., McKenna, J., Baker, J. & Wattie, N. (2009). How Pervasive Are Relative Age Effects in Secondary School Education? *Journal of Educational Psychology*, 101 (2), 520–528. <https://doi.org/10.1037/a0013845>
- Coburn, C. E. (2003). Rethinking Scale: Moving beyond Numbers to Deep and Lasting Change. *Educational Researcher*, 32 (6), 3–12.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2. ed.). Hillsdale NJ u.a.: Lawrence Erlbaum Associates Publishers.

- Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, 112 (1), 155–159.
<https://doi.org/10.1037/0033-2909.112.1.155>
- Collins, L. M. & Lanza, S. T. (2010). *Latent Class and Latent Transition Analysis. With Applications in the Social, Behavioral, and Health Sciences*. Hoboken, NJ: Wiley.
- Cools, W., Fraine, B. D., van den Noortgate, W. & Onghena, P. (2009). Multilevel design efficiency in educational effectiveness research. *School Effectiveness and School Improvement*, 20 (3), 357–373. <https://doi.org/10.1080/09243450902850176>
- Cortina, K. S., Baumert, J., Leschinsky, A., Mayer, K. U. & Trommer, L. (Hrsg.). (2005). *Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick* (2. Aufl.). Reinbek bei Hamburg: Rowohlt.
- Cronbach, L. J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, 16 (3), 297–334. <https://doi.org/10.1007/BF02310555>
- Crone, D. A. & Whitehurst, G. J. (1999). Age and schooling effects on emergent literacy and early reading skills. *Journal of Educational Psychology*, 91 (4), 604–614.
- Cumming, G. (2012). *Understanding The New Statistics. Effect Sizes, Confidence Intervals, and Meta-Analysis*. New York: Routledge.
- Cumming, G. & Finch, S. (2001). A Primer on the Understanding, Use, and Calculation of Confidence Intervals that are Based on Central and Noncentral Distributions. *Educational and Psychological Measurement*, 61 (4), 532–574.
- Cumming, G. & Finch, S. (2005). Inference by Eye: Confidence Intervals and How to Read Pictures of Data. *The American psychologist*, 60 (2), 170–180.
<https://doi.org/10.1037/0003-066X.60.2.170>
- Desimone, L. (2002). How Can Comprehensive School Reform Models Be Successfully Implemented? *Review of Educational Research*, 72 (3), 433–479.
- Deutsches PISA-Konsortium (Hrsg.). (2002). *PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich*. Wiesbaden: Springer Fachmedien Wiesbaden.
<https://doi.org/10.1007/978-3-663-11042-2>
- Diefenbach, H. (2009). Der Bildungserfolg von Schülern mit Migrationshintergrund im Vergleich zu Schülern ohne Migrationshintergrund. In R. Becker (Hrsg.), *Lehrbuch der Bildungssoziologie* (1. Aufl., S. 433–457). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91711-5_15
- Ditton, H. (2000). Qualität und Qualitätssicherung im Bildungsbereich; Schule, Sozialpädagogik, Hochschule. In *Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung* (Zeitschrift für Pädagogik. Beiheft. 41, S. 73–92). Beltz.
- Ditton, H. (2002). Evaluation und Qualitätssicherung. In R. Tippelt (Hrsg.), *Handbuch Bildungsforschung* (S. 775–790). Wiesbaden: VS Verlag für Sozialwissenschaften.
https://doi.org/10.1007/978-3-322-99634-3_43

- Downey, D. B., Hippel, P. T. von & Broh, B. A. (2004). Are Schools the Great Equalizer? Cognitive Inequality during the Summer Months and the School Year. *American Sociological Review*, 69 (5), 613–635. <https://doi.org/10.1177/000312240406900501>
- Dübbers, S. & Gewerkschaft Erziehung und Wissenschaft. (2005). *Jahrgangübergreifendes Lernen in der Berliner Schule: Argumente für die Einführung; Grundschule, Klassenstufen 1 bis 3; eine Handreichung der GEW Berlin*. Berlin: GEW Berlin.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P. et al. (2007). School Readiness and Later Achievement. *Developmental Psychology*, 43 (6), 1428. <https://doi.org/10.1037/0012-1649.43.6.1428>
- Ehm, J.-H. & Hasselhorn, M. (2017). Kompensatorische Zusatzförderung zur Erhöhung der Schulbereitschaft. In U. Hartmann, M. Hasselhorn & A. Gold (Hrsg.), *Entwicklungsverläufe verstehen – Individuelle Förderung wirksam gestalten. Erkenntnisse des Frankfurter IDeA-Zentrums* (S. 295–311). Stuttgart: Kohlhammer.
- Eid, M., Gollwitzer, M. & Schmitt, M. (2013). *Statistik und Forschungsmethoden. Lehrbuch; mit Online-Materialien* (3., korrigierte Aufl.). Weinheim ua: Beltz.
- Eid, M., Gollwitzer, M. & Schmitt, M. (2015). *Statistik und Forschungsmethoden* (4. Aufl.). Weinheim: Beltz.
- Elder, T. E. & Lubotsky, D. H. (2009). Kindergarten Entrance Age and Children's Achievement Impacts of State Policies, Family Background, and Peers. *Journal of Human Resources*, 44 (3), 641–683. <https://doi.org/10.3368/jhr.44.3.641>
- Faust, G. (2006). Zum Stand der Einschulung und der neuen Schuleingangsstufe in Deutschland. *Zeitschrift für Erziehungswissenschaft*, 9 (3), 328–347. <https://doi.org/10.1007/s11618-006-0054-8>
- Faust, G. (Hrsg.). (2013). *Einschulung. Ergebnisse aus der Studie Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter (BiKS)*. Münster: Waxmann.
- Faust, G., Kluczniok, K. & Pohlmann-Rother, S. (2013). Wie sich bei den Eltern die Entscheidung zur vorzeitigen Einschulung herausbildet. In G. Faust (Hrsg.), *Einschulung. Ergebnisse aus der Studie Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter (BiKS)* (S. 185–194). Münster: Waxmann.
- Fend, H. (2009). *Neue Theorie der Schule*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Fertig, M. & Kluge, J. (2005). *The Effect of Age at School Entry on Educational Attainment in Germany*. RWI Discussion Papers, No. 27.
- Fletcher, J. & Kim, T. (2016). The effects of changes in kindergarten entry age policies on educational achievement. *Economics of Education Review*, 50, 45–62. <https://doi.org/10.1016/j.econedurev.2015.11.004>
- Fredriksson, P. & Ockert, B. (2005, 21. Juli). *Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labor Market Performance*. IZA Discussion Paper No. 1659.

- Fullan, M. (2000). The Return of Large-Scale Reform. *Journal of Educational Change*, 1 (1), 5–27. <https://doi.org/10.1023/A:1010068703786>
- Fullan, M. (2007). *The new meaning of educational change* (4. ed.). New York: Teachers College Press.
- Fullan, M. G. (1993). Why teachers must become change agents. *Educational Leadership*, 50 (6), 12.
- Gallaher, J. & Wentling, T. L. (2004). The Adoption of e-learning Across Professional Groups. *Performance Improvement Quarterly*, 17 (3), 66–85. <https://doi.org/10.1111/j.1937-8327.2004.tb00314.x>
- Geiser, C. (2010). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung* (1. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gewerkschaft Erziehung und Wissenschaft. (Juni 2006). *Jahrgangsübergreifendes Lernen in der Berliner Schule. Argumente für die Einführung* (GEW Berlin, Hrsg.). Berlin.
- GEW Berlin. (2008). *Jahrgangsmischung in der Schulanfangsphase. Beschluss des Landesvorstandes vom 28.01.08*, Gewerkschaft Erziehung und Wissenschaft. Verfügbar unter https://www.gew-berlin.de/3914_3985.php
- Gillwald, K. (2000). Konzepte sozialer Innovation. WZB Discussion Paper, No. P 00-519.
- Gold, A., Duzy, D., Rauch, W. A. & Murcia, C. Q. (2012). Relatives Lebensalter und die Entwicklung schulischer Leistungen. *Zeitschrift für Bildungsforschung*, 2 (3), 193–208. <https://doi.org/10.1007/s35834-012-0046-0>
- Goldenbaum, A. (2012). *Innovationsmanagement in Schulen: Eine empirische Untersuchung zur Implementation eines Sozialen Lernprogramms*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Goldenbaum, A. (2013). Implementation von Schulinnovationen. In M. Rürup & I. Bormann (Hrsg.), *Innovationen im Bildungswesen* (Educational Governance, S. 149–172). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-531-19701-2_7
- Goldenbaum, A., Schmittat, M., Stockmann, P. & Kuhn, H. P. (2012). *Bericht zur Evaluation des Transfer-Bildungsprojekts „Schulen übernehmen Verantwortung – Bildungswege der Lernenden“* Stiftung Brandenburger Tor und der Senatsverwaltung für Bildung, Jugend und Wissenschaft (Senatsverwaltung für Bildung, J. u. W., Hrsg.). Universität Kassel. Verfügbar unter http://www.uni-kassel.de/fb01/fileadmin/datas/fb01/Institut_fuer_Erziehungswissenschaft/Dateien/EmpirischeBildungsforschung/Evaluationsbericht-Transferprojekt_SchueV_2012.pdf
- Gräsel, C. (2010). Stichwort: Transfer und Transferforschung im Bildungsbereich. *Zeitschrift für Erziehungswissenschaft*, 13 (1), 7–20. <https://doi.org/10.1007/s11618-010-0109-8>
- Gräsel, C. (2014). Lernumwelten in Schulen. In T. Seidel & A. Krapp (Hrsg.), *Pädagogische Psychologie* (6. Aufl., S. 407–432). Weinheim: Beltz.

- Gräsel, C., Jäger, M. & Willke, H. (2006). Konzeption einer übergreifenden Transferforschung und Einbeziehung des internationalen Forschungsstandes. In R. Nickolaus & C. Gräsel (Hrsg.), *Innovation und Transfer. Expertisen zur Transferforschung* (S. 445–566). Baltmannsweiler: Schneider Verlag Hohengehren.
- Gräsel, C. & Parchmann, I. (2004). Implementationsforschung – oder: der steinige Weg, Unterricht zu verändern. *Unterrichtswissenschaft*, 32 (3), 196–214.
- Gudjons, H. (1999). *Pädagogisches Grundwissen. Überblick – Kompendium – Studienbuch* (6. Aufl.). Bad Heilbrunn: Klinkhardt.
- Gutierrez, R. & Slavin, R. E. (1992). Achievement Effects of the Nongraded Elementary School. A Best Evidence Synthesis. *Review of Educational Research*, 62 (4), 333–376. <https://doi.org/10.3102/00346543062004333>
- Hall, G. E. & Hord, S. M. (2015). *Implementing Change. Patterns, Principles, and Potholes* (4. ed.). Boston, Mass: Pearson.
- Hameyer, U. (1978). *Innovationsprozesse. Analysemodell und Fallstudien zum sozialen Konflikt in der Curriculumrevision*. Weinheim: Beltz.
- Hanke, P. (2008). Jahrgangübergreifender Unterricht – auf die Unterrichtsqualität kommt es an. *Die Grundschulzeitschrift*, 22 (217), 22.
- Hannover, B. & Kessels, U. (2011). Sind Jungen die neuen Bildungsverlierer? Empirische Evidenz für Geschlechterdisparitäten zuungunsten von Jungen und Erklärungsansätze. *Zeitschrift für Pädagogische Psychologie*, 25 (2), 89–103. <https://doi.org/10.1024/1010-0652/a000039>
- Hasselhorn, M., Goldhammer, A. von & Weber, A. (2008). Belohnungsaufschub als volitionale Kompetenz. Ein relevanter Bereich für die Schuleingangsdagnostik? *Psychologie in Erziehung und Unterricht*, 55, 123–131.
- Hasselhorn, M., Köller, O., Maaz, K. & Zimmer, K. (2014). Implementation wirksamer Handlungskonzepte im Bildungsbereich als Forschungsaufgabe. *Psychologische Rundschau*, 65 (3), 140–149. <https://doi.org/10.1026/0033-3042/a000216>
- Hasselhorn, M. & Lohaus, A. (2008). Entwicklungsvoraussetzungen und Herausforderungen des Schuleintritts. In M. Hasselhorn & R. K. Silbereisen (Hrsg.), *Entwicklungspsychologie des Säuglings- und Kindesalters* (Enzyklopädie der Psychologie Theorie und Forschung Entwicklungspsychologie, S. 409–428). Göttingen: Hogrefe.
- Hattie, J. A. C. (2009). *Visible Learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hecht, P. (2010, 28. September). Einschulung: Debatte um frühe Schulpflicht reißt nicht ab. *Der Tagesspiegel*. Zugriff am 27.03.2017. Verfügbar unter <http://www.tagesspiegel.de/berlin/schule/einschulung-debatte-um-fruehe-schulpflicht-reisst-nicht-ab/1943718.html>
- Heinig, E., Thoren, K. & Brunner, M. (2014). *Erste Ergebnisse zur Evaluation der Früheinschulung in Berlin*. Berlin: Freie Universität Berlin; Institut für Schulqualität der Länder Berlin und Brandenburg.

- Helmke, A. (2014). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (Unterricht verbessern – Schule entwickeln, 5. Aufl.). Seelze-Velber: Klett/Kallmeyer.
- Hill, C. J., Bloom, H. S., Rebeck Black, A. & Lipsey, M. W. (2008). Empirical Benchmarks for Interpreting Effect Sizes in Research. *Child Development Perspectives*, 2 (3), 172–177. <https://doi.org/10.1111/j.1750-8606.2008.00061.x>
- Hill, P. W. & Rowe, K. J. (1996). Multilevel Modelling in School Effectiveness Research. *School Effectiveness and School Improvement*, 7 (1), 1–34. <https://doi.org/10.1080/0924345960070101>
- Holtappels, H. G. (1995). Schulkultur und Innovation – Ansätze, Trends und Perspektiven der Schulentwicklung. In H. G. Holtappels (Hrsg.), *Entwicklung von Schulkultur: Ansätze und Wege schulischer Erneuerung*. Neuwied ua: Luchterhand.
- Hox, J. J. (2010). *Multilevel Analysis. Techniques and Applications* (Quantitative methodology series, 2. ed.). New York, NY: Routledge.
- Hunneshagen, H. (2005). *Innovationen in Schulen: Identifizierung implementationsfördernder und -hemmender Bedingungen des Einsatzes neuer Medien*. Münster: Waxmann.
- IQB. (2017a). *Bildungsstandards*, Institut zur Qualitätsentwicklung im Bildungswesen. Zugriff am 28.03.2017. Verfügbar unter <https://www.iqb.hu-berlin.de/bista>
- IQB. (2017b). *Ländervergleich und IQB-Bildungstrend*, Institut zur Qualitätsentwicklung im Bildungswesen. Verfügbar unter <https://www.iqb.hu-berlin.de/bt>
- ISB. (2016). *Vergleichsarbeiten - ISB - Staatsinstitut für Schulqualität und Bildungsforschung*. Zugriff am 22.04.2016. Verfügbar unter <https://www.isb.bayern.de/grundschule/vergleichsarbeiten/>
- ISQ. (2017). *Vergleichsarbeiten*, Institut für Schulqualität der Länder Berlin und Brandenburg e.V. Verfügbar unter <https://www.isq-bb.de/wordpress/vergleichsarbeiten/>
- Jäger, M. (2004). *Transfer in Schulentwicklungsprojekten*. Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-322-83388-4>
- Kammermeyer, G. (2000). *Schulfähigkeit. Kriterien und diagnostische/prognostische Kompetenz von Lehrerinnen, Lehrern und Erzieherinnen*. Bad Heilbrunn/Obb.: Klinkhardt.
- Kastirke, N. (Hrsg.). (2006). *Die neue Schuleingangsphase als Thema der Schulentwicklung: Forschung – Stolpersteine – Praxisempfehlungen*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Kastirke, N. & Jennessen, S. (2006). Praxisempfehlungen und Gelingensbedingungen. In N. Kastirke (Hrsg.), *Die neue Schuleingangsphase als Thema der Schulentwicklung: Forschung – Stolpersteine – Praxisempfehlungen* (S. 275–291). Baltmannsweiler: Schneider Verlag Hohengehren.
- Kawaguchi, D. (2011). Actual age at school entry, educational outcomes, and earnings. *Journal of the Japanese and International Economies*, 25 (2), 64–80. <https://doi.org/10.1016/j.jjie.2009.02.002>

- Kemmler, L. & Heckhausen, H. (1962). Ist die sogenannte „Schulreife“ ein Reifungsproblem. In K. Ingenkamp & L. Schenk-Danzinger (Hrsg.), *Praktische Erfahrungen mit Schulreifetests* (Psychologische Praxis, Bd. 30, S. 52–89). Basel: Karger.
- Kern, A. (1951). *Sitzenbleiberelend und Schulreife. Ein psychologisch-pädagogischer Beitrag zu einer inneren Reform der Grundschule*. Freiburg: Herder.
- Kern, A. (Hrsg.). (1970a). *Die Schulreife in pädagogischer und psychologischer Sicht*. Frankfurt am Main: Akademische Verlagsgesellschaft.
- Kern, A. (1970b). Tendenzen zur Frühbeschulung in Deutschland. In A. Kern (Hrsg.), *Die Schulreife in pädagogischer und psychologischer Sicht* (S. 159–162). Frankfurt am Main: Akademische Verlagsgesellschaft.
- Kerstan, T. & Spiewak, M. (2008, 19. September). Schulsystem: »Schule ist die große Gleichmacherin«. Was muss getan werden, um Deutschlands Bildungssystem gerechter zu machen? Ein Gespräch mit dem Erziehungswissenschaftler Jürgen Baumert. *Die Zeit*. Zugriff am 05.03.2017. Verfügbar unter <http://www.zeit.de/2008/39/C-Interview-Baumert>
- Klein, K. J. & Knight, A. P. (2005). Innovation Implementation. *Current Directions in Psychological Science*, 14 (5), 243–246. <https://doi.org/10.1111/j.0963-7214.2005.00373.x>
- Klesmann, M. (2011, 17. September). Zwei Schulleiterinnen streiten über das jahrgangsübergreifende Lernen in der Grundschule. Die eine will es ausweiten - die andere will es loswerden: "Von oben herab verordnet". *Berliner Zeitung*. Zugriff am 28.03.2017. Verfügbar unter <http://www.berliner-zeitung.de/zwei-schulleiterinnen-streiten-ueber-das-jahrgangsuebergreifende-lernen-in-der-grundschule--die-eine-will-es-ausweiten---die-andere-will-es-loswerden--von-oben-herab-verordnet--15151678>
- Klesmann, M. (2013, 28. Januar). GEW für Auswertung der frühen Einschulung. *Berliner Zeitung*. Zugriff am 27.03.2017. Verfügbar unter <http://www.berliner-zeitung.de/gew-fuer-auswertung-der-fruehen-einschulung-6681830>
- KMK. (1997). *Empfehlungen zum Schulanfang – Beschluß der Kultusministerkonferenz vom 24.10.1997*.
- KMK. (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. München: LinkLuchterhand.
- KMK. (2014). *Standards für die Lehrerbildung. Bildungswissenschaften*. Beschluss der Kultusministerkonferenz vom 16.12.2004 i. d. F. vom 12.06.2014.
- KMK. (2015). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring. Beschluss der 350. Kultusministerkonferenz vom 11.06.2015*. Berlin: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland.
- Koalitionsvereinbarung zwischen SPD Landesverband Berlin und CDU Landesverband Berlin für die Legislaturperiode 2011-2016. (2011). *Berliner Perspektiven für starke Wirtschaft, gute Arbeit und sozialen Zusammenhalt*.
- Koch, S. & Gräsel, C. (2004). Schulreform und Neue Steuerung – erziehungs- und verwaltungswissenschaftliche Perspektiven. In S. Koch (Hrsg.), *Schulen für die Zukunft*:

- neue Steuerung im Bildungswesen* (S. 3–24). Baltmannsweiler: Schneider Verlag Hohengehren.
- Kolleck, N. (2013). Social network analysis in innovation research. Using a mixed methods approach to analyze social innovations. *European Journal of Futures Research*, 1 (1), 93. <https://doi.org/10.1007/s40309-013-0025-2>
- Kolleck, N. (2014). Innovations through networks: understanding the role of social relations for educational innovations. *Zeitschrift für Erziehungswissenschaft*, 17 (5), 47–64. <https://doi.org/10.1007/s11618-014-0547-9>
- Konsortium Bildungsberichterstattung. (2006). *Bildung in Deutschland. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: W. Bertelsmann Verlag.
- Kristen, C. (2006). Ethnische Diskriminierung in der Grundschule? Die Vergabe von Noten und Bildungsempfehlungen. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 58 (1), 79–97. <https://doi.org/10.1007/s11575-006-0004-y>
- Kristen, C. & Granato, N. (2007). The educational attainment of the second generation in Germany. Social origins and ethnic inequality. *Ethnicities*, 7 (3), 343–366. <https://doi.org/10.1177/1468796807080233>
- Kucharz, D. & Wagener, M. (2007). *Jahrgangsübergreifendes Lernen: eine empirische Studie zu Lernen, Leistung und Interaktion von Kindern in der Schuleingangsphase* (Bd. 4). Baltmannsweiler: Schneider Verlag Hohengehren.
- Kuhl, P., Felbrich, A., Richter, D., Stanat, P. & Pant, H. A. (2013). Die Jahrgangsmischung auf dem Prüfstand: Effekte jahrgangsübergreifenden Lernens auf Kompetenzen und sozio-emotionales Wohlbefinden von Grundschülerinnen und Grundschulern. In R. Becker & A. Schulze (Hrsg.), *Bildungskontexte. Strukturelle Voraussetzungen und Ursachen ungleicher Bildungschancen* (S. 299–324). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-531-18985-7_11
- Kuhl, P. & Harych, P. (2007). *Vergleichsarbeiten der Jahrgangsstufe 2 im Schuljahr 2006/2007. Länderbericht Berlin*. Berlin: Institut für Schulqualität der Länder Berlin und Brandenburg.
- Kuhl, P. & Harych, P. (2008). *VERA 3: Vergleichsarbeiten in der Jahrgangsstufe 3 im Schuljahr 2007/2008. Länderbericht Berlin*. Berlin: Institut für Schulqualität der Länder Berlin und Brandenburg.
- Laging, R. (Hrsg.). (2010). *Altersgemischtes Lernen in der Schule* (Grundlagen der Schulpädagogik, Bd. 28, 4. Aufl.). Baltmannsweiler: Schneider Verlag Hohengehren.
- Liebers, K., Prengel, A. & Bieber, G. (Hrsg.). (2008). *Die flexible Schuleingangsphase. Evaluationen zur Neugestaltung des Anfangsunterrichts*. Weinheim und Basel: Beltz.
- Lincove, J. A. & Painter, G. (2006). Does the Age That Children Start Kindergarten Matter? Evidence of Long-Term Educational and Social Outcomes. *Educational Evaluation and Policy Analysis*, 28 (2), 153–179. <https://doi.org/10.3102/01623737028002153>

- Loogma, K., Kruusvall, J. & Ümarik, M. (2012). E-learning as innovation. Exploring innovativeness of the VET teachers' community in Estonia. *Computers & Education*, 58 (2), 808–817. <https://doi.org/10.1016/j.compedu.2011.10.005>
- Luchte, K. (2005). *Implementierung pädagogischer Konzepte in sozialen Systemen. Ein systemtheoretischer Beratungsansatz*. Weinheim: Beltz.
- Lüdtke, O., Robitzsch, A., Köller, O. & Winkelmann, H. (2010). Kausale Effekte in der Empirischen Bildungsforschung: Ein Vergleich verschiedener Ansätze zur Schätzung des Effekts des Einschulungsalters. In W. Bos, E. Klieme, O. Köller & J. Baumert (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung* (S. 257–284). Münster u.a.: Waxmann.
- Lumley, T. (2014). *Survey: Analysis of Complex Survey Samples. R package version, 3, 5*.
- Maag Merki, K., Emmerich, M. & Holmeier, M. (2015). Further development of educational effectiveness theory in a multilevel context. From theory to methodology and from empirical evidence back to theory. *School Effectiveness and School Improvement*, 26 (1), 4–9. <https://doi.org/10.1080/09243453.2014.938930>
- Maaz, K., Baumert, J., Neumann, M., Becker, M. & Dumont, H. (Hrsg.). (2013). *Die Berliner Schulstrukturreform: Bewertung durch die beteiligten Akteure und Konsequenzen des neuen Übergangsverfahrens von der Grundschule in die weiterführenden Schulen*. Münster u.a.: Waxmann.
- Martínez, J. F. (2012). Consequences of omitting the classroom in multilevel models of schooling. An illustration using opportunity to learn and reading achievement. *School Effectiveness and School Improvement*, 23 (3), 305–326. <https://doi.org/10.1080/09243453.2012.678864>
- Maxwell, S. E. & Delaney, H. D. (2004). *Designing Experiments and Analyzing Data. A Model Comparison Perspective*. New York, London: Psychology Press.
- Mayer, S. E. & Knutson, D. (1999). Does the Timing of School Affect How Much Children Learn? In S. E. Mayer & P. E. Peterson (Hrsg.), *Earning and Learning. How Schools Matter* (S. 79–102). Washington, D.C.: Brookings Institution Press.
- Morgan, D. L. & Morgan, R. K. (2009). *Single-Case Research Methods for the Behavioral and Health Sciences*. Thousand Oaks, California: Sage.
- Müller-Lissner, A. (2015, 2. März). Wann ist der beste Zeitpunkt für den Schulanfang?: Abwarten ist nicht die beste Lösung. *Der Tagesspiegel*. Zugriff am 27.03.2017. Verfügbar unter <http://www.tagesspiegel.de/wissen/wann-ist-der-beste-zeitpunkt-fuer-den-schulanfang-abwarten-ist-nicht-die-beste-loesung/11441496.html>
- Mullis, I. V., Martin, M. O., Gonzalez, E. J. & Chrostowski, S. J. (2004). *TIMSS 2003 – International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, Mass.: TIMSS & PIRLS International Study Center; Lynch School of Education, Boston College.

- Musch, J. & Grondin, S. (2001). Unequal Competition as an Impediment to Personal Development. A Review of the Relative Age Effect in Sport. *Developmental Review*, 21 (2), 147–167. <https://doi.org/10.1006/drev.2000.0516>
- Muthén, L. K. & Muthén, B. O. (1998-2014) Mplus User's Guide [Computer software]. Los Angeles, CA: Muthén & Muthén.
- Nickel, H. (1990). Das Problem der Einschulung aus ökologisch-systemischer Perspektive. *Psychologie in Erziehung und Unterricht*, 37 (3), 217–227.
- Nickel, H. & Schmidt, U. (1975). *Vorschulkind und Schulanfänger* (Goldmann-Sachbücher, Bd. 11129). München: Goldmann.
- Nickel, H. & Schmidt-Denter, U. (1995). *Vom Kleinkind zum Schulkind. Eine entwicklungspsychologische Einführung für Erzieher, Lehrer und Eltern* (5. Aufl.). München: Reinhardt.
- Nylund, K. L., Asparouhov, T. & Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling. A Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal*, 14 (4), 535–569. <https://doi.org/10.1080/10705510701575396>
- O'Day, J. (2002). Complexity, Accountability, and School Improvement. *Harvard educational review*, 72 (3), 293–329. <https://doi.org/10.17763/haer.72.3.021q742t8182h238>
- Odden, A. R. (1991). New Patterns of Education Policy Implementation and Challenges for the 1990s. In A. R. Odden (Hrsg.), *Education policy implementation*. Albany: State University of New York Press.
- OECD. (2001a). *Education at a Glance 2001. OECD Indicators*. Paris: Organisation for Economic Co-operation and Development.
- OECD. (2001b). *Knowledge and Skills for Life*. Paris: Organisation for Economic Co-operation and Development.
- OECD. (2002). *Reading for Change. Performance and Engagement Across Countries. Results from PISA 2000*. Paris: Organisation for Economic Co-operation and Development.
- OECD. (2006). *Where Immigrant Students Succeed. A Comparative Review of Performance and Engagement in PISA 2003*. Paris: Organisation for Economic Co-operation and Development.
- OECD. (2007). *PISA 2006. Science Competencies for Tomorrow's World (Volume 1)*. Paris: Organisation for Economic Co-operation and Development. <https://doi.org/10.1787/9789264040014-en>
- OECD. (2010a). *PISA 2009 Results: Learning Trends: Changes in Student Performance Since 2000 (Volume V)*. Paris: Organisation for Economic Co-operation and Development.
- OECD. (2010b). *PISA 2009 Results: Overcoming Social Background. Equity in Learning Opportunities and Outcomes (Volume II)*. Paris: Organisation for Economic Co-operation and Development.

- OECD. (2010c). *PISA 2009 Results: What Makes a School Successful?: Resources, Policies and Practices (Volume IV)*. Paris: Organisation for Economic Co-operation and Development.
- OECD. (2011). *Starting Strong III. A Quality Toolbox for Early Childhood Education and Care*. Paris: Organisation for Economic Co-operation and Development.
- OECD. (2015). *Education Policy Outlook 2015. Making Reforms Happen*. Paris: Organisation for Economic Co-operation and Development.
- Pant, H. A., Vock, M., Pöhlmann, C. & Köller, O. (2008a). Eine modellbasierte Erfassung der Auseinandersetzung von Lehrkräften mit den länderübergreifenden Bildungsstandards. *Pädagogische Professionalität als Gegenstand empirischer Forschung*, 245–260.
- Pant, H. A., Vock, M., Pöhlmann, C. & Köller, O. (2008b). Offenheit für Innovationen. Befunde aus einer Studie zur Rezeption der Bildungsstandards bei Lehrkräften und Zusammenhänge mit Schülerleistungen. *Zeitschrift für Pädagogik*, 54 (6), 827–845.
- Pietsch, M., Scholand, B. & Schulte, K. (Hrsg.). (2015). *Schulinspektion in Hamburg. Der erste Zyklus 2007-2013: Grundlagen, Befunde und Perspektiven* (HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Bd. 15). Münster: Waxmann.
- Pohlmann-Rother, S., Kratzmann, J. & Faust, G. (2013). Schulfähigkeit aus der Sicht von Eltern, Erzieherinnen und Lehrkräften. In G. Faust (Hrsg.), *Einschulung. Ergebnisse aus der Studie Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter (BiKS)* (S. 153–170). Münster: Waxmann.
- Prenzel, A. (2006). *Pädagogik der Vielfalt. Verschiedenheit und Gleichberechtigung in interkultureller, feministischer und integrativer Pädagogik*. 3. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Prenzel, M., Blum, W. & Klieme, E. (2015). The Impact of PISA on Mathematics Teaching and Learning in Germany. In K. Stacey & R. Turner (Eds.), *Assessing Mathematical Literacy* (pp. 239–248). Cham: Springer International Publishing. Zugriff am 03.02.2016. Verfügbar unter http://link.springer.com/10.1007/978-3-319-10121-7_12
- Puhani, P. A. & Weber, A. M. (2005). *Does the Early Bird Catch the Worm? Instrumental Variable Estimates of Educational Effects of Age of School Entry in Germany*. *Diskussionspapiere des Fachbereichs Wirtschaftswissenschaften, Universität Hannover*, No. 332.
- R Core Team. (2016) R: A Language and Environment for Statistical Computing [Computer software]. Wien: R Foundation for Statistical Computing. Verfügbar unter <https://www.R-project.org>.
- Ramseger, J. (2004). *Grundschulen entwickeln sich. Ergebnisse des Berliner Schulversuchs Verlässliche Halbtagsgrundschule*. Münster: Waxmann.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks: Sage.

- Reardon, S. F. & Galindo, C. (2009). The Hispanic-White Achievement Gap in Math and Reading in the Elementary Grades. *American Educational Research Journal*, 46 (3), 853–891. <https://doi.org/10.3102/0002831209333184>
- Reinmann, G. & Vohle, F. (2004). Implementation als Designprozess. In Reinmann, G., & Mandl, H. (Hrsg.), *Psychologie des Wissensmanagements: Perspektiven, Theorien und Methoden* (S. 234–247). Göttingen: Hogrefe.
- Reinmann, G. (2005). Innovation ohne Forschung? Ein Plädoyer für den Design-Based Research-Ansatz in der Lehr-Lernforschung. *Unterrichtswissenschaft*, 33 (1), 52–69.
- Rekkor, S., Ümarik, M. & Loogma, K. (2013). Adoption of national curricula by vocational teachers in Estonia. *Journal of Vocational Education & Training*, 32 (4), 1–18. <https://doi.org/10.1080/13636820.2013.841277>
- Reynolds, D., Sammons, P., Fraine, B. D., van Damme, J., Townsend, T., Teddlie, C. et al. (2014). Educational effectiveness research (EER): a state-of-the-art review. *School Effectiveness and School Improvement*, 25 (2), 197–230. <https://doi.org/10.1080/09243453.2014.885450>
- Rogers, E. M. (2003). *Diffusion of Innovations* (5. ed., Free Press trade paperback ed.). New York u.a.: Free Press.
- Rogers, E. M. & Shoemaker, F. F. (1971). *Communication of Innovations. A Cross-Cultural Approach*. New York: Free Press u.a.
- Roßbach, H. G. (2010). Empirische Vergleichsuntersuchungen zu den Auswirkungen von jahrgangsheterogenen und jahrgangshomogenen Klassen. In R. Laging (Hrsg.), *Altersgemischtes Lernen in der Schule* (Grundlagen der Schulpädagogik, Bd. 28, 4. Aufl., S. 80–91). Baltmannsweiler: Schneider Verlag Hohengehren.
- Rost, J. (1996). *Lehrbuch Testtheorie, Testkonstruktion*. Bern: Huber.
- Scheerens, J. & Bosker, R. J. (1997). *The Foundations of Educational Effectiveness* (1. Aufl.). Oxford u.a.: Pergamon.
- Schenk-Danzinger, L. (1969). *Schuleintrittsalter, Schulfähigkeit und Lesereife. Untersuchungen über die Bedeutung von Schulreifegruppentests und vorschulischer Förderung* (Deutscher Bildungsrat: Gutachten und Studien der Bildungskommission, Bd. 7, 1. Aufl.). Stuttgart: Klett.
- Schneider, M. (2013, 23. Januar). Kommentar: Warum die Abkehr der Berliner Schulen von Jül gut ist. *Berliner Morgenpost*. Zugriff am 28.03.2017. Verfügbar unter <http://www.morgenpost.de/berlin-aktuell/article113070168/Warum-die-Abkehr-der-Berliner-Schulen-von-Juel-gut-ist.html>
- Schnell, R., Hill, P. B. & Esser, E. (2005). *Methoden der empirischen Sozialforschung* (7., völlig überarb. und erw. Aufl.). München: Oldenbourg.
- Schnepf, S. V. (2004). *How Different Are Immigrants? A Cross-Country and Cross-Survey Analysis of Educational Achievement*. IZA Discussion Paper No. 1398.

- Schober, P. S. & Spieß, C. K. (2012). Frühe Förderung und Betreuung von Kindern: Bedeutende Unterschiede bei der Inanspruchnahme besonders in den ersten Lebensjahren. *DIW Wochenbericht*, 79 (43), 17-28.
- Schober, P. S. & Stahl, J. F. (2014). Childcare Trends in Germany: Increasing Socio-Economic Disparities in East and West. *DIW Economic Bulletin*, 4 (11), 51-58.
- SenBJF. (2017). *Anmeldung der Schulanfänger*; Senatsverwaltung für Bildung, Jugend und Familie. Zugriff am 28.03.2017. Verfügbar unter <https://www.berlin.de/sen/bildung/schule/bildungswege/grundschule/anmeldung>
- SenBJS. Verordnung über Auswertungen von Vergleichsarbeiten und zentralen Prüfungsarbeiten. VergleichsVO, Berlin.
- SenBJW. (2011). *Zweite Runde Schulinspektion in Berlin. Informationen für Eltern, Schülerinnen und Schüler, Lehrkräfte und alle Interessierten* (Senatsverwaltung für Bildung, J. u. W., Hrsg.). Senatsverwaltung für Bildung, Jugend und Wissenschaft.
- SenBJW. (2013a). *Reformen auf einen Blick - Berlin.de*.
- SenBJW. (2013b, 28. Mai). *Fachbrief Grundschule Nr. 4. Unterschiedliche Verweildauer in der flexiblen Schulanfangsphase* (SenBJW, Hrsg.). Berlin: Senatsverwaltung für Bildung, Jugend und Wissenschaft.
- SenBJW. (2016). *Schulverzeichnis*, Senatsverwaltung für Bildung, Jugend und Wissenschaft. Verfügbar unter <https://www.berlin.de/sen/bildung/schule/berliner-schulen/schulverzeichnis/>
- SenBWF. (2004 i.d.F.v. 2016). Schulgesetz für das Land Berlin. SchulG.
- SenBWF. (2004). Schulgesetz für das Land Berlin. Schulgesetz - SchulG.
- SenBWF. (2005). Verordnung über den Bildungsgang der Grundschule. Grundschulverordnung - GsVO, Berlin.
- SenBWF. (2005 i.d.F.v. 2012). Verordnung über den Bildungsgang der Grundschule. Grundschulverordnung - GV, Berlin.
- SenBWF. (Februar 2012). *Zweite Runde Schulinspektion in Berlin. Handbuch ab Schuljahr 2011/12* (Senatsverwaltung für Bildung, J. u. W., Hrsg.). Berlin: Senatsverwaltung für Bildung, Jugend und Wissenschaft.
- SenSJS (2004). Grundschulreform 2000 – Neue Perspektiven für die Berliner Grundschule. In J. Ramseger (Hrsg.), *Grundschulen entwickeln sich: Ergebnisse des Berliner Schulversuchs Verlässliche Halbtagsgrundschule*. Münster u.a.: Waxmann.
- Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference* [Nachdr.]. Belmont, CA: Wadsworth Cengage Learning.
- Singhal, A. & Dearing, J. W. (Eds.). (2006). *Communication of Innovations. A Journey With Everett Rogers*. New Delhi: Sage Publications.
- Snijders, T. A. B. & Bosker, R. J. (2012). *Multilevel analysis. An Introduction to Basic and Advanced Multilevel Modeling* (2. ed.). Los Angeles u.a.: Sage Publications.

- Spillane, J. P. (2000). Cognition and Policy Implementation. District Policymakers and the Reform of Mathematics Education. *Cognition and Instruction*, 18 (2), 141–179. https://doi.org/10.1207/S1532690XCI1802_01
- Stahl, J. F. (2015). *Wer nutzt welche Qualität? Zusammenhänge zwischen sozioökonomischer Herkunft und Kita-Qualität* (DIW Roundup: Politik im Fokus Nr. 73). Berlin: Deutsches Institut für Wirtschaftsforschung (DIW). Verfügbar unter <https://www.econstor.eu/bitstream/10419/117371/1/834458748.pdf>
- Stamm, M. (2013). Das Konzept der Schulfähigkeit. In M. Stamm & D. Edelmann (Hrsg.), *Handbuch frühkindliche Bildungsforschung* (S. 273–284). Wiesbaden: Springer Fachmedien Wiesbaden.
- Stanat, P., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M. et al. (2002). *PISA 2000. Die Studie im Überblick: Grundlagen, Methoden und Ergebnisse*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Stipek, D. (2002). *At what age should children enter kindergarten?: A question for policy makers and parents*: Society for Research in Child Development.
- Stipek, D. & Byler, P. (2001). Academic achievement and social behaviors associated with age of entry into kindergarten. *Journal of Applied Developmental Psychology*, 22 (2), 175–189. [https://doi.org/10.1016/S0193-3973\(01\)00075-2](https://doi.org/10.1016/S0193-3973(01)00075-2)
- Strom, B. (2004). *Student achievement and birthday effects*. Unpublished manuscript, Norwegian University of Science and Technology.
- Stuchlik, E. (2002). Schulanfang auf neuen Wegen - pädagogische Reflexionen zu einem Modellversuch. *Grundschulunterricht*, 49 (9), 35–38.
- Supovitz, J. A. & Weinbaum, E. H. (Hrsg.). (2008). *The Implementation Gap: Understanding Reform in High Schools*. New York: Teachers College Press.
- Terhart, E. (2002). *Nach Pisa. Bildungsqualität entwickeln*. Hamburg: Europäische Verlagsanstalt.
- Terhart, E. & Klieme, E. (2006). Kooperation im Lehrerberuf: Forschungsproblem und Gestaltungsaufgabe. Zur Einführung in den Thementeil. *Zeitschrift für Pädagogik*, 52 (2), 163–166.
- Tervooren, A. (2010). Expertendiskurse zur Schulfähigkeit im Wandel. Zur Ausbreitung von Diagnostik. In D. Bühler-Niederberger, J. Mierendorff & A. Lange (Hrsg.), *Kindheit zwischen fürsorglichem Zugriff und gesellschaftlicher Teilhabe* (VS Research. Kindheit als Risiko und Chance, 1. Aufl., S. 253–271). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92382-6_12
- Thoren, K., Bröder, C., Gelke, A.-M. & Brunner, M. (2016). *Evaluation des Jahrgangsübergreifenden Lernens in der Berliner Schulanfangsphase*. Berlin: Freie Universität Berlin, Institut für Schulqualität der Länder Berlin und Brandenburg.
- Thoren, K., Heinig, E. & Brunner, M. (2016). Relative Age Effects in Mathematics and Reading: Investigating the Generalizability across Students, Time and Classes. *Frontiers in psychology*, 7.

- Trempler, K., Schellenbach-Zell, J. & Gräsel, C. (2012). Effekte des Transfermodellversuchsprogramms „Transfer-21“ auf Unterrichts- und Schulebene. In BMBF (Hrsg.), *Bildung für nachhaltige Entwicklung. Beiträge der Bildungsforschung* (Bildung, Bd. 39, Bd. 39, S. 25–42). Bonn: Bundesministerium für Bildung und Forschung, Referat Bildungsforschung.
- Trempler, K., Schellenbach-Zell, J. & Gräsel, C. (2013). Der Einfluss der Motivation von Lehrpersonen auf den Transfer von Innovationen. In M. Rürup & I. Bormann (Hrsg.), *Innovationen im Bildungswesen* (Educational Governance, S. 329–347). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-531-19701-2_14
- Veenman, S. (1995). Cognitive and Noncognitive Effects of Multigrade and Multi-Age Classes. A Best-Evidence Synthesis. *Review of Educational Research*, 65 (4), 319–381. <https://doi.org/10.3102/00346543065004319>
- Vieth-Entus, S. (2011, 15. September). Brandbrief: Immer mehr Schulen gegen Jahrgangsmischung. *Der Tagesspiegel*. Verfügbar unter <http://www.tagesspiegel.de/berlin/brandbrief-immer-mehr-schulen-gegen-jahrgangsmischung/4613800.html>
- Vieth-Entus, S. (2012, 4. Oktober). Rückstellungen in Berlin: Zur Schule mit fünf: Widerstand nimmt deutlich zu. *Der Tagesspiegel*. Zugriff am 27.03.2017. Verfügbar unter <http://www.tagesspiegel.de/berlin/schule/rueckstellungen-in-berlin-zur-schule-mit-fuenf-widerstand-nimmt-deutlich-zu-/7211564.html>
- Vieth-Entus, S. (2013a, 28. Januar). Einschulung von Fünfjährigen: Stress bleibt Kindersache. *Der Tagesspiegel*. Zugriff am 27.03.2017. Verfügbar unter <http://www.tagesspiegel.de/berlin/einschulung-von-fuenfjaehrigen-stress-bleibt-kindersache/7695070.html>
- Vieth-Entus, S. (2013b, 12. März). Umstrittene Früheinschulung: Viele Eltern klagen über Willkür. *Der Tagesspiegel*. Zugriff am 27.03.2017. Verfügbar unter <http://www.tagesspiegel.de/berlin/schule/umstrittene-frueheinschulung-viele-eltern-klagen-ueber-willkuer/7911368.html>
- Wehner, F. (2013). Entscheidungsprozesse im Zusammenhang mit einer Rückstellung. In G. Faust (Hrsg.), *Einschulung. Ergebnisse aus der Studie Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter (BiKS)* (S. 195–208). Münster: Waxmann.
- Weick, K. E. (1976). Educational Organizations as Loosely Coupled Systems. *Administrative Science Quarterly*, 21 (1), 1–19. <https://doi.org/10.2307/2391875>
- Wendt, P. (2007). Auf dem Weg zum jahrgangsübergreifenden Unterrichten. Fünf Maximen zu Schulentwicklung, Strukturen und Qualität. *Praxis Grundschule*, 30 (6), 4–5.
- Wilde, D. (2006). *Schulversuch „Unterricht in jahrgangsübergreifenden Lerngruppen (JÜL)“*. Ergebnisse des Schulversuchs und Schlussfolgerungen für die flexible Schulanfangsphase. Berlin. Verfügbar unter http://www.dagmarwilde.de/Saph/pdfs/sv_juel_ergebn280406.pdf

- Wilde, D. (2008, 6. Juni). *Entwicklung des jahrgangsübergreifenden Lernens in der Berliner Grundschule*. Berlin. Verfügbar unter http://bildungsserver.berlin-brandenburg.de/fileadmin/bbb/schule/lehren_und_lernen/schulanfang/EntwSAPH.Juelseit92.pdf
- Wilde, D. (2014). *Jahrgangsübergreifendes Lernen in der Berliner Schulanfangsphase*. Persönliches Gespräch am 07.05.2014. SenBJW.
- Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54 (8), 594–604.
- Wottawa, H. & Thierau, H. (2003). *Lehrbuch Evaluation* (Psychologie-Lehrbuch, 3. Aufl.). Bern: Huber.
- Wu, M. L., Adams, R. J. & Wilson, M. R. (1997). ConQuest: Multi-aspect test software [computer program]. *Camberwell, Vic.: Australian Council for Educational Research*.
- Wurster, S. & Gärtner, H. (2013). Erfassung von Bildungsprozessen im Rahmen von Schulinspektion und deren potenzieller Nutzen für die empirische Bildungsforschung. *Unterrichtswissenschaft*, 41 (3), 217–236.
- Yildirim, S., Zayim, N. & Saka, O. (2006). Technology Adoption of Medical Faculty in Teaching: Differentiating Factors in Adopter Categories. *Journal of Educational Technology & Society*, 9 (2), 213–222.
- Zander, L., Kolleck, N. & Hannover, B. (Hrsg.). (2014). *Soziale Netzwerkanalyse in Bildungsforschung und Bildungspolitik* (Zeitschrift für Erziehungswissenschaft Sonderheft, Bd. 26). Wiesbaden: Springer Fachmedien Wiesbaden.

Anhang

A.1.Übersicht der Schulgesetze und -verordnungen in den Bundesländern

Baden-Württemberg

SchG, vom 01.08.1983 (letzte berücksichtigte Änderung: § 15 geändert durch Artikel 1 des Gesetzes vom 23.02.2016 (GBl. S. 163): Schulgesetz für Baden-Württemberg.

Bayern

BayEUG, vom 31.05.2000 (GVBl. S. 414, 632, BayRS 01.01.2230-K), zuletzt geändert durch Gesetz vom 13.12.2016 (GVBl. S. 371): Bayerisches Gesetz über das Erziehungs- und Unterrichtswesen.

Brandenburg

Brandenburgisches Schulgesetz - BbgSchulG, vom 02.08.2002 (GVBl.I/02, [Nr. 08], S.78) zuletzt geändert durch Artikel 7 des Gesetzes vom 25.01.2016 (GVBl.I/16, [Nr. 5]): Gesetz über die Schulen im Land Brandenburg.

Grundschulverordnung - GV, vom 02.08.2007 (GVBl.II/07, [Nr. 16], S.190) zuletzt geändert durch Verordnung vom 06.01.2017 (GVBl.II/17, [Nr. 1]): Verordnung über den Bildungsgang der Grundschule, Brandenburg.

Bremen

BremSchulG, vom 28.06.2005 (Brem.GBl. S. 260, 388, 398 – 223-a-5), zuletzt geändert durch Gesetz vom 22.07.2014 (Brem.GBl. S. 362): Bremisches Schulgesetz.

Grundschulverordnung, vom 01.08.2012 (Brem.GBl. S. 369, ber. S. 426) Sa BremR 223–a–21 geändert durch Art. 1 VO zur Änderung schulrechtlicher Verordnungen vom 19.12.2014 (Brem.GBl. 2015 S. 5): Verordnung über die Organisation des Bildungsgangs der Grundschule (Grundschulverordnung) im Land Bremen.

Hamburg

HmbSG, vom 16.04.1997 (HmbGVBl. S. 97), zuletzt geändert am 15.09.2016 (HmbGVBl. S. 441): Hamburgisches Schulgesetz.

Hessen

Schulgesetz - HSchG, vom 14.06.2005 (GVBl. I S. 441), zuletzt geändert durch Artikel 2 des Gesetzes vom 24.03.2015 (GVBl. S. 118): Hessisches Schulgesetz (GVBl. II 72-123).

Mecklenburg-Vorpommern

Schulgesetz - SchulG M-V, vom 10.09.2010, letzte berücksichtigte Änderung: § 128a angepasst durch Bekanntmachung vom 27.06.2016 (GVOBl. M-V S. 522): Schulgesetz für das Land Mecklenburg-Vorpommern (GVOBl. M-V 2010, S. 462).

Niedersachsen

NSchG, vom 03.03.1998 (Nds. GVBl. S. 137), zuletzt geändert durch Artikel 3 des Gesetzes vom 26.10.2016 (Nds. GVBl. S. 226): Niedersächsisches Schulgesetz.

Nordrhein-Westfalen

Schulgesetz NRW - SchulG, vom 15.02.2005 (GV. NRW. S. 102), zuletzt geändert durch Gesetz vom 25.06.2015 (GV. NRW. S. 499): Schulgesetz für das Land Nordrhein-Westfalen.

Ausbildungsordnung Grundschule – AO-GS, vom 23.03.2005: Verordnung über den Bildungsgang in der Grundschule, Nordrhein-Westfalen.

Rheinland-Pfalz

SchulG, vom 30.03.2004, letzte berücksichtigte Änderung: § 101 geändert durch Artikel 10 des Gesetzes vom 16.02.2016 (GVBl. S. 37): Schulgesetz Rheinland-Pfalz.

GSchO, vom 10.10.2008 in der Fassung vom 09.12.2013: Schulordnung für die öffentlichen Grundschulen in Rheinland-Pfalz.

Saarland

Schulordnungsgesetz SchoG, vom 21.08.1996 (Amtsbl. S. 846, ber. 1997 S. 147), zuletzt geändert durch das Gesetz vom 20.01.2016 (Amtsbl. I S. 120): Gesetz Nr. 812 zur Ordnung des Schulwesens im Saarland.

Schulpflichtgesetz, vom 21.08.1996 (Amtsbl. S. 864, ber. 1997 S. 147), zuletzt geändert durch das Gesetz vom 17.06.2015 (Amtsbl. I S. 446): Gesetz Nr. 826 über die Schulpflicht im Saarland.

Sachsen

SchulG, vom 16.07.2004 (SächsGVBl. S. 298), zuletzt geändert durch das Gesetz vom 01.02.2017 (SächsGVBl. S. 18): Schulgesetz für den Freistaat Sachsen.

SOGS, vom 03.08.2004 (SächsGVBl. S. 312), zuletzt geändert durch Artikel 1 der Verordnung vom 15.02.2017 (SächsGVBl. S. 87): Schulordnung Grundschulen, Sachsen.

Sachsen-Anhalt

SchulG LSA, vom 22.02.2013, letzte berücksichtigte Änderung: § 30 geändert durch Artikel 8 des Gesetzes vom 25.02.2016 (GVBl. LSA S. 89, 94): Schulgesetz des Landes Sachsen-Anhalt.

Schleswig-Holstein

GrundSchulV SH 2007, vom 22.06.2007: Landesverordnung über Grundschulen, Schleswig-Holstein.

Schulgesetz – SchulG, vom 24.01.2007: Schleswig-Holsteinisches Schulgesetz.
Fundstelle: GVOBl. 2007, 39, ber. S. 276.

Thüringen

ThürSchulG, vom 30.04.2003, zuletzt geändert durch das Gesetz vom 31.01.2013 (GVBl. S. 22, 23): Thüringer Schulgesetz.

Thüringer Schulordnung - ThürSchulO, vom 20.01.1994, letzte berücksichtigte Änderung: mehrfach geändert durch Verordnung vom 07.07.2011 (GVBl. S. 208): Thüringer Schulordnung für die Grundschule, die Regelschule, die Gemeinschaftsschule, das Gymnasium und die Gesamtschule.

A.2., „Reformen auf einen Blick“ (SenBJW)

Einige wichtige Reformen der letzten Jahre auf einen Blick ...

- ▶ Neues Schulgesetz - erstes Bundesland nach Pisa
- ▶ Umgestaltung der Kitas zu Bildungseinrichtungen
- ▶ Übertragung von Kitas an freie Träger
- ▶ Deutschtest und Deutschpflichtkurs vor der Einschulung
- ▶ Flexible Schulanfangsphase
- ▶ Senkung des Einschulungsalters
- ▶ Keine Rückstellungen mehr
- ▶ Verlagerung der Horte an die Schulen
- ▶ Alle Grundschulen sind verlässliche Halbtagschulen
- ▶ Zusätzliche Stunden für Leseförderung
- ▶ Unterstützung für Lehrkräfte in migrantenstarken Jahrgängen
- ▶ Alle Gesamtschulen zu Ganztagschulen
- ▶ Arbeitsprogramm Hauptschule
- ▶ Mittlerer Schulabschluss für alle
- ▶ Verkürzung der Zeit zum Abitur auf 12 Jahre
- ▶ Zentralabitur
- ▶ Regelmäßige Vergleichsarbeiten in allen Schulformen
- ▶ Überprüfung der Schulqualität
- ▶ Konzept zur Förderung Hochbegabter
- ▶ Stärkung der Mitbestimmung von Eltern und Schülerinnen und Schülern
- ▶ Schulen erhalten Vertretungsreserve (2007)
- ▶ Konzept für Kinderschutz beschlossen (2007)
- ▶ Drei Projektgruppen zur Optimierung der Verwaltungsarbeit (2006)



Neues modernes Schulgesetz (2004) – erstes Bundesland nach PISA

Der PISA-Befund für Deutschland war verheerend. Berlin hat sofort reagiert – mit einem neuen, modernen Schulgesetz, das tief greifende Reformen einleitete. Zu viel Reform, zu wenig Reform – beide Vorwürfe wurden hernach erhoben. Beides ist falsch:

1. Die Berliner Bildungsreform musste umfassend sein, denn ein Bildungssystem ist ein komplexes Bedingungsgefüge – hier können nicht einzelne Teile isoliert voneinander reformiert werden.
2. Die Reformen hätten gewiss auch für zwei Legislaturperioden gereicht – aber im Interesse der nachwachsenden Generation war keine Zeit zu verlieren.
3. Mehr und schnellere Reformen hätten alle Akteure von Schule überfordert. Eine Reform folgt sinnvollen Schrittfolgen, die ineinander greifen wie Zahnräder. Eine Reform muss sich prozesshaft entfalten können – dafür braucht sie Zeit und Ruhe.
4. Wer aus einer Bildungsreform der letzten 4 Jahre schnelle sichtbare Erfolge in Vergleichsstudien erwartet, hat keine Ahnung von Bildungspolitik. (Die bei PISA getesteten 15-Jährigen sind Mitte der 90er Jahre in die Schule eingetreten.)

Umgestaltung der Kitas zu Bildungseinrichtungen

Kitazeit ist in Berlin Bildungszeit. Denn nie wieder sind Kinder so lernbereit und lernfähig wie in den frühen Kindheitsjahren. Nichts interessiert die Kleinen so sehr wie zu wissen, warum die Dinge so sind, wie sie sind. Das für alle Berliner Kitas verbindliche Kita-Bildungsprogramm greift diese kindliche Neugierde auf. Es wurde erarbeitet von der Arbeitsgruppe Berliner Fachwissenschaftler und Praxisberater unter Leitung der INA – FU Berlin, Leitung: Dr. Ch. Preissing). Grundlegende Aussagen zum Bildungsbegriff, zu den Zielen und Aufgaben der Erzieherinnen und Erzieher werden mit der ausführlichen Beschreibung von sieben Bildungsbereichen verknüpft. Die konzeptionelle Rahmenvorgabe und konkretes Arbeitsmaterial für

die Praxis unterstützt die pädagogischen Fachkräfte bei Orientierung an der Lebenssituation der Kinder. Alle Kitas erhielten die Bildungsprogramme. Um die Eltern mitzunehmen wurden Kurzinfos in fünf Sprachen verteilt. Das Bildungsprogramm stößt auf große Zustimmung in Berlin, bundesweit und auch im Ausland. Das Saarland und Hamburg haben es übernommen, die Bernard van Leer Foundation lässt es ins Englische übersetzen. Alle Berliner Erzieherinnen und Erzieher wurden über ein Stufensystem zu diesem Programm fortgebildet. Die Überprüfung der Erfolge der Anwendung erfolgt in mehreren Schritten bis 2009 in jeder Kita.

Übertragung von Kitas an freie Träger - Ziel: mehr Vielfalt, Qualität

Berlin hatte sich 2003 vorgenommen, das Verhältnis von Kindertageseinrichtungen von zwei Dritteln in kommunaler zu einem Drittel in freier Trägerschaft umzukehren. Denn mehr Konkurrenz bringt mehr Qualität, und Eltern sollen aus einer Vielfalt von Kita-Profilen wählen können – von der sportbetonten über die Kneipp-Kita bis zur bilingualen Kita.

Das ist nun geschafft:

Seit 2004 wurden 273 Kitas mit 30.000 Plätzen in freie Trägerschaft übertragen. Insgesamt sind damit dann zwei Drittel aller Kitaplätze in freier Trägerschaft. Alle verbleibenden kommunalen Kitas wurden in 5 Eigenbetrieben, in denen sich jeweils 2-3 Bezirke zusammenschließen, zusammengefasst. Die Finanzierung wurde über Kostenblätter gesichert und damit eine einheitliche Qualität ermöglicht. Besser früh investieren als spät reparieren ist Berlins Devise. Im Kitabereich ist Berlin Spitze – in Vielfalt, in Qualität und in Quantität. In Berlin sind Beruf und Familie vereinbar.

Deutschtest und Deutschpflichtkurs vor der Einschulung (ab 2004)

Seit 2004 gibt es den für alle verpflichtenden Sprachtest Deutsch-Plus. Der Test erfolgt im Zusammenhang mit der Schulanmeldung – bei Kitakindern durch die Kitaerzieher, bei Nicht-Kita-Kindern in der Grundschule. Alle Kinder, die Sprachdefizite aufweisen, sind zur Teilnahme an der daran anschließenden Sprachförderung verpflichtet. Kita-Kinder werden in der Kita gefördert, denn Sprachförderung ist ein Schwerpunkt im Kita-Bildungsprogramm. Zusammen mit dem grünen Sprachlernordner zur Sprachförderung, dem Sprachkoffer und dem Sprachlernortagebuch für jedes Kind stehen den Erzieherinnen entsprechende Materialien zur Verfügung.

Kinder mit Sprachdefiziten, die keine Kita besuchen, müssen am halbjährigen Sprachförderkurs in einer Grundschule teilnehmen, den Lehrerinnen mit der Ausbildung „Deutsch als Zweitsprache“ und Erzieherinnen verantworten: 2005 waren es 2 Stunden, seit 2006 sind es 3 Stunden Sprachförderung täglich (insgesamt 330 Stunden). Ziel dieser Pflicht: Jedes Kind soll vom ersten Schultag an dem Unterricht folgen können, damit keine Lücken entstehen.

Flexible Schulanfangsphase

Nicht das Kind muss „schulreif“ sein, sondern Schule muss sich so wandeln, dass sie jedes Kind vom ersten Tag an individuell fördern kann. Ein hohes Ziel – wie kann man sich dem nähern?

Das 1. und 2. Schuljahr werden zur flexiblen Schulanfangsphase zusammengefasst, in der die Kinder – je nach Voraussetzungen beim Schuleintritt und folgendem individuellem Lernfortschritt – ein, zwei oder drei Jahre bleiben.

Die flexible Schulanfangsphase gibt es auch in anderen Bundesländern; Berlin führt sie schrittweise seit 2004/05 ein. Seit 2007/2008 ist sie für alle Grundschulen Pflicht. Genug Zeit also für Lehrkräfte, sich in den speziellen Fortbildungen, die das LISUM dazu anbietet, fortzubilden und in Schulen, die das z. T. als Schulversuch bereits seit vielen Jahren erfolgreich praktizieren, zu hospitieren.

Vorteile: Das individuelle Lerntempo entscheidet, Schnelle werden nicht ausgebremst, bei langsamer Lernenden entstehen keine Lücken, die sie dann durch die Schullaufbahn schleppen. Kein Kind wechselt allein die vertraute soziale Gruppe.

Das soziale Lernen wird intensiviert – jeder kann etwas anderes gut und hilft dem anderen. Auch zunächst lernschwächere Kinder können in bestimmten Bereichen Jüngeren helfen. Die Klassen sind ruhiger, denn Erstklässler werden durch die erfahrenen Kinder in die Abläufe eingeführt.

Auch Skeptiker haben diese positiven Erfahrungen inzwischen überzeugt.

Bereits jetzt gibt es in einem altershomogenen Erstklässlerjahrgang bereits Entwicklungsunterschiede von bis zu drei Jahren. Die flexible Schulanfangsphase ermöglicht einen besseren Umgang damit.

Senkung des Einschulungsalters von 6,7 auf 6,2 Jahre (2005)

Im europäischen Vergleich wurden deutsche Kinder spät eingeschult und damit wurde wertvolle

Bildungszeit im schulischen Kontext verschenkt. Mit der Neustruktur des Schulbeginns (flexible Schulanfangsphase) und anderer methodisch-didaktischer Arbeitsansätze ist Schule fit auch für jüngere Kinder. Seit dem Schuljahr 2005/06 werden alle Kinder eingeschult, die im laufenden Kalenderjahr 6 Jahre alt werden. Somit kommen Berliner Kinder nun mit durchschnittlich 6,2 Jahren in die Schule.

Keine Rückstellungen mehr

Grundsätzlich werden alle Kinder in dem Jahr schulpflichtig, in dem sie das sechste Lebensjahr vollenden. Seit dem Schuljahr 2005/2006 gibt es daher in der Regel keine Zurückstellungen mehr. Die Grundschule hat sich in der Phase Schulanfangs mit der flexiblen Schulanfangsphase, in der die Jahrgangsstufen 1/2 eine pädagogische Einheit bilden, kindgerecht verändert. In der Schulanfangsphase wird der Fokus stärker als je zuvor auf die unterschiedlichen Voraussetzungen und Bedürfnisse der Kinder gelegt. Die Erfahrung zeigte, dass Entwicklungsverzögerungen, die zu Rückstellungen führten, häufig soziale Gründe hatten. Rückstellungen führten häufig dazu, dass sich diese Defizite um ein weiteres Jahr verfestigten, weil das Kind in diesem Umfeld verblieb. Die flexible Schulanfangsphase nimmt das Kind fördernd in seinen Lernvoraussetzungen und -möglichkeiten an und ermöglicht ihm - angepasst an seine Lernvoraussetzungen und sein individuelles Lerntempo - anschlussfähige Lernschritte. Kinder, die mehr Zeit benötigen, können drei Jahre in der Schulanfangsphase verweilen.

Sollte der Entwicklungsstand eines Kindes allerdings so starke Verzögerungen aufweisen, dass eine bessere Förderung in einer Kindertagesstätte zu erwarten ist, kann auf Antrag der Eltern die regional zuständige Schulaufsicht - auf Grundlage einer Begründung der Eltern, einer Stellungnahme der besuchten Kindertagesstätte sowie eines Gutachtens des zuständigen Schularztes oder des Schulpsychologischen Dienstes - eine Zurückstellung von der Schulbesuchspflicht für ein Jahr genehmigen, sofern dann eine entsprechende Förderung in einer Kita erfolgt. Eine Rückstellung nach Beginn des Schulbesuchs ist ausgeschlossen.

Verlagerung der Horte an die Schulen - alle Grundschulen sind offene oder gebundene Ganztagschulen (2005)

Ganztagschule – Bildung, Erziehung und Betreuung aus einem Guss; Unterricht und Freizeit kindgemäß rhythmisiert und aufeinander bezogen. Berlin verlagerte alle Horte in die Zuständigkeit – und wo irgend möglich auch räumlich – in die Grundschulen. 162 Mio € wurden und werden dafür in Bau und Ausstattung investiert. Ein schwieriger Umstellungsprozess über Jahre, dessen Erfolge und Chancen bereits jetzt offensichtlich sind. Grundschule ist in Berlin flächendeckend Ganztagschule – von 6 – 18 Uhr. Von solcher Qualität und Quantität in der Vereinbarkeit von Beruf und Familie träumen Eltern anderswo.

Alle Grundschulen sind verlässliche Halbtagschulen (2005) – 7.30-13.30 Uhr kostenlos

Montags erst zur 3. Stunde Unterricht oder sommers hitzefrei – nach der 4. Stunde Schluss - aber die Eltern arbeiten: Wohin mit dem Kind?

In Berlin kein Problem mehr. Seit dem Schuljahr 2005/06 bietet jede Grundschule verlässliche Betreuung auch außerhalb der Unterrichtszeiten in der Zeit von 7.30 – 13.30 Uhr an – kostenlos für alle und ohne gesonderte Anmeldung durch die Eltern. Nicht jeder will die Ganztagschule nutzen und trotzdem.

Zusätzliche Stunden für Leseförderung, Fremdsprache ab Klasse 3 (2003) und Naturwissenschaften in der Grundschule (2005)

Grundschule stärken – denn dort werden, neben der Kita – die entscheidenden Weichen für den Bildungserfolg gestellt. Die flächendeckende Ganztagsgrundschule ist ein tragendes Element dabei. Berlin hat darüber hinaus auf diagnostizierte Defizite reagiert:

Eine Stunde mehr Leseförderung im 2. Jahrgang für alle, denn die Ergebnisse der Orientierungstests Leseverständnis konnten vielerorts nicht zufrieden stellen. Jüngere Kinder lernen Sprachen schneller, deshalb wurde der Start der ersten Fremdsprache auf die 3. Klasse vorgezogen – zur Wahl steht Englisch und Französisch.

Vergleichstests wiesen auf Mängel in der naturwissenschaftlichen Bildung. Berlin führte daher das Fach Naturwissenschaften in der Grundschule ein.

Unterstützung für Lehrkräfte in migrantenstarken Jahrgängen

Die Quote von Kindern nicht-deutscher Herkunftssprache in den Einschulungsjahrgängen steigt steil an. Häufig kommen Kinder aus bildungsfernen Elternhäusern und haben neben massiven Sprachproblemen auch Defizite in ganz basalen Fähigkeiten wie dem Umgang mit einem Stift oder gar einer Schere. Um Lehrkräfte hier zu unterstützen, wurde die sonst übliche Zumessung in Grundschulen (1 Lehrkraft für 24 Kinder) bei Jahrgängen mit mehr als 40 % Kindern nicht-deutscher Herkunftssprache abgesenkt auf 1

Lehrkraft für 20 Kinder. Im Schuljahr 2005/06 zunächst für die 1. Klassen, im Schuljahr 2006/07 wurde das auch auf die 2. Klassen ausgedehnt. Die Schulen können diese zusätzlichen Lehrerstunden für eine Frequenzabsenkung, für Doppelsteckung oder Teilungsstunden verwenden – auf jeden Fall kommt sie Kindern und Lehrkräften zugute. Zusätzlich gibt es in solchen Jahrgängen eine weitere zusätzliche Stunde Deutschunterricht.

Alle Gesamtschulen zu Ganztagschulen ausgebaut (2006)

Alle gemeinsam zu bestmöglichem Bildungserfolg – das ist das Kurzkonzept der Gesamtschule. Für ein Mehr an Zeit sind in Berlin seit dem 1. August 2006 nicht nur alle 400 Grund-, sondern auch alle 49 Gesamtschulen Ganztagschulen. Rhythmisierter Unterricht, Entspannungszeiten und Zeiten für Hausaufgabenbetreuung und vielfältige Anregung zu anspruchsvoll gestalteter Freizeit mit Arbeitsgemeinschaften und Clubs, die auch von außerschulischen Trägern angeboten werden machen die Schule zum Lern und Lebensort – für mehr Spaß am Lernen und ein Mehr an Bildungserfolg.

Arbeitsprogramm Hauptschule (2005)

In Berlin schaffen 10 % mehr Jugendliche das Abitur als in Bayern, in Berlin besuchen 10% eines Jahrgangs die Hauptschule – in Bayern mehr als das Doppelte. Hauptschule ist eine Schule mit Chancen – sie ist weitaus besser als andere Schulformen mit Personal ausgestattet: 1 Lehrkraft für ca. 9 Schüler. Das ist gut und richtig so – denn hier kommen Kinder und Jugendliche zusammen, die familiär, sozial und bildungsbiographisch nicht die besten Startchancen hatten. Um sie noch besser zu Unterstützung, wurde 2005 das Arbeitsprogramm Hauptschule aufgelegt: mehr Praxisorientierung, mehr Förderung und Forderung in Bildung und Erziehung, Sozialarbeiter an jede Hauptschule, Zusammenwachsen mit Realschulen. So geht es vorwärts in einer schwierigen Schulform, die – das hat das Beispiel Rütli-Schule gezeigt – mit den richtigen Personen und der richtigen Motivation einen guten und selbstbewussten Weg gehen kann.

Mittlerer Schulabschluss für alle (ab 2006) - ersetzt den Realschulabschluss

Mehr Transparenz: Der Mittlere Schulabschluss (MSA) wird – für alle gleich – in allen Schularten der weiterführenden Schule abgelegt. Er ersetzt den Realschulabschluss.

Vorteile: Schulen, Schülerinnen und Schüler, Eltern und Ausbildungsbetriebe wissen, woran sie sind, denn die Abschlüsse sind nun vergleichbar.

Die neu eingeführte Präsentationsprüfung birgt zusätzliche Chancen für die Jugendlichen, erworbene Kompetenzen einzubringen.

Die Erfahrungen mit dem ersten MSA waren positiv.

Eine Prüfung, die alle bestehen, ist keine Prüfung; auch im MSA zeigen sich Berlins Ansprüche an Qualität: 78 % haben den MSA bestanden. Besonders bemerkenswert: Gut vorbereitet haben 3 % mehr Hauptschüler als bislang diesen Abschluss geschafft. Die Schulen erhielten genaue Rückmeldung z. B. über Defizite in mathematischen Kompetenzen – hier werden jetzt gezielt Fortbildungen angeboten.

Verkürzung der Zeit zum Abitur auf 12 Jahre (an Gesamtschulen weiter 13 Jahre möglich)

Unsere Kinder kommen zu spät aus der Schule, zu spät in den Job – im europäischen Vergleich. Neben der Stellschraube „frühere Einschulung“ wurde auch am Ende der Schulzeit optimiert: Die Siebtklässler des Schuljahres 2006/07 werden an Gymnasien das Abitur bereits nach 12 Jahren ablegen. Die Stunden des 11. Jahrgangs wurden auf die Jahrgänge 5-10 als zusätzliche Stunden verlagert – für alle. In Berlin kommen also auch Kinder, die nicht das Abitur anstreben, in den Genuss dieses Mehr an Unterricht.

Wer nicht so schnell ist oder später durchstartet, behält seine Chancen: An Gesamtschulen ist das Abitur weiterhin nach 13 Jahren möglich.

Zentralabitur (ab 2007)

Viele Länder haben es schon – ab Frühjahr 2007 nun auch Berlin: Das Zentralabitur. In den Fächern Deutsch, Mathematik und den Fremdsprachen wird mit zentraler Aufgabenstellung geprüft. Das schafft Vergleichbarkeit und Gerechtigkeit. Und es entlastet die Lehrkräfte, die selbst keine Abituraufgaben mehr entwerfen müssen. Berlin ist mit seiner Abiturquote von 39,6 % eines Jahrgangs Spitze. Die Quote stieg stetig an. Allein um mehr als 3 % in den letzten drei Jahren – ein Zeichen für bessere Qualität und mehr Leistungsbereitschaft.

Regelmäßige Vergleichsarbeiten in allen Schulformen

Messen allein hilft nicht – natürlich, aber eine valide Datenlage gibt detaillierte Rückmeldung, die der erste Schritt ist, um Verbesserungen in Gang zu setzen. Berlin hat ein System von Vergleichsarbeiten und

Lernstandstests installiert, das die Leistungsentwicklung jedes einzelnen Kindes, der Klasse, der Schule regelmäßig überprüft und dieses an die Schule zurückmeldet. Dieses Wissen muss Konsequenzen haben – Diskussion in schulischen Gremien, problemangemessene Maßnahmen zur Veränderung, erneute Überprüfung.

Transparenz sowie interne und externe Überprüfung der Schulqualität

Was gab es vor 5 Jahren an Maßnahmen zur Qualitätsüberprüfung und –entwicklung an Berliner Schulen? Fehlanzeige! In einem Land, das für Qualitätsarbeit bekannt ist, kann auch Schule nicht länger ausgenommen werden.

Neben der gewachsenen Bereitschaft, sich internationalen Vergleichsstudien zu stellen und dem System von Orientierungs- und Vergleichsarbeiten sind viele wirksame Qualitätsinstrumente installiert worden, die einem Paradigmenwechsel nahe kommen: Schulen schärfen mit ihrem -> Schulprogramm ihr Profil, setzen sich Ziele und überprüfen, ob sie sie erreichen. Seit 2006 ist eine -> Schulinspektion etabliert. Teams aus Lehrkräften, Schulleitern, Eltern, der IHK und den Unis überprüfen die Qualität einer jeden Schule im 5-Jahresturnus und geben Anregung zu Verbesserungen. Das neu eingeführte -> Schulporträt auf unseren Seiten im Internet veröffentlicht standardisierte Daten über die einzelne Schule und schafft so mehr Transparenz auch für Eltern bei der Suche nach einer geeigneten Schule für ihr Kind. Dort werden auch „harte“ Daten zum Unterrichtsausfall veröffentlicht, der in Berlin durch gemeinsame Anstrengung auf dem niedrigsten Stand seit 1999 ist – aber natürlich trotzdem noch zu hoch.

Konzept zur vielfältigen Förderung Hochbegabter (2004)

Berlin hat ein ausgefeiltes Konzept zur Förderung besonders begabter und hoch begabter Kinder und Jugendlichen. Dabei wird fein abgestimmt: So viel Separation wie nötig, so viel Integration wie möglich. Berlin bietet Superschnellläuferklassen, die in 11 Jahren zum Abitur führen, Durchlässigkeit, Klassen zu überspringen bzw. einzelne Fächer in nächst höheren Klassenstufen zu besuchen. Die vier Hochbegabten-Netzwerke Nord, Süd, West, und Ost ermöglichen kurze Weg und Zugang zu Projekten und zusätzlichem Nachmittagsunterricht in Begabtengruppen. Jährlich finden auf der Insel Scharfenberg erfolgreiche Sommercamps für Hochbegabte statt, von denen Kinder, Lehrkräfte, Wissenschaftler und Eltern begeistert sind.

Stärkung der Mitbestimmung von Eltern und Schülerinnen und Schülern sowie externer Kompetenz (Schulkonferenz)

Berlin hat mit seinem neuen Schulgesetz die Mitbestimmung von Eltern und Schülerinnen und Schülern weiter gestärkt und zeigt damit: Sie sind neben den Lehrkräften mitentscheidende und mitverantwortliche Akteure von Schule und können viel bewegen. Höchstes Entscheidungsgremium der Schule ist die Schulkonferenz, besetzt zu jeweils einem Drittel mit Lehrkräften, Eltern und Schüler/innen sowie einer externen, nicht zur Schule gehörenden kompetenten Person.

Schulen erhalten Vertretungsreserve (2007)

Schulen erhalten mehr Mittel und Verantwortung mit neuer Vertretungsreserve. Ab dem Schuljahr 2007/2008 kann jede Schule auf freiwilliger Basis in eigener Verantwortung für kurzfristige Vertretung sorgen. Schulen, die sich für diese dezentrale Vertretungssteuerung entscheiden, erhalten einen Etat in Höhe von drei Prozent ihres Unterrichtsbedarfs – entsprechend etwa 1.350 Euro pro Vollzeitstelle. Aus diesen Mitteln können sie befristete Einstellungen finanzieren, die bis zu drei Monate dauern können. Wenn Lehrerinnen und Lehrer länger als drei Monate ausfallen und die Ausstattung der Schule dadurch unter 100 Prozent fällt, finanziert das Land eine befristete Einstellung.

Konzept für Kinderschutz beschlossen (2007)

Jeder Fall von Verwahrlosung und Vernachlässigung von Kindern ist einer zu viel. Die schlimmen Fälle in der Vergangenheit zeigen, dass die Kinder und Jugendlichen in Hilfesystemen betreut wurden, die aber nicht immer wie erforderlich funktionierten. An irgendeiner Stelle wurden Informationen nicht weitergereicht, andere Behörden nicht einbezogen oder es wurde zu einem späteren Zeitpunkt nicht noch einmal nachgeschaut. Klar ist: Es darf keine Sackgassen geben und keiner darf sich die Hände in Unschuld waschen.

Der Senat verfolgt das Ziel, den Kinderschutz weiter zu verbessern, risikohafte Entwicklungen frühzeitiger zu erkennen und schneller zu handeln. Er hat deshalb ein Konzept für ein Netzwerk Kinderschutz beschlossen. Das erfordert eine verbesserte Zusammenarbeit insbesondere zwischen Kinder- und Jugendgesundheitsdiensten, Kinderärzten, Jugendämtern, Kindertageseinrichtungen, Schulen, Gerichten und Polizei. Dafür werden Haushaltsmittel in Höhe von mehr als 1,1 Mio. Euro eingesetzt. Mit diesem Netzwerk nimmt Berlin eine Vorreiterrolle im Kampf gegen Verwahrlosung und Vernachlässigung ein.

Außerdem wurde eine zentrale „Hotline-Kinderschutz“ unter der Telefonnummer 610066 freigeschaltet. Die Hotline ist rund um die Uhr von Kinderschutzberatern besetzt. Bei jeder eingehenden Meldung nehmen die Beraterinnen und Berater eine Risikoeinschätzung vor und leiten gegebenenfalls Kinderschutzmaßnahmen ein.

© Senatsverwaltung für Bildung, Jugend und Wissenschaft Berlin

Kontakt

Senatsverwaltung für Bildung, Jugend und Wissenschaft
Bernhard-Weiß-Str. 6
D-10178 Berlin-Mitte

 [Stadtplan](#)

Telefon 030 90227 5050

A.3. Übersicht zur Berichterstattung regionaler und überregionaler Medien

JüL_Brief. Online verfügbar unter

http://www.tagesspiegel.de/downloads/4614150/2/pdf_brandbrief.pdf, zuletzt geprüft am 12.12.2016.

Jeder sechste Erstklässler wird verspätet eingeschult (2014). In: Berliner Zeitung, 06.11.2014.

Online verfügbar unter <http://www.berliner-zeitung.de/berlin/jeder-sechste-erstklaessler-wird-verspaetet-ingeschult-479288>, zuletzt geprüft am 27.03.2017.

Bürgerinitiative überreicht Unterschriften für spätere Einschulungen (2015). In: Berliner

Zeitung, 27.02.2015. Online verfügbar unter <http://www.berliner-zeitung.de/berlin/buergerinitiative-ueberreicht-unterschriften-fuer-spaetere-einschulungen-2580238>, zuletzt geprüft am 27.03.2017.

Alberti, Stefan (2015): Bildungspolitik in Berlin: Senat verschiebt Einschulung. In: taz.de,

06.10.2015. Online verfügbar unter <https://www.taz.de/Archiv-Suche/%215235249&s=fr%C3%BChe+einschulung+berlin/>, zuletzt geprüft am 27.03.2017.

Anders, Florentine; Köhler, Regina (2009): Gemeinsames Lernen: Berliner Grundschulen

rebellieren gegen Reform. In: Berliner Morgenpost, 22.04.2009. Online verfügbar unter <http://www.morgenpost.de/berlin/article104070894/Berliner-Grundschulen-rebellieren-gegen-Reform.html>, zuletzt geprüft am 28.03.2017.

Anders, Florentine; Köhler, Regina (2010): Jahrgangsgemischter Unterricht: Die Schulen sind

mit der Reform überfordert. In: Berliner Morgenpost, 14.09.2010. Online verfügbar unter <http://www.morgenpost.de/berlin-aktuell/article104653324/Die-Schulen-sind-mit-der-Reform-ueberfordert.html>, zuletzt geprüft am 28.03.2017.

Bittner, Maria (2013): Erfahrungsbericht: Jahrgangsübergreifendes Lernen wirkt wie ein

Durchlauferhitzer. In: Berliner Morgenpost, 10.08.2013. Online verfügbar unter <http://www.morgenpost.de/berlin-aktuell/article118893199/Jahrgangsuebergreifendes-Lernen-wirkt-wie-ein-Durchlauferhitzer.html>, zuletzt geprüft am 28.03.2017.

Dückers, Tanja (2014): Frühe Einschulung: Keine Fünfjährigen in die ersten Klassen. In: Zeit

online, 16.09.2014. Online verfügbar unter <http://www.zeit.de/gesellschaft/schule/2014-09/frueheinschulung-bildungssystem/komplettansicht>, zuletzt geprüft am 27.03.2017.

Hecht, Patricia (2010): Einschulung: Debatte um frühe Schulpflicht reißt nicht ab. In: Der

Tagesspiegel, 28.09.2010. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/einschulung-debatte-um-fruehe-schulpflicht-reisst-nicht-ab/1943718.html>, zuletzt geprüft am 27.03.2017.

Hecht, Patricia (2011): Jahrgangsmischung: Gemeinsam oder doch getrennt. In: Der

Tagesspiegel, 27.09.2011. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/jahrgangsmischung-gemeinsam-oder-doch-getrennt/4665718.html>, zuletzt geprüft am 28.03.2017.

Jiménez, Fanny (2015): Warum es das richtige Einschulungsalter nicht gibt. In: WELT.de,

01.01.2015. Online verfügbar unter <https://www.welt.de/wissenschaft/article149074052/Warum-es-das-richtige-Einschulungsalter-nicht-gibt.html>, zuletzt geprüft am 27.03.2017.

Kerstan, Thomas; Spiewak, Martin (2008): Schulsystem: »Schule ist die große Gleichmacherin«.

Was muss getan werden, um Deutschlands Bildungssystem gerechter zu machen? Ein

- Gespräch mit dem Erziehungswissenschaftler Jürgen Baumert. In: Die Zeit, 19.09.2008. Online verfügbar unter <http://www.zeit.de/2008/39/C-Interview-Baumert>, zuletzt geprüft am 05.03.2017.
- Klesmann, Martin (2015): Ende der Früheinschulung: Berliner Kinder werden wieder mit sechs Jahren eingeschult. In: Berliner Zeitung, 06.10.15. Online verfügbar unter <http://www.berliner-zeitung.de/berlin/ende-der-frueheinschulung-berliner-kinder-werden-wieder-mit-sechs-jahren-eingeschult-22592590>, zuletzt geprüft am 27.03.2017.
- Klesmann, Martin (2010): Zuversichtliches für Erstklässler. In: Berliner Zeitung, 28.08.2010. Online verfügbar unter <http://www.berliner-zeitung.de/zuversichtliches-fuer-erstklaessler-15408806>, zuletzt geprüft am 28.03.2017.
- Klesmann, Martin (2011): Zwei Schulleiterinnen streiten über das jahrgangsübergreifende Lernen in der Grundschule. Die eine will es ausweiten - die andere will es loswerden: "Von oben herab verordnet". In: Berliner Zeitung, 17.09.2011. Online verfügbar unter <http://www.berliner-zeitung.de/zwei-schulleiterinnen-streiten-ueber-das-jahrgangsuuebergreifende-lernen-in-der-grundschule--die-eine-will-es-ausweiten---die-andere-will-es-loswerden--von-oben-herab-verordnet--15151678>, zuletzt geprüft am 28.03.2017.
- Klesmann, Martin (2012): Berlin-Mitte: Lehrer stellen Schulsenatorin ein Ultimatum. In: Berliner Zeitung, 20.03.2012. Online verfügbar unter <http://www.berliner-zeitung.de/berlin/berlin-mitte-lehrer-stellen-schulsenatorin-ein-ultimatum-10685462>, zuletzt geprüft am 28.03.2017.
- Klesmann, Martin (2012): Jahrgangsübergreifendes Lernen: Jede dritte Grundschule gibt Altersmischung auf. In: Berliner Zeitung, 20.12.2012. Online verfügbar unter <http://www.berliner-zeitung.de/berlin/jahrgangsuuebergreifendes-lernen-jede-dritte-grundschule-gibt-altersmischung-auf-5825022>, zuletzt geprüft am 28.03.2017.
- Klesmann, Martin (2013): Bildungspolitik: Senat hadert mit der frühen Einschulung. In: Berliner Zeitung, 26.01.2013. Online verfügbar unter <http://www.berliner-zeitung.de/berlin/bildungspolitik-senat-hadert-mit-der-fruehen-einschulung-5818690>, zuletzt geprüft am 27.03.2017.
- Klesmann, Martin (2013): GEW für Auswertung der frühen Einschulung. In: Berliner Zeitung, 28.01.2013. Online verfügbar unter <http://www.berliner-zeitung.de/gew-fuer-auswertung-der-fruehen-einschulung-6681830>, zuletzt geprüft am 27.03.2017.
- Klesmann, Martin (2013): Bildungspolitik: Senat lässt Früheinschulung überprüfen. In: Berliner Zeitung, 24.05.2013. Online verfügbar unter <http://www.berliner-zeitung.de/berlin/bildungspolitik-senat-laesst-frueheinschulung-ueberpruefen-5821808>, zuletzt geprüft am 27.03.2017.
- Klesmann, Martin (2013): Einschulung: Reif für die Schule. In: Berliner Zeitung, 09.08.2013. Online verfügbar unter <http://www.berliner-zeitung.de/einschulung-reif-fuer-die-schule-3810664>, zuletzt geprüft am 27.03.2017.
- Klesmann, Martin (2014): Leistungen in der Schule: Früh eingeschulte Kinder wiederholen öfter. In: Berliner Zeitung, 07.07.2014. Online verfügbar unter <http://www.berliner-zeitung.de/berlin/leistungen-in-der-schule-frueh-eingeschulte-kinder-wiederholen-oeffter-3094074>, zuletzt geprüft am 27.03.2017.
- Klesmann, Martin (2014): Schulbeginn: Probleme mit der frühen Einschulung. In: Berliner Zeitung, 17.10.2014. Online verfügbar unter <http://www.berliner-zeitung.de>

- zeitung.de/berlin/schulbeginn-probleme-mit-der-fruehen-einschulung-94194, zuletzt geprüft am 27.03.2017.
- Klesmann, Martin; Heide, Annett (2013): Erstklässler erzählen: S wie Schulkind. In: Berliner Zeitung, 27.09.2013. Online verfügbar unter <http://www.berliner-zeitung.de/berlin/erstklaessler-erzaehlen-s-wie-schulkind-3639616>, zuletzt geprüft am 28.03.2017.
- Klöpper, Anna (2015): Das war die Woche in Berlin I: Leider nur eine Trotzreaktion. In: taz.de, 10.10.2015. Online verfügbar unter <https://www.taz.de/Archiv-Suche/%215237121&s=fr%C3%BChe+einschulung+berlin/>, zuletzt geprüft am 27.03.2017.
- Kneist, Sigrid; Vieth-Entus, Susanne (2009): Bildung: Reform ohne Reife. In: Der Tagesspiegel, 06.07.2009. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/bildungsreform-ohne-reife/1551292.html>, zuletzt geprüft am 27.03.2017.
- Köhler, R.; Nauhaus, C. (2014): Eltern wehren sich gegen Früh-Einschulung ihrer Kinder. In: Berliner Morgenpost, 28.06.2014. Online verfügbar unter <http://www.morgenpost.de/berlin/article129565611/Eltern-wehren-sich-gegen-Frueh-Einschulung-ihrer-Kinder.html>, zuletzt geprüft am 27.03.2017.
- Köhler, Regina (2011): Jahrgangsübergreifendes Lernen: Alle in einem Klassenzimmer - geht das gut? In: Berliner Morgenpost, 30.10.2011. Online verfügbar unter <http://www.morgenpost.de/berlin-aktuell/article105211046/Alle-in-einem-Klassenzimmer-geht-das-gut.html>, zuletzt geprüft am 28.03.2017.
- Köhler, Regina (2012): Berliner Schulen: Lehrer schreiben Brandbrief an Bildungssenatorin. In: Berliner Morgenpost, 18.03.2012. Online verfügbar unter <http://www.morgenpost.de/berlin-aktuell/article106085415/Lehrer-schreiben-Brandbrief-an-Bildungssenatorin.html>, zuletzt geprüft am 28.03.2017.
- Köhler, Regina; Anders, Florentine (2013): Schulanfang: Viele Eltern lassen ihre Kinder ein Jahr später einschulen. In: Berliner Morgenpost, 10.08.2013. Online verfügbar unter <http://www.morgenpost.de/berlin-aktuell/article118883129/Viele-Eltern-lassen-ihre-Kinder-ein-Jahr-spaeter-einschulen.html>, zuletzt geprüft am 27.03.2017.
- Lange, Katrin (2014): Keine Rückstellung: Fünfjähriger wird gegen den Willen seiner Eltern eingeschult. In: Berliner Morgenpost, 22.07.2014. Online verfügbar unter <http://www.morgenpost.de/berlin/article130421458/Fuenfjaehriger-wird-gegen-den-Willen-seiner-Eltern-eingeschult.html>, zuletzt geprüft am 27.03.2017.
- Lehmann, Armin (2012): Schulinspektionsbericht: Berlins Schulen – Mängel überall. In: Der Tagesspiegel, 17.08.2012. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schulinspektionsbericht-berlins-schulen-maengel-ueberall/7016044.html>, zuletzt geprüft am 28.03.2017.
- Lehmann, Anna (2004): Evaluation wird helfen. In: taz. die tageszeitung, 23.09.2004. Online verfügbar unter <https://www.taz.de/Archiv-Suche/%21695951&s=jahrgangs%C3%BCbergreifendes+lernen+berlin&SuchRahmen=Print/>, zuletzt geprüft am 28.03.2017.
- Lehmann, Anna (2007): Jahrgangsübergreifender Unterricht in Grundschulen: Lehrer boykottieren Unterricht. In: taz.de, 21.12.2007. Online verfügbar unter <https://www.taz.de/Archiv-Suche/%215189501&s=jahrgangs%C3%BCbergreifendes+lernen+berlin/>, zuletzt geprüft am 28.03.2017.

- Lehmann, Anna (2008): Schule: Die Angst der Lehrer vor der neuen Lehre. In: taz.de, 28.02.2008. Online verfügbar unter <https://www.taz.de/Archiv-Suche/%215185882&s=jahrgangs%C3%BCbergreifendes+lernen+berlin/>, zuletzt geprüft am 28.03.2017.
- Martens, Daniela (2008): Berliner Schul-Dschungel (2): Auf Antrag schulreif. In: Der Tagesspiegel, 11.09.2008. Online verfügbar unter <http://www.tagesspiegel.de/berlin/berliner-schul-dschungel-2-auf-antrag-schulreif/1321930.html>, zuletzt geprüft am 27.03.2017.
- Memarnia, Susanne (2015): Später in die Schule: Senat zeigt Herz für Kinder. In: taz.de, 09.01.2015. Online verfügbar unter <https://www.taz.de/Archiv-Suche/%215024279&s=fr%C3%BChe+einschulung+berlin/>, zuletzt geprüft am 27.03.2017.
- Miller, Tobias (2002): An 13 Grundschulen lernen Ältere und Jüngere in einer Klasse - mit Erfolg: Die moderne Dorfschule. In: Berliner Zeitung, 06.09.2002. Online verfügbar unter <http://www.berliner-zeitung.de/an-13-grundschulen-lernen-aeltere-und-juengere-in-einer-klasse---mit-erfolg-die-moderne-dorfschule-16621818>, zuletzt geprüft am 28.03.2017.
- Müller-Lissner, Adelheid (2015): Wann ist der beste Zeitpunkt für den Schulanfang?: Abwarten ist nicht die beste Lösung. In: Der Tagesspiegel, 02.03.2015. Online verfügbar unter <http://www.tagesspiegel.de/wissen/wann-ist-der-beste-zeitpunkt-fuer-den-schulanfang-abwarten-ist-nicht-die-beste-loesung/11441496.html>, zuletzt geprüft am 27.03.2017.
- Nolte, Barbara (2012): Erstklässler in Berlin: Sie sind doch erst fünf. In: Der Tagesspiegel, 12.08.2012. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/erstklaessler-in-berlin-sie-sind-doch-erst-fuenf/6992080-all.html>, zuletzt geprüft am 27.03.2017.
- Nolte, Barbara (2012): Einschulung: Schreiend und weinend in die erste Klasse. In: Zeit online, 13.08.2012. Online verfügbar unter <http://www.zeit.de/gesellschaft/schule/2012-08/einschulung-berlin-kinder>, zuletzt geprüft am 27.03.2017.
- Rusch, Hilke (2014): Schule: Jünger! Schneller! Klüger? In: taz.de, 07.07.2014. Online verfügbar unter <https://www.taz.de/Archiv-Suche/%215038178&s=fr%C3%BChe+einschulung+berlin/>, zuletzt geprüft am 27.03.2017.
- Schaar, Katrin (2008): Unterrichtsform: Erste Klassen abgeschafft. In: taz.de, 04.09.2008. Online verfügbar unter <https://www.taz.de/Archiv-Suche/%215176299&s=jahrgangs%C3%BCbergreifendes+lernen+berlin/>, zuletzt geprüft am 28.03.2017.
- Schneider, Marius (2013): Kommentar: Warum die Abkehr der Berliner Schulen von Jül gut ist. In: Berliner Morgenpost, 23.01.2013. Online verfügbar unter <http://www.morgenpost.de/berlin-aktuell/article113070168/Warum-die-Abkehr-der-Berliner-Schulen-von-Juel-gut-ist.html>, zuletzt geprüft am 28.03.2017.
- Vieth-Entus, Susanne (2003): Viel Zustimmung für frühere Einschulung. In: Der Tagesspiegel, 14.11.2003. Online verfügbar unter <http://www.tagesspiegel.de/berlin/viel-zustimmung-fuer-fruehere-einschulung/465240.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2007): Zöllner gegen Pflicht zur Früheinschulung. In: Der Tagesspiegel, 31.01.2007. Online verfügbar unter <http://www.tagesspiegel.de/berlin/zoellner-gegen-pflicht-zur-frueheinschulung/804998.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2008): Hintergrund: Jahrgangsübergreifendes Lernen. In: Der Tagesspiegel, 16.09.2008. Online verfügbar unter

- <http://www.tagesspiegel.de/berlin/hintergrund-jahrgangsuibergreifendes-lernen/1326960.html>, zuletzt geprüft am 28.03.2017.
- Vieth-Entus, Susanne (2008): Interview: "Das ist nichts für schwierige Kinder". In: Der Tagesspiegel, 16.09.2008. Online verfügbar unter <http://www.tagesspiegel.de/berlin/interview-das-ist-nichts-fuer-schwierige-kinder/1325556.html>, zuletzt geprüft am 28.03.2017.
- Vieth-Entus, Susanne (2009): Grundschulreform: Altersgemischte Klassen bleiben umstritten. In: Der Tagesspiegel, 26.05.2009. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/grundschulreform-altersgemischte-klassen-bleiben-umstritten/1521500.html>, zuletzt geprüft am 28.03.2017.
- Vieth-Entus, Susanne (2009): Jahrgangsuibergreifendes Lernen: Wie lernen Ihre Kinder in altersgemischten Gruppen? In: Der Tagesspiegel, 15.09.2009. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/jahrgangsuibergreifendes-lernen-wie-lernen-ihre-kinder-in-altersgemischten-gruppen/1599888.html>, zuletzt geprüft am 28.03.2017.
- Vieth-Entus, Susanne (2010): Grunschulreform: Sehnsucht nach der Vorklasse. In: Der Tagesspiegel, 23.06.2010. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/grundschulreform-sehnsucht-nach-der-vorklasse/1865718.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2010): Jahrgangsuibergreifendes Lernen: Jetzt haben Schulen die Wahl. In: Der Tagesspiegel, 20.11.2010. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/jahrgangsuibergreifendes-lernen-jetzt-haben-schulen-die-wahl/2935984.html>, zuletzt geprüft am 28.03.2017.
- Vieth-Entus, Susanne (2010): Jahrgangsuibergreifendes Lernen: Jetzt haben Schulen die Wahl. In: Der Tagesspiegel, 20.11.2010. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/jahrgangsuibergreifendes-lernen-jetzt-haben-schulen-die-wahl/2935984.html>, zuletzt geprüft am 28.03.2017.
- Vieth-Entus, Susanne (2010): Zöllner will JüL lockern: Jahrgangsmischung ohne Zwang. In: Der Tagesspiegel, 15.12.2010. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/zoellner-will-juel-lockern-jahrgangsmischung-ohne-zwang/3623312.html>, zuletzt geprüft am 28.03.2017.
- Vieth-Entus, Susanne (2011): Brandbrief: Immer mehr Schulen gegen Jahrgangsmischung. In: Der Tagesspiegel, 15.09.2011. Online verfügbar unter <http://www.tagesspiegel.de/berlin/brandbrief-immer-mehr-schulen-gegen-jahrgangsmischung/4613800.html>, zuletzt geprüft am 12.12.2016.
- Vieth-Entus, Susanne (2011): Jahrgangsuibergreifendes Lernen: Fast 70 Schulen dürfen JüL abschaffen. In: Der Tagesspiegel, 18.10.2011. Online verfügbar unter <http://www.tagesspiegel.de/berlin/jahrgangsuibergreifendes-lernen-fast-70-schulen-duerfen-juel-abschaffen/5221646.html>, zuletzt geprüft am 28.03.2017.
- Vieth-Entus, Susanne (2012): Früheinschulung: Senat verteidigt Schule ab fünf. In: Der Tagesspiegel, 26.09.2012. Online verfügbar unter <http://www.tagesspiegel.de/berlin/frueheinschulung-senat-verteidigt-schule-ab-fuenf/7179886.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2012): Rückstellungen in Berlin: Zur Schule mit fünf: Widerstand nimmt deutlich zu. In: Der Tagesspiegel, 04.10.2012. Online verfügbar unter

- <http://www.tagesspiegel.de/berlin/schule/rueckstellungen-in-berlin-zur-schule-mit-fuenf-widerstand-nimmt-deutlich-zu-/7211564.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2013): Einschulung von Fünfjährigen: Stress bleibt Kindersache. In: Der Tagesspiegel, 28.01.2013. Online verfügbar unter <http://www.tagesspiegel.de/berlin/einschulung-von-fuenfjaehrigen-stress-bleibt-kindersache/7695070.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2013): Früheinschulung: Bin ich schon reif für eine Prüfung? In: Der Tagesspiegel, 12.02.2013. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/frueheinschulung-bin-ich-schon-reif-fuer-eine-pruefung/7768318.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2013): Umstrittene Früheinschulung: Viele Eltern klagen über Willkür. In: Der Tagesspiegel, 12.03.2013. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/umstrittene-frueheinschulung-viele-eltern-klagen-ueber-willkuer/7911368.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2013): Untersuchung zur frühen Schulpflicht: Früheinschulung: CDU und SPD wollen sich nicht abspeisen lassen. In: Der Tagesspiegel, 05.06.2013. Online verfügbar unter <http://www.tagesspiegel.de/berlin/untersuchung-zur-fruehen-schulpflicht-frueheinschulung-cdu-und-spd-wollen-sich-nicht-abspeisen-lassen/8299698.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2013): Ausblick auf das neue Schuljahr: 3300 Eltern verschieben die Schulpflicht. In: Der Tagesspiegel, 02.08.2013. Online verfügbar unter <http://www.tagesspiegel.de/berlin/ausblick-auf-das-neue-schuljahr-3300-eltern-verschieben-die-schulpflicht/8583476.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2014): Grundschulpreis für FU-Forscher Ramseger: Das Dilemma der Integration. In: Der Tagesspiegel, 12.05.2014. Online verfügbar unter <http://www.tagesspiegel.de/wissen/grundschulpreis-fuer-fu-forscher-ramseger-das-dilemma-der-integration/9881852.html>, zuletzt geprüft am 28.03.2017.
- Vieth-Entus, Susanne (2014): Berliner Eltern rebellieren gegen die Früheinschulung. In: Der Tagesspiegel, 30.06.2014. Online verfügbar unter <http://www.tagesspiegel.de/berlin/bildung-berliner-eltern-rebellieren-gegen-die-frueheinschulung/10112436.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2014): Neue Zahlen: Noch mehr Eltern boykottieren die Früheinschulung. In: Der Tagesspiegel, 17.10.2014. Online verfügbar unter <http://www.tagesspiegel.de/berlin/neue-zahlen-noch-mehr-eltern-boykottieren-die-frueheinschulung/10849868.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2015): Schule mit sechs Jahren: Ende der Früheinschulung erhöht Druck auf Kitas. In: Der Tagesspiegel, 13.01.2015. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/schule-mit-sechs-jahren-ende-der-frueheinschulung-erhoeht-druck-auf-kitas/11219480.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2015): Neuer Rekord bei Zurückstellungen: 6400 Kinder blieben in der Kita. In: Der Tagesspiegel, 18.10.2015. Online verfügbar unter <http://www.tagesspiegel.de/berlin/neuer-rekord-bei-zurueckstellungen-6400-kinder-blieben-in-der-kita/12462168.html>, zuletzt geprüft am 27.03.2017.
- Vieth-Entus, Susanne (2016): Neue Zahlen für Berliner Erstklässler: 8500 Schüler werden später eingeschult. In: Der Tagesspiegel, 02.12.2016. Online verfügbar unter

<http://www.tagesspiegel.de/berlin/neue-zahlen-fuer-berliner-erstklaessler-8500-schueler-werden-spaeter-ingeschult/14923624.html>, zuletzt geprüft am 27.03.2017.

Vieth-Entus, Susanne; Vogt, Sylvia (2012): Überforderte Kinder: Eltern und Ärzte warnen vor Früheinschulungen. In: Der Tagesspiegel, 07.08.2012. Online verfügbar unter <http://www.tagesspiegel.de/berlin/schule/ueberforderte-kinder-eltern-und-aerzte-warnen-vor-frueheinschulungen/6968890.html>, zuletzt geprüft am 27.03.2017.

Vieth-Entus, Susanne; Warnecke, Tilmann (2012): Nach Grundschul-Studie: CDU stellt Früheinschulung infrage. In: Der Tagesspiegel, 05.10.2012. Online verfügbar unter <http://www.tagesspiegel.de/berlin/nach-grundschul-studie-cdu-stellt-frueheinschulung-infrage/7221882.html>, zuletzt geprüft am 27.03.2017.

Vogt, Sylvia (2012): Grundschulen: Schulanfänger bleiben auf der Reform sitzen. In: Der Tagesspiegel, 31.10.2012. Online verfügbar unter <http://www.tagesspiegel.de/berlin/grundschulen-schulanfaenger-bleiben-auf-der-reform-sitzen/7323350.html>, zuletzt geprüft am 27.03.2017.

Wierth, Alke (2007): Jahrgangsübergreifendes Lernen: Aus zwei mach eins. In: taz.de, 01.09.2007. Online verfügbar unter <https://www.taz.de/Archiv-Suche/%215195692&s=jahrgangs%C3%BCbergreifendes+lernen+berlin/>, zuletzt geprüft am 28.03.2017.

A.4.Ergänzende Tabellen Studie II

Tab. A- 1: Varianzanalysen zur Schulqualität mit vier latenten Klassen

Merkmal	4 latente Klassen				df (3,)	F	p	η^2
	M (SD)							
	1	2	3-5	6				
Schulqualität								
Arbeits- und Kommunikationskultur im Kollegium	3,22 (,90)	3,36 (,73)	3,40 (,82)	3,64 (,63)	278	3,58	,014	,04
Beteiligung der Schülerinnen und Schüler und der Eltern	3,00 (,77)	2,78 (1,01)	2,82 (,76)	3,07 (,77)	277	2,12	,098	,02
Evaluation schulischer Entwicklungsvorhaben	2,77 (1,05)	2,85 (,99)	3,07 (,86)	3,27 (,86)	274	4,14	,007	,04
Fortschreibung des Schulprogramms	2,58 (1,06)	2,77 (1,04)	2,72 (1,05)	3,05 (,91)	278	2,58	,054	,03
Personalentwicklung und Personaleinsatz	2,67 (,85)	2,72 (,91)	2,86 (,91)	3,09 (,75)	276	3,69	,012	,04
Schulleistungsdaten und Schullaufbahn	3,49 (,66)	3,39 (,78)	3,32 (,78)	3,68 (,54)	277	3,93	,009	,04
Schulleitungshandeln und Qualitätsmanagement	2,98 (,92)	2,96 (,95)	3,07 (,84)	3,41 (,79)	277	4,67	,003	,05
Schulleitungshandeln und Schulgemeinschaft	3,53 (,85)	3,48 (,84)	3,43 (,87)	3,78 (,50)	268	3,09	,028	,03

Tab. A- 2: Varianzanalysen zur Schulqualität mit sechs latenten Klassen und den Klassen der „allg. Nicht-Überzeugten“

Merkmal	6 latente Klassen						"allg. Nicht-Überzeugte"										
	M (SD)						df (5,)	F	p	η^2	M (SD)			df (2,)	F	p	η^2
	1	2	3	4	5	6					3	4	5				
	Schulqualität																
Arbeits- und Kommunikationskultur im Kollegium	3,22 (,90)	3,42 (,58)	3,28 (,77)	3,43 (,79)	3,40 (,82)	3,64 (,63)	276	2,31	,044	,04	3,28 (,77)	3,43 (,79)	3,40 (,82)	125	,41	,667	,01
Beteiligung der Schülerinnen und Schüler und der Eltern	3,00 (,77)	2,71 (1,08)	2,88 (,98)	2,67 (1,00)	2,82 (,76)	3,07 (,77)	275	1,52	,182	,03	2,88 (,98)	2,67 (1,00)	2,82 (,76)	124	,50	,606	,01
Evaluation schulischer Entwicklungsvorhaben	2,77 (1,05)	3,04 (1,00)	2,81 (1,05)	2,75 (,89)	3,07 (,86)	3,27 (,86)	272	2,76	,019	,05	2,81 (1,05)	2,75 (,89)	3,07 (,86)	123	1,47	,234	,02
Fortschreibung des Schulprogramms	2,58 (1,06)	2,96 (1,00)	2,65 (1,04)	2,79 (1,07)	2,72 (1,05)	3,05 (,91)	276	1,84	,106	,03	2,65 (1,04)	2,79 (1,07)	2,72 (1,05)	125	,14	,867	,00
Personalentwicklung und Personaleinsatz	2,67 (,85)	2,87 (,92)	2,65 (,97)	2,70 (,82)	2,86 (,91)	3,09 (,75)	274	2,40	,037	,04	2,65 (,97)	2,70 (,82)	2,86 (,91)	124	,69	,504	,01
Schulleistungsdaten und Schullaufbahn	3,49 (,66)	3,50 (,78)	3,33 (,81)	3,41 (,75)	3,32 (,78)	3,68 (,54)	275	2,55	,028	,04	3,33 (,81)	3,41 (,75)	3,32 (,78)	124	,13	,875	,00
Schulleitungshandeln und Qualitätsmanagement	2,98 (,92)	2,96 (,93)	3,00 (,93)	2,89 (1,03)	3,07 (,84)	3,41 (,79)	275	2,83	,016	,05	3,00 (,93)	2,89 (1,03)	3,07 (,84)	125	,36	,701	,01
Schulleitungshandeln und Schulgemeinschaft	3,53 (,85)	3,70 (,63)	3,38 (,94)	3,44 (,85)	3,43 (,87)	3,78 (,50)	266	2,37	,039	,04	3,38 (,94)	3,44 (,85)	3,43 (,87)	122	,05	,949	,00

Tab. A- 3: Varianzanalysen zur Unterrichtsqualität mit vier latenten Klassen

Merkmal	4 latente Klassen				df (3,)	F	p	η^2
	M (SD)							
	1	2	3-5	6				
Unterrichtsqualität								
Fachimmanentes, fachübergreifendes und fächerverbindendes Lernen	3,13 (,34)	3,21 (,41)	3,07 (,26)	3,21 (,41)	278	2,22	,086	,02
Förderung von Selbstvertrauen und Selbsteinschätzung	3,09 (,36)	3,13 (,33)	3,11 (,36)	3,12 (,32)	278	,14	,937	,00
Innere Differenzierung	1,56 (,66)	1,72 (,68)	1,68 (,76)	1,92 (,73)	278	2,96	,033	,03
Kooperation des pädagogischen Personals	3,54 (,69)	3,63 (,59)	3,60 (,54)	3,70 (,63)	192	,53	,664	,01
Kooperatives Lernen	1,47 (,55)	1,54 (,58)	1,56 (,63)	1,52 (,63)	278	,23	,877	,00
Lehr- und Lernzeit	3,93 (,25)	3,97 (,18)	3,89 (,36)	3,96 (,19)	278	1,36	,255	,01
Leistungs- und Anstrengungsbereitschaft	3,62 (,61)	3,74 (,47)	3,63 (,56)	3,79 (,47)	278	1,64	,180	,02
Lern- und Arbeitsbedingungen	3,98 (,15)	4,00 (,00)	4,00 (,00)	4,00 (,00)	278	1,77	,153	,02
Methoden- und Medienwahl	3,60 (,69)	3,72 (,52)	3,56 (,50)	3,72 (,50)	278	1,46	,226	,02
Pädagogisches Klima im Unterricht	3,91 (,36)	3,95 (,22)	3,95 (,23)	4,00 (,00)	278	1,86	,136	,02
Problemorientiertes Lernen	1,07 (,25)	1,16 (,37)	1,25 (,47)	1,26 (,52)	278	2,51	,059	,03
Schulinternes Curriculum und Abstimmung des Lehr- und Lernangebots	3,18 (,91)	3,16 (,73)	3,16 (,70)	3,31 (,79)	278	,69	,561	,01
Selbstständiges Lernen	1,29 (,46)	1,27 (,49)	1,40 (,65)	1,52 (,59)	278	3,36	,019	,03
Sprach- und Kommunikationsförderung	2,64 (,61)	2,66 (,59)	2,70 (,68)	2,66 (,61)	278	,09	,968	,00
Strukturierung und transparente Zielausrichtung	3,71 (,46)	3,73 (,45)	3,75 (,43)	3,73 (,45)	278	,09	,968	,00
Systematische Unterstützung, Förderung und Beratung	3,53 (,66)	3,57 (,65)	3,58 (,53)	3,72 (,53)	278	1,39	,245	,01
Verhalten im Unterricht	4,00 (,00)	4,00 (,00)	3,95 (,23)	3,99 (,11)	278	2,73	,044	,03

Tab. A- 4: Varianzanalysen zur Unterrichtsqualität mit sechs latenten Klassen und den Klassen der „allg. Nicht-Überzeugten“

Merkmal	6 latente Klassen										"allg. Nicht-Überzeugte"						
	M (SD)						df (5,)	F	p	η^2	M (SD)			df (2,)	F	p	η^2
	1	2	3	4	5	6					3	4	5				
Unterrichtsqualität																	
Fachimmanentes, fachübergreifendes und fächerverbindendes Lernen	3,13 (,34)	3,25 (,44)	3,19 (,39)	3,21 (,42)	3,07 (,26)	3,21 (,41)	276	1,42	,218	,03	3,19 (,39)	3,21 (,42)	3,07 (,26)	125	2,19	,116	,03
Förderung von Selbstvertrauen und Selbsteinschätzung	3,09 (,36)	3,13 (,34)	3,12 (,32)	3,14 (,36)	3,11 (,36)	3,12 (,32)	276	,10	,991	,00	3,12 (,32)	3,14 (,36)	3,11 (,36)	125	,11	,897	,00
Innere Differenzierung	1,56 (,66)	1,67 (,64)	1,72 (,73)	1,75 (,65)	1,68 (,76)	1,92 (,73)	276	1,80	,113	,03	1,72 (,73)	1,75 (,65)	1,68 (,76)	125	,08	,921	,00
Kooperation des pädagogischen Personals	3,54 (,69)	3,75 (,45)	3,59 (,64)	3,57 (,65)	3,60 (,54)	3,70 (,63)	190	,48	,788	,01	3,59 (,64)	3,57 (,65)	3,60 (,54)	83	,01	,987	,00
Kooperatives Lernen	1,47 (,55)	1,67 (,64)	1,49 (,55)	1,50 (,58)	1,56 (,63)	1,52 (,63)	276	,44	,822	,01	1,49 (,55)	1,50 (,58)	1,56 (,63)	125	,21	,807	,00
Lehr- und Lernzeit	3,93 (,25)	4,00 (,00)	3,95 (,21)	3,96 (,19)	3,89 (,36)	3,96 (,19)	276	,93	,462	,02	3,95 (,21)	3,96 (,19)	3,89 (,36)	125	,78	,461	,01
Leistungs- und Anstrengungsbereitschaft	3,62 (,61)	3,79 (,41)	3,72 (,50)	3,71 (,46)	3,63 (,56)	3,79 (,47)	276	1,05	,386	,02	3,72 (,50)	3,71 (,46)	3,63 (,56)	125	,44	,643	,01
Lern- und Arbeitsbedingungen	3,98 (,15)	4,00 (,00)	4,00 (,00)	4,00 (,00)	4,00 (,00)	4,00 (,00)	276	1,05	,386	,02	4,00 (,00)	4,00 (,00)	4,00 (,00)	125	,62	,540	-
Methoden- und Medienwahl	3,60 (,69)	3,75 (,44)	3,67 (,57)	3,75 (,52)	3,56 (,50)	3,72 (,50)	276	,96	,441	,02	3,67 (,57)	3,75 (,52)	3,56 (,50)	125	1,34	,267	,02

Merkmal	6 latente Klassen						"allg. Nicht-Überzeugte"										
	M (SD)						df (5,)	F	p	η^2	M (SD)			df (2,)	F	p	η^2
	1	2	3	4	5	6					3	4	5				
Unterrichtsqualität																	
Pädagogisches Klima im Unterricht	3,91 (,36)	3,96 (,20)	3,91 (,29)	4,00 (,00)	3,95 (,23)	4,00 (,00)	276	1,76	,121	,03	3,91 (,29)	4,00 (,00)	3,95 (,23)	125	1,42	,245	,02
Problemorientiertes Lernen	1,07 (,25)	1,13 (,34)	1,14 (,35)	1,21 (,42)	1,25 (,47)	1,26 (,52)	276	1,64	,150	,03	1,14 (,35)	1,21 (,42)	1,25 (,47)	125	,78	,460	,01
Schulinternes Curriculum und Abstimmung des Lehr- und Lernangebots	3,18 (,91)	3,33 (,64)	3,05 (,79)	3,18 (,72)	3,16 (,70)	3,31 (,79)	276	,84	,522	,01	3,05 (,79)	3,18 (,72)	3,16 (,70)	125	,38	,687	,01
Selbstständiges Lernen	1,29 (,46)	1,33 (,56)	1,23 (,48)	1,29 (,46)	1,40 (,65)	1,52 (,59)	276	2,11	,064	,04	1,23 (,48)	1,29 (,46)	1,40 (,65)	125	1,21	,301	,02
Sprach- und Kommunikationsförderung	2,64 (,61)	2,79 (,41)	2,65 (,69)	2,57 (,57)	2,70 (,68)	2,66 (,61)	276	,38	,860	,01	2,65 (,69)	2,57 (,57)	2,70 (,68)	125	,37	,693	,01
Strukturierung und transparente Zielausrichtung	3,71 (,46)	3,67 (,48)	3,70 (,46)	3,82 (,39)	3,75 (,43)	3,73 (,45)	276	,43	,830	,01	3,70 (,46)	3,82 (,39)	3,75 (,43)	125	,69	,504	,01
Systematische Unterstützung, Förderung und Beratung	3,53 (,66)	3,58 (,65)	3,63 (,66)	3,46 (,64)	3,58 (,53)	3,72 (,53)	276	1,10	,363	,02	3,63 (,66)	3,46 (,64)	3,58 (,53)	125	,64	,528	,01
Verhalten im Unterricht	4,00 (,00)	4,00 (,00)	4,00 (,00)	4,00 (,00)	3,95 (,23)	3,99 (,11)	276	1,63	,153	,03	4,00 (,00)	4,00 (,00)	3,95 (,23)	125	1,93	,150	,03

A.5. Ergänzende Tabellen zu Studie IV

Tab. A- 5: Effektgrößen Cohens h innerhalb der Hauptkohorte und Jahrgangsstufen

Jahrgang	Altersgruppe	P ₁	P ₂	Ø ₁	Ø ₂	h
Gesamt						
2	jung vs. mittel	35,5%	36,5%	1,28	1,30	-,02
	mittel vs. alt	36,5%	28,1%	1,30	1,12	,18
	jung vs. alt	35,5%	28,1%	1,28	1,12	,16
3	jung vs. mittel	33,6%	37,6%	1,24	1,32	-,08
	mittel vs. alt	37,6%	28,8%	1,32	1,13	,19
	jung vs. alt	33,6%	28,8%	1,24	1,13	,10
8	jung vs. mittel	31,4%	39,6%	1,19	1,36	-,17
	mittel vs. alt	39,6%	29,0%	1,36	1,14	,22
	jung vs. alt	31,4%	29,0%	1,19	1,14	,05
Mädchen						
2	jung vs. mittel	50,9%	49,2%	1,59	1,56	,03
	mittel vs. alt	49,2%	46,2%	1,56	1,49	,06
	jung vs. alt	50,9%	46,2%	1,59	1,49	,09
3	jung vs. mittel	51,4%	49,4%	1,60	1,56	,04
	mittel vs. alt	49,4%	46,3%	1,56	1,50	,06
	jung vs. alt	51,4%	46,3%	1,60	1,50	,10
8	jung vs. mittel	53,8%	49,1%	1,65	1,55	,09
	mittel vs. alt	49,1%	47,0%	1,55	1,51	,04
	jung vs. alt	53,8%	47,0%	1,65	1,51	,13
nicht deutsch						
2	jung vs. mittel	34,8%	34,9%	1,26	1,26	,00
	mittel vs. alt	34,9%	35,4%	1,26	1,27	-,01
	jung vs. alt	34,8%	35,4%	1,26	1,27	-,01
3	jung vs. mittel	28,2%	29,3%	1,12	1,14	-,02
	mittel vs. alt	29,3%	29,3%	1,14	1,14	,00
	jung vs. alt	28,2%	29,3%	1,12	1,14	-,02
8	jung vs. mittel	28,2%	30,2%	1,12	1,16	-,04
	mittel vs. alt	30,2%	32,6%	1,16	1,21	-,05
	jung vs. alt	28,2%	32,6%	1,12	1,21	-,10

$h \geq ,20$ klein, $h \geq ,50$ mittel, $h \geq ,80$ groß

Tab. A- 6: Effektgrößen Cohens h innerhalb der Vergleichskohorte I und Jahrgangsstufen

Jahrgang	Altersgruppe	P ₁	P ₂	Ø ₁	Ø ₂	h
Gesamt						
2	mittel vs. alt	52,6%	47,4%	1,62	1,52	,11
8	mittel vs. alt	51,5%	48,5%	1,60	1,54	,06
Mädchen						
2	mittel vs. alt	50,8%	46,1%	1,59	1,49	,09
8	mittel vs. alt	51,0%	46,3%	1,59	1,50	,09
nicht deutsch						
2	mittel vs. alt	32,6%	32,5%	1,21	1,21	,00
8	mittel vs. alt	32,6%	33,4%	1,21	1,23	-,02

$h \geq ,20$ klein, $h \geq ,50$ mittel, $h \geq ,80$ groß

Tab. A- 7: Effektgrößen Cohens h innerhalb der Vergleichskohorte II und Jahrgangsstufen

Jahrgang	Altersgruppe	P ₁	P ₂	Ø ₁	Ø ₂	h
Gesamt						
3	jung vs. mittel	47,4%	52,6%	1,52	1,62	-,10
Mädchen						
3	jung vs. mittel	51,5%	48,2%	1,60	1,53	,07
nicht deutsch						
3	jung vs. mittel	29,8%	28,8%	1,16	1,13	,02

$h \geq ,20$ klein, $h \geq ,50$ mittel, $h \geq ,80$ groß

Tab. A- 8: Effektgrößen Cohens h innerhalb der Hauptkohorte und über die Jahrgangsstufen

Altersgruppe	Jahrgang	P ₁	P ₂	Ø ₁	Ø ₂	h
Gesamt						
jung	2 vs. 3	35,5%	33,6%	1,28	1,24	,04
	2 vs. 8	35,5%	31,4%	1,28	1,19	,09
	3 vs. 8	33,6%	31,4%	1,24	1,19	,05
mittel	2 vs. 3	36,5%	37,6%	1,30	1,32	-,02
	2 vs. 8	36,5%	39,6%	1,30	1,36	-,06
	3 vs. 8	37,6%	39,6%	1,32	1,36	-,04
alt	2 vs. 3	28,1%	28,8%	1,12	1,13	-,02
	2 vs. 8	28,1%	29,0%	1,12	1,14	-,02
	3 vs. 8	28,8%	29,0%	1,13	1,14	,00
Mädchen						
jung	2 vs. 3	50,9%	51,4%	1,59	1,60	-,01
	2 vs. 8	50,9%	53,8%	1,59	1,65	-,06
	3 vs. 8	51,4%	53,8%	1,60	1,65	-,05
mittel	2 vs. 3	49,2%	49,4%	1,56	1,56	,00
	2 vs. 8	49,2%	49,1%	1,56	1,55	,00
	3 vs. 8	49,4%	49,1%	1,56	1,55	,01
alt	2 vs. 3	46,2%	46,3%	1,49	1,50	,00
	2 vs. 8	46,2%	47,0%	1,49	1,51	-,02
	3 vs. 8	46,3%	47,0%	1,50	1,51	-,02
nicht deutsch						
jung	2 vs. 3	34,8%	28,2%	1,26	1,12	,14
	2 vs. 8	34,8%	28,2%	1,26	1,12	,14
	3 vs. 8	28,2%	28,2%	1,12	1,12	,00
mittel	2 vs. 3	34,9%	29,3%	1,26	1,14	,12
	2 vs. 8	34,9%	30,2%	1,26	1,16	,10
	3 vs. 8	29,3%	30,2%	1,14	1,16	-,02
alt	2 vs. 3	35,4%	29,3%	1,27	1,14	,13
	2 vs. 8	35,4%	32,6%	1,27	1,21	,06
	3 vs. 8	29,3%	32,6%	1,14	1,21	-,07

$h \geq ,20$ klein, $h \geq ,50$ mittel, $h \geq ,80$ groß

Tab. A- 9: Effektgrößen Cohens h innerhalb der Vergleichskohorte I und über die Jahrgangsstufen

Altersgruppe	Jahrgang	P ₁	P ₂	Ø ₁	Ø ₂	h
Gesamt						
mittel	2 vs. 8	52,6%	51,5%	1,62	1,60	,02
alt	2 vs. 8	47,4%	48,5%	1,52	1,54	-,02
Mädchen						
mittel	2 vs. 8	50,8%	51,0%	1,59	1,59	,00
alt	2 vs. 8	46,1%	46,3%	1,49	1,50	,00
nicht deutsch						
mittel	2 vs. 8	32,6%	32,6%	1,21	1,21	,00
alt	2 vs. 8	32,5%	33,4%	1,21	1,23	-,02

$h \geq ,20$ klein, $h \geq ,50$ mittel, $h \geq ,80$ groß

Tab. A- 10: Effektgrößen Cohens h über die Hauptkohorte und Vergleichskohorte I und innerhalb der Jahrgangsstufen

Jahrgang	Altersgruppe	Kohorte	P ₁	P ₂	Ø ₁	Ø ₂	h
Gesamt							
2	mittel	HK vs. VKH 1	56,6%	52,6%	1,70	1,62	,08
	alt	HK vs. VKH 1	43,4%	47,4%	1,44	1,52	-,08
8	mittel	HK vs. VKH 1	57,8%	51,5%	1,73	1,60	,13
	alt	HK vs. VKH 1	42,2%	48,5%	1,41	1,54	-,13
Mädchen							
2	mittel	HK vs. VKH 1	49,2%	50,8%	1,56	1,59	-,03
	alt	HK vs. VKH 1	46,2%	46,1%	1,49	1,49	,00
8	mittel	HK vs. VKH 1	49,1%	51,0%	1,55	1,59	-,04
	alt	HK vs. VKH 1	47,0%	46,3%	1,51	1,50	,01
nicht deutsch							
2	mittel	HK vs. VKH 1	34,9%	32,6%	1,26	1,21	,05
	alt	HK vs. VKH 1	35,4%	32,5%	1,27	1,21	,06
8	mittel	HK vs. VKH 1	30,2%	32,6%	1,16	1,21	-,05
	alt	HK vs. VKH 1	32,6%	33,4%	1,21	1,23	-,02

$h \geq ,20$ klein, $h \geq ,50$ mittel, $h \geq ,80$ groß

Tab. A- 11: Effektgrößen Cohens h über die Hauptkohorte und Vergleichskohorte II und innerhalb der Jahrgangsstufen

Jahrgang	Altersgruppe	Kohorte	P ₁	P ₂	Ø ₁	Ø ₂	h
Gesamt							
3	jung	HK vs. VKH 2	47,2%	47,4%	1,52	1,52	,00
	mittel	HK vs. VKH 2	52,8%	52,6%	1,63	1,62	,00
Mädchen							
3	jung	HK vs. VKH 2	51,4%	51,5%	1,60	1,60	,00
	mittel	HK vs. VKH 2	49,4%	48,2%	1,56	1,53	,02
nicht deutsch							
3	jung	HK vs. VKH 2	28,2%	29,8%	1,12	1,16	-,04
	mittel	HK vs. VKH 2	29,3%	28,8%	1,14	1,13	,01

$h \geq ,20$ klein, $h \geq ,50$ mittel, $h \geq ,80$ groß

Tab. A- 12: mittlere Leistung (z-standardisiert) der Altersgruppen der Hauptkohorte in Deutsch / Lesen und Mathematik

Altersgruppe	Deutsch / Lesen					Mathematik				
	alle	weiblich	männlich	deutsch	nichtdeutsch	alle	weiblich	männlich	deutsch	nichtdeutsch
Jahrgangsstufe 2										
jung	-,10 [-,14; -,07]	-,01 [-,05; ,02]	-,20 [-,24; -,16]	,12 [,09; ,15]	-,58 [-,63; -,53]	-,12 [-,16; -,08]	-,21 [-,25; -,17]	-,03 [-,07; ,01]	,08 [,04; ,11]	-,52 [-,57; -,47]
mittel	,05 [,02; ,08]	,14 [,10; ,17]	-,03 [-,07; ,01]	,27 [,24; ,30]	-,42 [-,47; -,37]	,06 [,02; ,09]	-,03 [-,07; ,02]	,13 [,10; ,17]	,27 [,23; ,30]	-,38 [-,43; -,33]
alt	,06 [,03; ,10]	,14 [,11; ,18]	-,01 [-,04; ,03]	,27 [,23; ,30]	-,34 [-,39; -,29]	,08 [,04; ,12]	-,03 [-,08; ,01]	,18 [,13; ,22]	,27 [,23; ,31]	-,29 [-,34; -,24]
Jahrgangsstufe 3										
jung	-,05 [-,09; -,02]	-,03 [-,07; ,01]	-,08 [-,12; -,04]	,14 [,11; ,18]	-,62 [-,68; -,57]	-,07 [-,11; -,04]	-,20 [-,25; -,16]	,06 [,02; ,11]	,09 [,05; ,12]	-,52 [-,57; -,46]
mittel	,04 [,01; ,08]	,09 [,05; ,13]	,00 [-,04; ,04]	,25 [,22; ,28]	-,52 [-,58; -,46]	,05 [,01; ,08]	-,08 [-,12; -,04]	,17 [,13; ,21]	,21 [,17; ,25]	-,39 [-,44; -,33]
alt	,00 [-,03; ,04]	,06 [,01; ,10]	-,04 [-,08; ,00]	,20 [,16; ,24]	-,52 [-,58; -,46]	,02 [-,01; ,06]	-,12 [-,16; -,07]	,15 [,10; ,19]	,18 [,14; ,22]	-,39 [-,45; -,33]
Jahrgangsstufe 8										
jung	,11 [-,03; ,26]	,13 [-,03; ,30]	,09 [-,06; ,23]	,28 [,14; ,42]	-,37 [-,58; -,15]	,02 [-,14; ,18]	-,08 [-,25; ,08]	,14 [-,07; ,35]	,23 [,06; ,39]	-,56 [-,79; -,34]
mittel	-,02 [-,15; ,11]	,03 [-,12; ,18]	-,08 [-,23; ,08]	,18 [,05; ,32]	-,53 [-,70; -,35]	-,01 [-,14; ,12]	-,11 [-,26; ,04]	,08 [-,07; ,23]	,22 [,08; ,35]	-,61 [-,79; -,43]
alt	-,06 [-,21; ,08]	,01 [-,15; ,17]	-,13 [-,29; ,03]	,13 [-,02; ,28]	-,59 [-,79; -,39]	-,05 [-,18; ,09]	-,23 [-,38; -,07]	,11 [-,05; ,27]	,14 [,02; ,27]	-,60 [-,82; -,38]

Tab. A- 13: mittlere Leistung (z-standardisiert) der Altersgruppen der Vergleichskohorte I in Deutsch / Lesen und Mathematik

Altersgruppe	Deutsch / Lesen					Mathematik				
	alle	weiblich	männlich	deutsch	nichtdeutsch	alle	weiblich	männlich	deutsch	nichtdeutsch
Jahrgangsstufe 2										
mittel	,00 [-,04; ,04]	,03 [-,02; ,07]	-,04 [-,08; ,01]	,24 [,20; ,27]	-,55 [-,60; -,50]	-,01 [-,06; ,04]	-,10 [-,15; -,05]	,08 [,03; ,13]	,19 [,14; ,23]	-,44 [-,51; -,38]
alt	,00 [-,04; ,04]	,03 [-,01; ,08]	-,02 [-,07; ,02]	,24 [,20; ,28]	-,51 [-,56; -,46]	,01 [-,03; ,06]	-,08 [-,13; -,03]	,09 [,04; ,14]	,19 [,15; ,23]	-,37 [-,44; -,31]
Jahrgangsstufe 8										
mittel	,04 [-,05; ,13]	,04 [-,07; ,14]	,04 [-,06; ,14]	,29 [,20; ,37]	-,51 [-,62; -,40]	,02 [-,08; ,11]	-,09 [-,20; ,01]	,13 [,02; ,23]	,26 [,17; ,34]	-,53 [-,65; -,42]
alt	-,04 [-,12; ,05]	-,02 [-,11; ,07]	-,05 [-,15; ,05]	,19 [,10; ,28]	-,52 [-,63; -,42]	-,05 [-,14; ,05]	-,18 [-,29; -,07]	,07 [-,04; ,18]	,16 [,06; ,26]	-,49 [-,62; -,36]

Tab. A- 14: mittlere Leistung (z-standardisiert) der Altersgruppen der Vergleichskohorte II in Deutsch / Lesen und Mathematik

Altersgruppe	Deutsch / Lesen					Mathematik				
	alle	weiblich	männlich	deutsch	nichtdeutsch	alle	weiblich	männlich	deutsch	nichtdeutsch
Jahrgangsstufe 3										
jung	-,05 [-,17; ,07]	,01 [-,14; ,16]	-,11 [-,24; ,02]	,19 [,09; ,29]	-,65 [-,87; -,43]	-,03 [-,17; ,10]	-,11 [-,27; ,06]	,04 [-,10; ,19]	,16 [,03; ,30]	-,52 [-,74; -,30]
mittel	,02 [-,10; ,14]	,13 [,00; ,25]	-,07 [-,22; ,07]	,25 [,16; ,35]	-,56 [-,77; -,34]	,03 [-,12; ,17]	-,03 [-,18; ,12]	,08 [-,08; ,24]	,29 [,15; ,43]	-,59 [-,79; -,39]

IV.3 Sind Überlappungen der altersspezifischen Leistungen festzustellen und, wenn ja, wie lassen sich diese beschreiben?

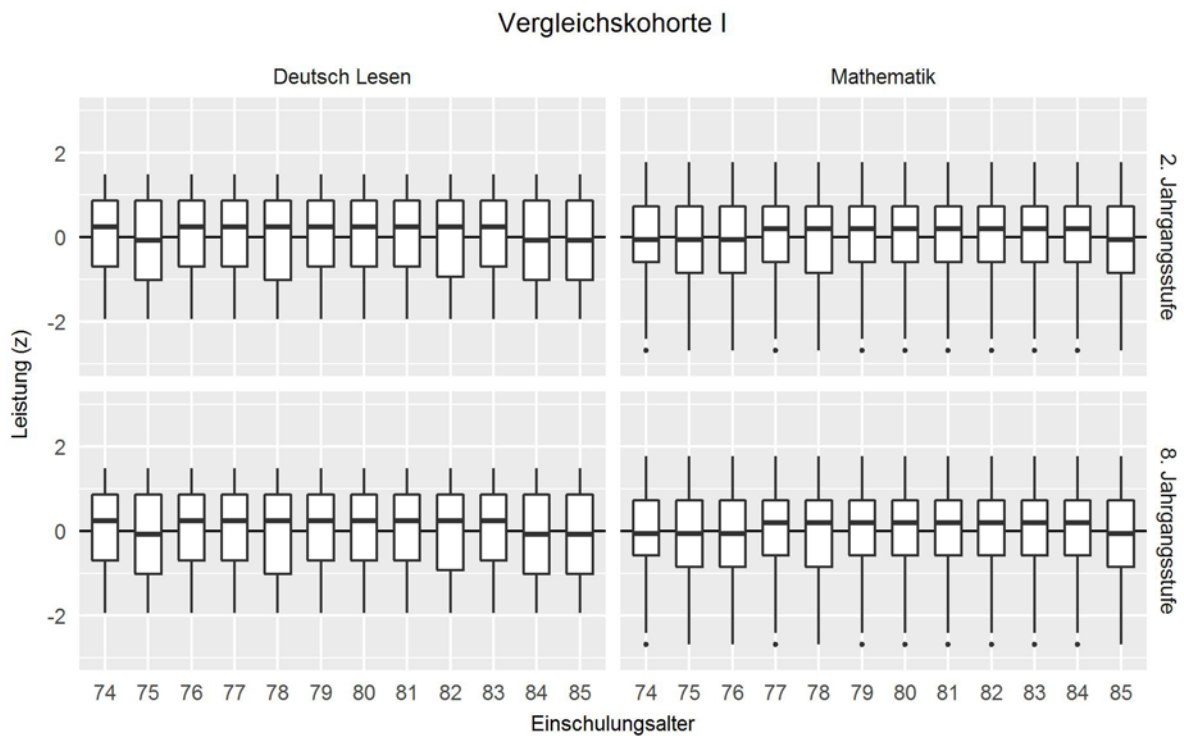


Abb. A- 1: Leistungsverteilung (Boxplots) der Vergleichskohorte I nach Einschulungsalter (in Monaten) für Deutsch / Lesen und Mathematik

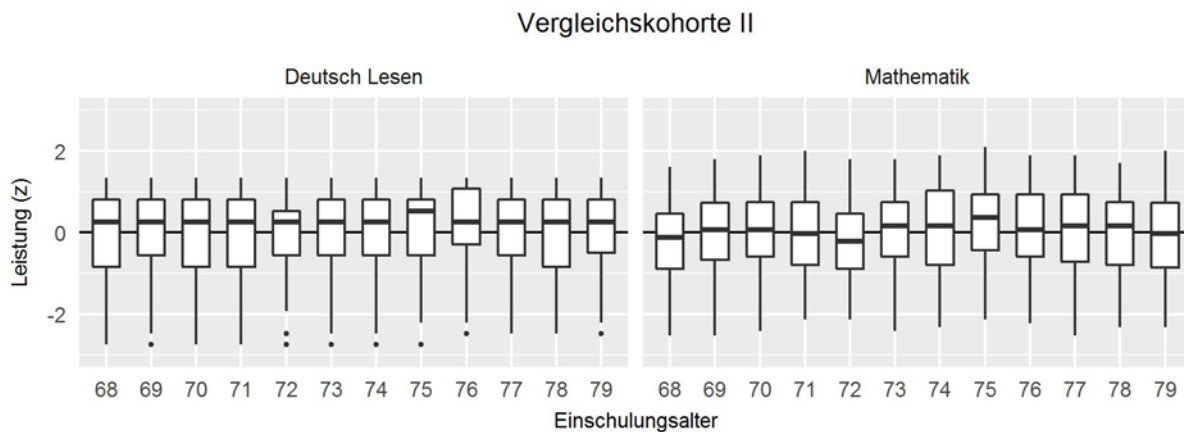


Abb. A- 2: Leistungsverteilung (Boxplots) der Vergleichskohorte II nach Einschulungsalter (in Monaten) für Deutsch / Lesen und Mathematik

Erklärung

Hiermit versichere ich, dass ich die Dissertation „Implementationserfolg von Schulreformen in der Berliner Schulanfangsphase“ selbstständig verfasst habe. Alle Hilfsmittel, die ich verwendet habe, sind angegeben. Die Arbeit ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Berlin, im März, 2017

Katharina Thoren