

Technische Universität Dortmund

Fakultät Erziehungswissenschaft, Psychologie und Soziologie

Ganztagsschule in Deutschland: Effekte auf Leistungsniveau und soziale
Bildungsungleichheit

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

Vorgelegt von Isa Steinmann

Geboren am 19.10.1989 in Schwerte

Matrikelnummer 182545

Erstgutachter: Prof. Dr. Wilfried Bos

Zweitgutachter: PD Dr. Rolf Strietholt

Februar 2019

Dies ist eine Veröffentlichung als Dissertation
an der Fakultät Erziehungswissenschaft, Psychologie und Soziologie
an der Technischen Universität Dortmund.

Die Arbeit wurde am 12.12.2018 schriftlich eingereicht und
am 08.02.2019 in Dortmund mündlich verteidigt.

Danksagung

Ich möchte die Gelegenheit nutzen und mich bei all jenen bedanken, die mich auf dem Weg zu dieser Dissertation unterstützt haben. Als erstes möchte ich meinen Betreuern Wilfried Bos und Rolf Strietholt für ihre langjährige Betreuung meiner Dissertation aber auch für die vielfältigen Erfahrungen danken, die ich darüber hinaus am Institut für Schulentwicklungsforschung (IFS) sammeln konnte. In den vier Jahren konnte ich an einer Vielzahl von Seminaren und Workshops teilnehmen und meine Arbeiten sowohl innerhalb des Instituts als auch auf nationalen und internationalen Fachtagungen präsentieren und hierbei von den konstruktiven Rückmeldungen profitieren. Vor allem Rolf Strietholt möchte ich für die unermüdliche konstruktive Beratung und enge Betreuung danken. Mein Dank gilt auch allen weiteren Kolleginnen und Kollegen am IFS, die mir mit Rat und Tat zur Seite standen, insbesondere Svenja Hartwig. Meine Familie hat die Grundsteine für meinen Weg gelegt und mich auf diesem immer wieder bestärkt. Darüber hinaus gilt mein Dank auch meinem Freund, meinen Freundinnen und Freunden, die sich so manches anhören mussten und mir trotzdem fortwährend den Rücken gestärkt haben. Danke!

Inhaltsverzeichnis

Inhaltsverzeichnis

Tabellenverzeichnis.....	3
Abbildungsverzeichnis	4
Zusammenfassung	5
Deutsche Zusammenfassung	5
Abstract in English	8
1 Einleitung.....	11
1.1 Ganztagsschule in Deutschland.....	11
1.1.1 Politische Definition und Wirkungserwartungen.....	11
1.1.2 Pädagogische Wirkungsannahmen.....	13
1.1.3 Ausgestaltung von Ganztagsschulen.....	14
1.2 Effekte auf Leistung und Bildungsungleichheit.....	16
1.2.1 Effekte der Ganztagsschule in Deutschland.....	16
1.2.2 Effekte erweiterter Bildung und Betreuung im internationalen Raum	20
1.2.3 Effekte erweiterter Unterrichtszeit im internationalen Raum	22
1.3 Forschungsanliegen und Forschungsfragen	24
1.4 Literaturverzeichnis.....	25
2 Beiträge der Arbeit.....	33
2.1 Beitrag I: Ganztagsschulen als neue Organisationsformen – Entwicklungen und Evaluation.....	33
2.2 Beitrag II: Effekte nachmittäglicher Bildungsangebote auf sprachliche Kompetenzentwicklungen.....	67
2.3 Beitrag III: Student achievement and educational inequality in half- and all-day schools: Evidence from Germany	95
2.4 Beitrag IV: Participation in extracurricular activities and student achievement: Evidence from German all-day schools	124

Inhaltsverzeichnis

3	Gesamtdiskussion	152
3.1	Zentrale Befunde	152
3.2	Reichweite und Limitationen	156
3.2.1	Stichproben.....	156
3.2.2	Methodisches Vorgehen.....	157
3.2.3	Unabhängige Variablen.....	158
3.2.4	Abhängige Variablen.....	159
3.3	Implikationen	161
3.4	Fazit.....	164
3.5	Literaturverzeichnis.....	166
4	Anhang.....	175
4.1	Veröffentlichungsstatus der Einzelbeiträge.....	175
4.2	Eigenanteile bei den Einzelbeiträgen	176
4.2.1	Beitrag I.....	176
4.2.2	Beitrag II	176
4.2.3	Beitrag III	177
4.2.4	Beitrag IV	177
4.3	Eidesstattliche Erklärung.....	178

Tabellenverzeichnis

Tabellenverzeichnis

Beitrag	Titel	S.
Beitrag II	Tabelle 1 Deskriptive Statistiken der Schüler/-innen und Schulen	76
Beitrag II	Tabelle 2 Ergebnisse der Mehrebenenmodelle für die Sprachbewusstheit in Deutsch	80
Beitrag II	Tabelle 3 Ergebnisse der Mehrebenenmodelle für den C-Test in Englisch	82
Beitrag III	Table 1 Descriptive statistics for student and school characteristics	102
Beitrag III	Table 2 Predictors of growth in achievement at half- and all-day schools	109
Beitrag III	Table 3 Predictors of growth in achievement inequality related to parental education at half- and all-day schools	110
Beitrag III	Table 4 Predictors of growth in achievement inequality related to language of origin at half- and all-day schools	111
Beitrag III	Appendix A1 Descriptive statistics of student and school characteristics prior to imputation	121
Beitrag III	Appendix A2 Variables in the imputation model on the student and school level: Sources, measurement points, and percentages of missing values	122
Beitrag III	Appendix A3 Descriptive statistics for the subsamples of all-day schools with high and low student participation profiles for robustness analyses	123
Beitrag IV	Table 1. Treatment and control groups' comparison statistics of covariates in sample 5–7.	135
Beitrag IV	Table 2. Treatment effects on mathematics and reading in sample 5–7.	138
Beitrag IV	Table 3. Treatment effects on mathematics and reading in sample 7–9.	139
Beitrag IV	Appendix A1. Descriptive covariate statistics for sample 5–7 and sample 7–9.	149
Beitrag IV	Appendix A2. Treatment and control groups' comparison statistics of covariates in sample 7–9.	150
Beitrag IV	Appendix A3. Variables in the imputation model.	150

Abbildungsverzeichnis

Abbildungsverzeichnis

Beitrag	Titel	S.
Beitrag I	Abbildung 1: Prozentuale Anteile an schulischen Verwaltungseinheiten in öffentlicher und privater Trägerschaft und an Schüler/-innen mit Ganztagsbetrieb zwischen 2002 und 2016	36
Beitrag I	Abbildung 2: Prozentuale Anteile an schulischen Verwaltungseinheiten in öffentlicher Trägerschaft und an Schüler/-innen mit Ganztagsbetrieb in den Bundesländern im Jahr 2016	37
Beitrag I	Abbildung 3: Modell der Bildungsqualität außerunterrichtlicher Angebote in der Ganztagschule	50
Beitrag III	Figure 1 Measurement model for latent growth in mathematics achievement	105
Beitrag III	Figure 2 Measurement model for latent growth in social inequality in mathematics achievement	107107

Zusammenfassung

Deutsche Zusammenfassung

Ganztägige Beschulung wird in Deutschland seit 2003 massiv ausgebaut. Während im vergangenen Jahrhundert noch Halbtagschulen die Regel waren, sind dies nun Ganztagschulen. Der politisch forcierte Ganztagschulausbau ist eine der umfangreichsten und teuersten Bildungsreformen in Deutschland. Von Seiten der Politik wurden hieran hohe Erwartungen gerichtet, die sowohl im Bereich der nachmittäglichen Betreuung als auch Bildung lagen. Zu den Bildungszielen ganztägiger Beschulung gehörten unter anderem die Verbesserung der Leistungen von Schüler/-innen sowie die Verminderung sozialer Bildungsungleichheiten. Die vorliegende Dissertation evaluiert diese beiden Bildungsziele und fokussiert hierbei standardisiert gemessene Testleistungen.

Als Kern der Dissertation wird untersucht, inwieweit Ganztagschulen gegenüber Halbtagschulen Mehrwerte für die Leistungsentwicklung von Schüler/-innen und den Abbau sozialer Bildungsungleichheiten in diesen Leistungen bilden. Zunächst wurde in diesem Zusammenhang der Frage nachgegangen, inwiefern Ganztagschulen über Halbtagschulen hinausgehende Bildungspotenziale aufweisen. In einem grundlegenden Handbuchbeitrag wurden zunächst konzeptionelle Grundlagen und Wirkungserwartungen diskutiert, um herauszuarbeiten, wie der Begriff *Ganztagschule* in der Praxis ausgestaltet wird. Es wurde herausgestellt, dass sich Ganztagschulen in vielerlei Hinsicht eher graduell von Halbtagschulen mit Nachmittagsangebot abgrenzen lassen, beispielsweise in Hinblick auf die angebotenen Ganztagsprogramme oder die Teilnahmemerkmale der Schüler/-innen. Vor dem Hintergrund pädagogischer Wirkungsmodelle wurde geschlussfolgert, dass das Ganztagskonzept zwar ein über Halbtagschulen hinausgehendes theoretisches Potenzial zur Leistungssteigerung und zum Bildungsungleichheitsabbau aufweist. Gleichzeitig scheint es in der Praxis häufig qualitativ nicht so ausgestaltet zu werden, dass sich entsprechende Effekte ganztägiger Beschulung erwarten lassen.

Effekte des Ganztags wurden in drei empirischen Studien geprüft. Zunächst wurde für den Beginn des Ganztagschulausbaus untersucht, ob das schulische Anbieten nachmittäglicher Förderungsprogramme Effekte auf Leistungsentwicklungen von Neuntklässler/-innen oder auf soziale Bildungsungleichheiten aufwies. Es wurde eine Stichprobe von $N=8958$ Schüler/-innen an 184 Schulen der repräsentativen DESI-Studie

(Deutsch-Englisch-Schülerleistungen-International) aus dem Schuljahr 2003/04 analysiert. In Mehrebenenregressionen wurden Leistungen von Schüler/-innen am Ende der neunten Klasse auf das schulische Angebot nachmittäglicher Förderprogramme sowie die Interaktion mit individuellen Herkunftsmerkmalen regressiert. Kontrolliert wurden die individuelle Vorleistung zu Beginn der neunten Klasse sowie die Schulform und Schulkomposition. Es zeigten sich weder Unterschiede in der sprachlichen Entwicklung in Deutsch und Englisch als Fremdsprache zwischen Schulen mit und ohne Nachmittagsangebot, noch in der Entwicklung herkunftsbedingter Ungleichheiten in diesen Leistungsmaßen.

In einer zweiten Studie wurde für den späteren Verlauf des Ganztagschulausbaus untersucht, ob sich Halb- und Ganztagschulen hinsichtlich der Leistungsentwicklungen zwischen fünfter und neunter Klasse unterschieden, oder hinsichtlich sozialer Ungleichheiten in diesen Entwicklungen. Der Studie lag eine Stichprobe von $N=3024$ Fünftklässler/-innen an 144 Schulen zugrunde, die aus der repräsentativen NEPS-Studie (National Educational Panel Study) stammte. In linearen latenten Mehrebenenwachstumsmodellen wurden die Entwicklungen der Leistungen der Schüler/-innen und der sozialen Bildungsungleichheit zwischen drei Testzeitpunkten in den Jahrgangsstufen 5, 7 und 9 analysiert. In separaten Analysen wurden die Leistungs- und Ungleichheitsentwicklungen durch die Organisationsform (Halb- versus Ganztagschule) erklärt. Kontrolliert wurden hierbei die Schulform und Schulkomposition. Der Untersuchungszeitraum reichte vom Schuljahr 2010/11 bis 2014/15. Es wurden weder Effekte der ganztägigen Beschulung auf die Entwicklung der Mathematik- oder Leseleistung von Fünftklässler/-innen gefunden, noch auf die Entwicklung herkunftsbedingter Ungleichheiten.

Die dritte empirische Studie, die Teil der Dissertation war, untersuchte auf Ebene der Schüler/-innen, ob die individuelle Teilnahme an lernförderlich ausgerichteten Ganztagsangeboten offener Ganztagschulen leistungssteigernd wirkte. Ebenfalls basierend auf Daten aus NEPS zwischen 2010/11 und 2014/15 wurden zwei separate Stichproben an offenen Ganztagschulen untersucht: $N=1131$ Fünftklässler/-innen an 43 Schulen wurden bis zur siebten Jahrgangsstufe verfolgt und $N=1545$ Siebtklässler/-innen an 64 Schulen bis zur neunten Klassenstufe. Für beide Stichproben wurde die Leistung der Schüler/-innen in Regressionen mit schulfixierten Effekten auf die Teilnahme an lernförderlich ausgerichteten Ganztagsangeboten zurückgeführt, während die Vorleistung und weitere Schüler/-innenmerkmale kontrolliert wurden. An solchen Angeboten teilnehmende Schüler/-innen

wiesen weder zwischen fünfter und siebter noch zwischen siebter und neunter Klassenstufe bessere Mathematik- und Leseleistungsentwicklungen auf, als nichtteilnehmende.

Die Befunde der vorliegenden Dissertation sind konform mit früheren Studien und erweitern den Forschungsstand in verschiedener Hinsicht. Besondere methodische Mehrwerte der Dissertation bestehen im Vergleich zu früheren Untersuchungen darin, dass a) die durchgeführten Studien auf für Deutschland repräsentativen Datensätzen beruhen, b) die Untersuchungszeiträume sowohl den Beginn als auch weiteren Verlauf des Ganztagschulenausbaus abdeckten und c) dass Stichproben zwischen fünfter und neunter Jahrgangsstufe untersucht wurden. Es wurden d) längsschnittliche Untersuchungen vorgenommen, bei denen verschiedene Ursachen etwaiger Konfundierungen berücksichtigt wurden. Aus inhaltlicher Sicht bietet die Dissertation vielfältige Operationalisierungen e) ganztägiger Beschulung ebenso wie f) der abhängigen Maße der Leistung und Bildungsungleichheit von Schüler/-innen. Limitationen der vorliegenden Arbeiten werden diskutiert, die unter anderem mit der Sekundäranalyse von Beobachtungsdaten und der Operationalisierung ganztägiger Beschulung zusammenhängen. In Bezug auf den Fokus der Dissertation auf Leistungen und Ungleichheiten in standardisierten Testmaßen bietet insbesondere der Handbuchbeitrag eine weitergehende Einordnung. Hier wurde ergänzend herausgestellt, dass frühere Studien zu Wirkungen ganztägiger Beschulung beispielsweise eine chancengerechtere Schüler/-innenteilhabe an Nachmittagsprogrammen aufzeigten und unter bestimmten Bedingungen Effekte auf Motivationen, Selbstkonzepte, Sozialverhalten oder Schulnoten.

Die Befunde der vorliegenden Dissertation deuten in der Gesamtschau mit vorherigen Arbeiten darauf hin, dass Ganztagschulen zumindest durchschnittlich nicht so ausgestaltet werden, dass sich Effekte auf die untersuchten Leistungs- und Ungleichheitsmaße zeigen. Es wird diskutiert, inwiefern die definitorischen und regulatorischen Rahmenbedingungen des Ganztagschulenausbaus zu der heterogenen qualitativen Ausgestaltung des Ganztagskonzepts beigetragen haben könnten. Mögliche Implikationen für zukünftige Investitionen in eine qualitative Weiterentwicklung der Ganztagschulen werden auch vor dem Hintergrund pädagogischer Wirkungsmodelle thematisiert. Hierbei wird herausgestellt, dass weitere Forschung notwendig ist, um Qualitätsbedingungen von effektiv lernförderlichen und bildungsungleichheitsreduzierenden Ganztagschulen zu identifizieren. Die vorliegende Dissertation stellt eine Bereicherung des summativ evaluativen Forschungsstands dar und bietet Ansatzpunkte für künftige formativere Studien.

Abstract in English

All-day schooling has been massively expanded in Germany since 2003. While half-day schools were most common in the last century, the majority of German schools are organized as all-day schools nowadays. The expansion of all-day schools is one of the most extensive and expensive educational reforms in Germany. Political expectations were high, both in terms of childcare and education. The educational aims included, among others, the improvement of student performance and the reduction of social educational inequalities. The present dissertation evaluates these two educational aims, focusing on standardized test scores.

The main objective of the dissertation is to examine the extent to which all-day schools provide an additional value over half-day schools in terms of student achievement developments and the reduction of social educational inequalities in achievement. In this context, the first research question is to what extent all-day schools provide an educational capacity over and above half-day schools. In a comprehensive handbook article, conceptual principles and effectiveness expectations were first discussed in order to elaborate how the term *all-day school* is applied in practice. It was demonstrated that in many respects all-day schools can be distinguished rather gradually from half-day schools with afternoon programs, for example with regard to the all-day programs or the participation characteristics of the students. Against the background of pedagogical effectiveness models, it was concluded that the all-day concept has a theoretical potential beyond half-day schools of increasing performance and reducing educational inequality. In practice at the same time, it often does not seem to be qualitatively implemented in such a way that corresponding effects of all-day schooling can be expected.

The effects of all-day schooling were investigated in three empirical studies. First, for the beginning of the all-day school expansion, it was examined whether the provision of afternoon support at school had effects on ninth graders' performance developments or on social educational inequalities. A sample of $N=8958$ students at 184 schools of the representative DESI study (Deutsch-Englisch-Schülerleistungen-International) from the school year 2003/04 was analyzed. In multi-level regressions, student achievement at the end of ninth grade was regressed on the schools' provision of afternoon programs as well as on the interaction with individual background characteristics. Individual prior achievement at the beginning of ninth grade as well as the school type and school composition were controlled.

Neither differences in the language development in German and English as a foreign language were found between schools with and without afternoon support options, nor in the development of social inequalities in performance.

In a second study, for the later course of the all-day school extension, it was examined whether half-day and all-day schools differed in terms of performance developments between fifth and ninth grade and in terms of social inequalities in these developments. The study based on a sample of $N=3024$ fifth graders at 144 schools from the representative NEPS study (National Educational Panel Study). In linear latent multi-level growth models, the development of student achievement and social educational inequality were analyzed using three test points in grades 5, 7, and 9. In separate analyses, the achievement and inequality developments were predicted by the school organization form (half- versus all-day school), controlling for school type and school composition. The investigation period was from 2010/11 until 2014/15. Neither effects of all-day schooling on the development of mathematics or reading performance of fifth graders nor on the development of social inequalities were found.

The third empirical study that was part of the dissertation examined at student level, whether individual participation in academic all-day programs of nonmandatory all-day schools had a learning-promoting effect. Also based on NEPS data from 2010/11 to 2014/15, two separate samples of nonmandatory all-day schools were examined: $N=1131$ fifth graders at 43 schools were followed to seventh grade and $N=1545$ seventh graders at 64 schools to ninth grade. For both samples, student performance was explained by the participation in academic extracurricular activities in regressions with school fixed effects, while prior achievement and other characteristics were controlled. Students participating in such programs did not show better mathematics and reading performance developments between fifth and seventh grade or between seventh and ninth grade, than non-participants.

These findings are consistent with earlier studies and extend the state of research in various respects. Particularly, methodical added value in contrast to previous studies results from the fact that the conducted studies a) were based on data sets that were representative for Germany, b) covered both the beginning and further course of the all-day school expansion, and c) examined samples between the fifth and ninth grades. Also, d) longitudinal studies were carried out in which various causes of possible confounding were taken into account. From a content perspective, the dissertation provides various operationalizations of e) all-day schooling as well as f) the dependent measures of student achievement and educational

inequality. Limitations of the present work are discussed, which are among other things related to the secondary analysis of observational data and the operationalization of all-day schooling. With regard to dissertation's focus on achievements and inequalities in standardized test scores, particularly the handbook article offers a more comprehensive perspective. For example, it was pointed out that earlier studies on the effects of all-day schooling showed more equal opportunities for afternoon program participation of students and, under certain conditions, effects on motivations, self-concepts, social behavior or school grades.

The findings of the present dissertation indicate together with previous research that all-day schools are, at least on average, not designed in such a way that they affect the examined performance and inequality measures. It is discussed to which extent the definitory and regulatory framework conditions of the all-day school extension could have contributed to the heterogeneous qualitative organization of all-day schooling. Possible implications for future investments in the qualitative further development of all-day schools are also discussed against the background of pedagogical effectiveness models. Here it is stressed that additional research is necessary in order to identify quality conditions of all-day schools that actually promote learning and reduce educational inequality. The present dissertation provides an added value to the summative evaluative state of research and offers starting points for future more formative studies.

1 Einleitung

1.1 Ganztagschule in Deutschland

Die vorliegende Dissertation untersucht Effekte der Schulorganisationsform Ganztagschule auf schulische Leistungen von Schüler/-innen sowie auf soziale Ungleichheiten in diesen Leistungen in Deutschland. Ganztagschulen waren vor der Jahrtausendwende im Vergleich zu Halbtagschulen, die typischerweise nur Vormittagsunterricht und kein Nachmittagsprogramm bereitstellen, die Ausnahme (Bargel & Kuthe, 1991). Dies änderte sich mit dem Investitionsprogramm Zukunft Bildung und Betreuung (IZBB), das aufgrund verschiedener Wirkungserwartungen den Ausbau von Ganztagsplätzen durch die Schaffung neuer und den Ausbau bestehender Ganztagschulen zum Ziel hatte. Das IZBB wurde 2003 von Bund und Ländern verabschiedet und umfasste eine Fördersumme von vier Milliarden Euro Bundes- und circa 400 Millionen Euro Ländermitteln, die bis 2009 abgerufen werden konnten (Bundesministerium für Bildung und Forschung [BMBF], 2003b, 2005). Es trug maßgeblich zu der massiven Umstrukturierung des deutschen Schulsystems von einem primär halbtägigen zu einem überwiegend ganztägigen bei: Laut Statistiken der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) (2008, 2018) waren im Jahr 2002 16,3% aller Schulen Ganztagschulen; in 2016 waren es bereits 67,5%.

Im Folgenden wird beleuchtet, welche Definitionen und Wirkungserwartungen aus politischer Sicht hinter dieser Bildungsreform standen, wie sich diese zu einer pädagogischen Perspektive verhalten und wie die Ganztagsdefinition in Deutschland umgesetzt wird. Im weiteren Verlauf der Einleitung werden empirische Befunde zu Effekten ganztägiger Beschulung zusammengefasst. Hiervon ausgehend wird beschrieben, welche Forschungsanliegen und Forschungsfragen die vorliegende Dissertation untersucht.

1.1.1 Politische Definition und Wirkungserwartungen

Die Vergabe der IZBB-Mittel und die KMK-Ganztagschulstatistik richtete sich nach der folgenden Definition der KMK: Ganztagschulen sind Schulen, an denen

- „an mindestens drei Tagen in der Woche ein ganztägiges Angebot für die Schülerinnen und Schüler bereitgestellt wird, das täglich mindestens sieben Zeitstunden umfasst;

- an allen Tagen des Ganztagschulbetriebs den teilnehmenden Schülerinnen und Schülern ein Mittagessen bereitgestellt wird;
- die Ganztagsangebote unter der Aufsicht und Verantwortung der Schulleitung organisiert und in enger Kooperation mit der Schulleitung durchgeführt werden sowie in einem konzeptionellen Zusammenhang mit dem Unterricht stehen“ (KMK, 2018, S. 4 f.).

Die KMK unterscheidet zwischen drei Ganztagschulformen: An gebundenen Ganztagschulen sind alle Schüler/-innen zur Teilnahme am Ganztagsbetrieb verpflichtet, an teilweise gebundenen ist die Teilnahme für einzelne Schüler/-innengruppen (wie einzelne Jahrgänge) verpflichtend und für die anderen freiwillig und an offenen Ganztagschulen besteht Wahlfreiheit für alle Schüler/-innen. Definitivisch kann seit dem Schuljahr 2016/17 an offenen Ganztagschulen auch lediglich eine „Mitverantwortung der Schulleitung“ (ebd., S. 6) für das Ganztagsangebot bestehen, während die Hauptverantwortung bei einem außerschulischen Träger liegt. Hierbei muss ein gemeinsames pädagogisches Konzept verfolgt werden. Die KMK-Definition für Ganztagschulen gibt insgesamt eher zeitliche Rahmen und die Verfügbarkeit eines Mittagessens vor, als dass sie beispielsweise Inhalte, Zielgruppen oder Programmformate für den Ganztag festlegt. Während die Qualität regulärer Beschulung unter anderem durch Curricula und klar geregelte professionelle Verantwortlichkeiten gesichert wird, haben die Schulen, beziehungsweise außerschulischen Träger für die Ausgestaltung des Ganztagsbetriebs demnach größere Freiheitsgrade. Darüber hinaus gestalten die Bundesländer die KMK-Definition verschieden aus, beispielsweise in Hinblick auf die Verbindlichkeit der Teilnahme oder die Öffnungszeiten (KMK, 2015, 2018).

Neben der politischen Definition, die hinter dem IZBB-basierten Ganztagschulausbau stand, sind die von der Politik geäußerten Zielsetzungen und Wirkungserwartungen für die vorliegende Dissertation relevant. Der Ganztagschulausbau resultierte einerseits aus gestiegenen Betreuungsbedarfen. Mit dem Bereitstellen verlässlicher ganztägiger Betreuung sollte die Vereinbarkeit von Familie und Beruf verbessert und so die (Vollzeit-)Beteiligung von Eltern am Arbeitsmarkt gesteigert werden (vgl. BMBF, 2003a, 2003b; Ottweiler, 2005). Andererseits wurden auch Bildungsziele für die Begründung des Ganztagsausbaus genannt: Die KMK beschloss im Nachgang der Veröffentlichung der Ergebnisse der PISA-Studie (Programme for International Student Assessment) im Jahr 2001 als eines von sieben Handlungsfeldern „Maßnahmen zum Ausbau von schulischen und außerschulischen Ganztagsangeboten mit dem Ziel erweiterter Bildungs- und Fördermöglichkeiten,

insbesondere für Schülerinnen und Schüler mit Bildungsdefiziten und besonderen Begabungen“ (KMK, 2002, S. 7). Das IZBB war die größte dieser Maßnahmen und zielte auf verbesserte Bildungsergebnisse, insbesondere auch für bildungsbenachteiligte Gruppen ab, sodass Bildung und soziale Herkunft entkoppelt würden (vgl. auch Holtappels, 2006; Steiner, 2009; Züchner & Fischer, 2014). Laut BMBF (2003a) könnten diese pädagogischen Zielsetzungen an Ganztagschulen aufgrund erweiterter zeitlicher und personeller Ressourcen besser als an Halbtagschulen erreicht werden. So könnte mehr individuelle Förderung stattfinden und die Lernkultur durch vielfältigere Lernaktivitäten und die Verknüpfung von Unterricht und Zusatzangeboten erweitert werden. Außerdem könnte die Schule geöffnet, das soziale Lernen und die Partizipation von Schüler/-innen und Eltern gestärkt werden und mehr multiprofessionelle Expertise in die Förderung der Schüler/-innen eingehen. Zusammenfassend lässt sich sagen, dass die Politik hohe Erwartungen an den IZBB-basierten Ganztagschulausbau richtete, auch in Hinblick auf pädagogische Mehrwerte wie eine optimierte Kompetenzförderung und die Entkopplung von Leistungs- und Herkunftsmerkmalen. Auch nach Beendigung des IZBB verfolgten die Länder diese Zielsetzungen und strebten einen weiteren Ganztagsausbau an (KMK, 2015).

1.1.2 Pädagogische Wirkungsannahmen

Von den Wirkungserwartungen der politischen Akteure lässt sich die pädagogische Perspektive abgrenzen. Modelle, die die potenziellen Wirkungen der Ganztagschule auf Schüler/-innen abbilden, betonen die Bedeutung von Qualitätsmerkmalen auf verschiedenen Ebenen (Fischer & Klieme, 2013; Miller & Truong, 2009; Radisch, 2009). Neben allgemeinen Merkmalen effektiver Schulen wie ein positives Schulklima oder eine gute Personalentwicklung müssen auch auf den Ganztage bezogene Voraussetzungen an den Schulen gegeben sein. Hierzu zählen beispielsweise stabile externe Kooperationen und ein ausreichendes Platzangebot. Als Idealmodell werden hierbei häufig sogenannte rhythmisierte und vollgebundene Ganztagschulen diskutiert, die ein verpflichtendes Ganztagsprogramm mit dem regulären Unterricht über den gesamten Schultag alternieren. Schulmerkmale, sowie allgemeine externe Kontextfaktoren bilden zusammen mit den individuellen und familiären Voraussetzungen der Schüler/-innen die Grundlage für Wirkungen innerhalb der Schulen. In diesem Zusammenhang wird auf Prozessebene angenommen, dass Qualitätsmerkmale der Ganztagsangebote (z.B. hochwertiges pädagogisches Konzept, Leitung durch qualifiziertes Personal) mit Merkmalen der Teilnahme (z.B. dauerhaft, intensiv) interagieren und die

Effekte auf die Schüler/-innen bedingen. Die Modelle gehen davon aus, dass sich verschiedene Wirkungen auf Ebene der Schüler/-innen hierbei gegenseitig beeinflussen, wie zum Beispiel Kompetenzen, Leistungsmotivation oder Sozialverhalten. Sind solche Bedingungen auf den Ebenen der Schule, Angebote und Teilnahme gegeben, wird davon ausgegangen, dass Ganztagschulen gesteigerte Potenziale haben, Schüler/-innen zu fördern (vgl. Höhmann, Holtappels & Schnetzer, 2004; Holtappels, 2006; Stecher, Radisch, Fischer & Klieme, 2007). Betrachtet man Schüler/-innengruppen, die aufgrund ihres sozialen Hintergrunds ohne Ganztagsangebote wenig lernförderliche oder anregungsreiche Nachmittagsgestaltungen zur Verfügung haben, können besondere Mehrwerte erwartet werden. Auch Schüler/-innen, deren Muttersprache nicht Deutsch ist, können in Ganztagschulen potenziell intensiver gefördert werden (z.B. durch Förderunterricht) und verbringen durch die Ganztagsbeteiligung mehr Zeit in deutschsprachigen Kontexten (Bellin & Tamke, 2010; Steiner, 2009; Züchner & Fischer, 2014).

Vor dem Hintergrund der wissenschaftlichen Annahmen zu Qualitätsbedingungen der Wirksamkeit des Ganztagsmodells wird in der Literatur vielfach diskutiert, dass die KMK-Ganztagschuldefinition relevante pädagogische Merkmale nicht hinreichend abbilden würde (z.B. Böttcher, 2015; Holtappels, 2006; StEG-Konsortium, 2015; Wendt, Goy, Walzebug & Valtin, 2016). Da die IZBB-Förderung von Ganztagschulen jedoch auf dieser Definition basierte, befürchtete zum Beispiel Holtappels (2006) eine Zunahme niedrigqualitativer Ganztagschulen. Insgesamt lässt sich schlussfolgern, dass den politischen Annahmen einer Leistungssteuerung und eines Bildungsungleichheitsabbaus durch Ganztagschulen aus pädagogischer Perspektive nur eingeschränkt zugestimmt wird. Es wird davon ausgegangen, dass für derartige Effekte in der Praxis bestimmte Qualitätsmerkmale des Ganztagsbetriebs gegeben sein müssen.

1.1.3 Ausgestaltung von Ganztagschulen

Aufgrund der Gestaltungsspielräume, die die KMK-Ganztagschuldefinition lässt und der verschiedenen Regelungen in den Bundesländern haben sich sehr unterschiedliche Ganztagschulformen entwickelt. Während in 2002 31,9% der Ganztagschulen voll gebunden, 7,7% teilweise gebunden und 60,3% offen organisiert waren, haben insbesondere teilweise gebundene und offene Formen mit dem Ganztagsausbau zugenommen. Im Jahr 2016 waren nur noch 16,6% der Ganztagschulen voll gebunden, während 15,7% teilweise gebunden und 67,7% offen organisiert waren (KMK, 2008, 2018). An gebundenen

Ganztagschulen finden sich im Gegensatz zu offenen eher rhythmisierte Zeitstrukturen, bei denen die Ganztags- und regulären Unterrichtselemente über den gesamten Schultag alterniert werden und die als lernförderlich diskutiert werden. Diese Rhythmisierung ist jedoch insgesamt selten, wie beispielsweise die Studie zur Entwicklung von Ganztagschulen (StEG) belegt (Holtappels, 2008; StEG-Konsortium, 2013, 2015). Außerdem besteht an gebundenen Ganztagschulen häufiger ein schriftliches Ganztagschulkonzept und eine stärkere Verknüpfung der Ganztagsinhalte mit dem regulären Unterricht. An offenen Ganztagschulen ist das Nachmittagsprogramm häufig additiv zum Vormittagsunterricht angelegt und hiermit inhaltlich wenig verbunden. Der Ganztagsbetrieb offener Modelle wird öfter durch externe Träger gesteuert. Über die Ganztagschultypen hinweg wird das Ganztagsangebot selten an fünf Wochentagen angeboten und oftmals nicht durch Schulleitungen gesteuert (ebd.).

Betrachtet man neben Strukturmerkmalen die Ebene der Ganztagsangebote, dann haben Ganztagschulen flächendeckend umfangreiche Programme und Kooperationen mit externen Partnern aufgebaut. Die Ganztagsangebote reichen von lernförderlich ausgerichteten (z.B. Förderunterricht, fachliche Arbeitsgemeinschaften) bis hin zu weniger kernfachbezogenen Programmen (z.B. Kunsthandwerk, Sportclub), während erstere durchschnittlich seltener angeboten werden als letztere. Das Mittagessen, das an Ganztagschultagen bereitgestellt wird, muss typischerweise von den Eltern bezahlt werden, während die Ganztagsangebote kostenfrei sind (Arnoldt, 2011; Rollett, Lossen, Jarsinski, Lüpschen & Holtappels, 2011; StEG-Konsortium, 2013, 2015). Durchschnittlich arbeiten weniger als die Hälfte der Lehrkräfte selbst am Ganztage mit; der Ganztage wird überwiegend von (verschieden qualifiziertem) zusätzlichem Personal umgesetzt (Höhmman, Bergmann & Gebauer, 2008; StEG-Konsortium, 2013, 2015). Untersucht man außerdem die Teilnahmequoten am Ganztage beziehungsweise an einzelnen Programmen näher, dann sind diese insbesondere an offenen Ganztagschulen und an fachlich orientierten Programmarten teils gering ausgeprägt (Holtappels, 2008; StEG-Konsortium, 2016; Steiner, 2011).

Der Ganztagsbetrieb wird demnach sehr heterogen ausgestaltet. Diese Heterogenität hat vor dem Hintergrund pädagogischer Wirkmodelle (s. Abschnitt 1.1.2) auch eine Qualitätsdimension. So sind Merkmale wie die Rhythmisierung von Angeboten und Unterricht, eine hohe Qualifikation des Personals oder hohe Teilnahmequoten nicht flächendeckend gegeben. Vor diesem Hintergrund ist unklar, inwiefern sich Ganztagschulen wirklich von Halbtagschulen abgrenzen und sich empirische Ganztageeffekte auf Ebene der Schüler/-innen erwarten lassen.

1.2 Effekte auf Leistung und Bildungsungleichheit

Ausgehend von den politischen Annahmen, dass Ganztagschulen im Vergleich zu Halbtagschulen Lernprozesse besser fördern und zu einer Entkopplung von sozialem Hintergrund und schulischer Leistung beitragen könnten, liegt der Fokus der vorliegenden Dissertation auf der Untersuchung von Effekten ganztägiger Beschulung auf Leistungs- und Bildungsungleichheitsmaße. Der folgende Überblick empirischer Befunde bezieht sich demnach, ebenso wie die eigenen empirischen Analysen in den Einzelbeiträgen, auf zwei Sorten abhängiger Variablen. Einerseits werden die Kompetenzen von Schüler/-innen in Form von standardisiert gemessenen Testleistungen in verschiedenen Domänen untersucht (z.B. Testleistungen in Mathematik oder Lesen).¹ Andererseits werden soziale Ungleichheiten in Testleistungen als abhängiges Maß betrachtet. Soziale Ungleichheiten betreffen sowohl verschiedene soziale Gruppen (z.B. sozioökonomischer Hintergrund oder Migrationshintergrund) als auch verschiedene statistische Maße (z.B. sozialer Gradient, differenzielle Effekte für soziale Gruppen).² Der Literaturüberblick beginnt mit Effekten der Ganztagschule in Deutschland. Diese werden anschließend um Befunde internationaler Modelle sowie um allgemeine Erkenntnisse zu Effekten erweiterter Unterrichtszeit ergänzt.

1.2.1 Effekte der Ganztagschule in Deutschland

Studien, die die Effekte von Ganztagschulen der Primar- und Sekundarstufe I in Deutschland untersuchen, lassen sich in zwei Gruppen unterteilen: Eine Gruppe von Studien untersuchte Effekte ganztägiger Betreuung und Bildung auf Schulebene und die andere auf Individualebene. Aus der ersten Gruppe lassen sich zunächst Studien mit einem

¹ In den zitierten Vorarbeiten und eigenen Analysen werden verschiedene standardisierte Tests verwendet, die auf jeweils unterschiedliche Leistungs- beziehungsweise Kompetenzkonzepte zurückgehen. Diese sind in den Primärquellen nachzulesen; in der Dissertation werden allgemeine Begriffe der Leistung und Kompetenz verwendet. Beispielsweise Hartig und Klieme (2006) oder Klieme und Hartig (2008) ordnen solche Verständnisse umfassender ein.

² Eine vertiefte Auseinandersetzung mit Gerechtigkeitstheorien und ihren Auswirkungen auf das Verständnis von Bildungsungleichheit und der Operationalisierung entsprechender Maße findet sich beispielsweise bei Berkemeyer, Bos, Hermstein, Abendroth und Semper (2017) oder Manitiuis, Hermstein, Berkemeyer und Bos (2015).

querschnittlichen Design herausgreifen, die repräsentative Daten internationaler Schulleistungsstudien nutzten. Drei Studien (Holtappels, Radisch, Rollett & Kowoll, 2010; Radisch, Klieme & Bos, 2006; Willems, Wendt & Radisch, 2015) verglichen Leistungsniveaus an Ganztags- und Halbtagsgrundschulen und fanden entweder gleich hohe Kompetenzlevels in Lesen, Mathematik und Naturwissenschaft oder höhere Levels an Halbtagschulen. Zwei (Radisch et al., 2006; Willems et al., 2015) untersuchten außerdem Bildungsungleichheitsmaße als abhängige Variablen. Sie fanden gleich starke Zusammenhänge zwischen Leistung und sozialer Herkunft beziehungsweise Migrationshintergrund an Halb- und an Ganztagschulen. In der Sekundarstufe I fanden sich kaum Unterschiede im Niveau der naturwissenschaftlichen Kompetenz im Vergleich zwischen Schulen mit und ohne Nachmittagsangebot (Hertel, Klieme, Radisch & Steinert, 2008). Bei der Interpretation der Befunde dieser querschnittlichen Studien müssen Selektionseffekte berücksichtigt werden. Wenn beispielsweise besonders Schulen in sozioökonomisch schlechtgestellter Lage einen Ganztagsbetrieb anbieten, lassen sich beobachtete Differenzen zwischen Halb- und Ganztagschulen nicht kausal auf ihre Organisationsform zurückführen. Tatsächlich belegten die genannten Studien, dass das schulische Angebot nachmittäglicher Betreuung und Förderung mit einem tendenziell bildungsbenachteiligten Klientel zusammenhängt (Hertel et al., 2008; Holtappels et al., 2010; Radisch et al., 2006; Willems et al., 2015).

Zwei weitere der Studien, die Ganztageeffekte auf Schulebene untersuchten, umgingen konfundierende Selektionsmechanismen mit (quasi-)längsschnittlichen Designs. Eine Studie (Strietholt, Manitius, Berkemeyer & Bos, 2015) basierte zwar auf querschnittlichen repräsentativen Schulleistungsstudien aus dem Grundschul- und weiterführenden Bereich, nutzte jedoch Matchingverfahren um Effektverzerrungen zu vermeiden. Durch das Matching wurden Ganztagschulen jeweils mit solchen Halbtagschulen verglichen, die in Hinblick auf Merkmale wie unter anderem ihre soziale Komposition und Approximationen des früheren Leistungsniveaus möglichst ähnlich waren. Diese Studie fand ebenfalls keine Unterschiede zwischen Halb- und Ganztagschulen in Hinblick auf das Leistungsniveau in Mathematik, Lesen und Naturwissenschaften oder in Hinblick auf das Ausmaß sozialer Ungleichheit in diesen Leistungsmaßen. Eine andere Studie (Linberg, Struck & Bäumer, 2018) nutzte längsschnittliche repräsentative Daten und verglich die Leistungsentwicklungen in Lesen und Mathematik von der fünften bis zur siebten Klasse zwischen Schulen, die zum ersten

Messzeitpunkt Halbtagschulen, offene Ganztagschulen und gebundene Ganztagschulen waren. Diese Studie fand ebenfalls keine Leistungseffekte der Organisationsform.

Zusammenfassend lässt sich sagen, dass die Studien auf aggregierter Schulebene keine Ganztagschuleffekte in Hinblick auf Leistungs- und Ungleichheitsmaße nachwiesen. Solche aggregierten Maße spiegeln den allgemeinen Effekt aller Ganztags- versus Halbtagschulen wider. Gleichzeitig nehmen nicht alle Ganztagschüler/-innen auch an entsprechenden zusätzlichen Programmen teil, insbesondere nicht nur an solchen, die auf eine fachliche Leistungsförderlichkeit ausgelegt sind (Holtappels, Jarsinski & Rollett, 2011; StEG-Konsortium, 2016; Wendt et al., 2016). Daher ist denkbar, dass sich Effekte des Ganztags zwar nicht auf Aggregat-, jedoch auf Individualebene nachweisen lassen. Als zweite Gruppe empirischer Studien sind demnach solche Untersuchungen von Interesse, die am Ganztags beziehungsweise an spezifischen Ganztagsprogrammen teilnehmende Schüler/-innen mit nichtteilnehmenden verglichen.

Betrachtet man im Bereich der Studien zu Teilnahmeeffekten zuerst wieder Ergebnisse querschnittlicher Analysen repräsentativer Schulleistungsstudien, so zeigt sich, dass am Ganztags teilnehmende Schüler/-innen zumeist niedrigere Leistungsniveaus aufwiesen als nichtteilnehmende (Hertel et al., 2008; Holtappels et al., 2010). Auch auf Ebene der Schüler/-innen sind Selektionseffekte bei der Interpretation dieser querschnittlichen Befunde zu berücksichtigen. Es ist davon auszugehen, dass die negative Korrelation von Ganztagsbeteiligung und Leistungsniveau Mechanismen einer remedialen Förderung leistungsschwacher und bildungsbenachteiligter Gruppen durch Ganztagsangebote widerspiegelt; keine Effekte im Sinne von Ursache und Wirkung (ebd.).

Um kausale Effekte der Ganztagsbeteiligung auf Individualebene zu untersuchen verwendeten verschiedene Studien längsschnittliche Designs. Bellin und Tamke (2010) verglichen die Leseleistungsentwicklungen von Schüler/-innen offener Ganztagschulen in Berlin, die an dem Ganztagsangebot teilnahmen, mit dem Lesefortschritt nichtteilnehmender. Unter Kontrolle relevanter Variablen, wie der Vorleistung und der kognitiven Grundfähigkeit, zeigten sich keine Teilnahmeeffekte zwischen der ersten und zweiten Klasse. Außerdem wurden keine kompensatorischen Effekte für Kinder mit Migrationshintergrund gefunden. Die generelle Teilnahme am Ganztagsangebot bedeutet jedoch nicht, dass auch auf eine Leistungsförderung ausgelegte Angebote besucht werden. Daher sind insbesondere Studien von Interesse, die die Effekte spezifischer Ganztagsangebote untersuchten. Zwei Studien (Fischer, Sauerwein, Theis & Wolgast, 2016; Tillmann et al., 2018) verglichen die

Leseleistungsentwicklungen von Schüler/-innen, die an lesebezogenen Ganztagsangeboten teilnahmen, mit der von nichtteilnehmenden. Sowohl zwischen der dritten und vierten Klasse, als auch im Verlauf der fünften Klasse unterschieden sich die Leistungsfortschritte nicht zwischen den beiden Gruppen. Eine Studie (Linberg et al., 2018), die die Lese- und Mathematikentwicklungen zwischen fünfter und siebter Klasse anhand repräsentativer Daten untersuchte, fand nach der Kontrolle vielfältiger individueller (u.a. Vorleistung, Schulnoten) und schulischer Faktoren (u.a. Ganztagsorganisationsform, soziale Komposition) keine Effekte der Teilnahme an Hausaufgabenhilfe oder Förderunterricht in der sechsten Klasse. Eine weitere Studie (Lossen, Tillmann, Holtappels, Rollett & Hannemann, 2016) fand keinen Effekt des Besuchs naturwissenschaftsbezogener Ganztagsangebote auf Lernfortschritte in Sachunterrichtskompetenzen. Hierbei wurden an den Angeboten teilnehmende und nichtteilnehmende Schüler/-innen zwischen der dritten und vierten Jahrgangsstufe in zehn Bundesländern verglichen.

Die bisherigen Effektivitätsstudien zur Ganztagsschule in Deutschland liefern demnach keine Belege für die Annahme, dass ganztägige Beschulung mit Blick auf Leistungs- und Bildungsungleichheitsmaße einen Mehrwert gegenüber halbtägiger bilden würde. Die Befunde deuten weiterhin daraufhin, dass sich Ganz- und Halbtagsschulen, aber auch am Ganztage teilnehmende und nichtteilnehmende Schüler/-innen innerhalb von Ganztagsschulen, deutlich voneinander unterscheiden (s. auch Fischer & Klieme, 2013; StEG-Konsortium, 2015, 2016; Steiner, 2011). Je nach Angebotstyp bestehen Selbstselektionseffekte in verschiedene Richtungen, da die Schüler/-innen beziehungsweise Eltern über die Teilnahme selbst entscheiden. So werden Leseangebote im Grundschulbereich beispielsweise vorwiegend von Mädchen und leseleistungstarken Kindern besucht, während naturwissenschaftsbezogene Angebote eher von Jungen und bildungsbenachteiligten Kindern frequentiert werden (Fischer et al., 2016; Lossen et al., 2016). Solche Selektionseffekte müssen in der Untersuchung von Effekten der Angebotsteilnahme berücksichtigt werden.

In der Gesamtschau ist der Forschungsstand nicht umfangreich und in seiner Reichweite beschränkt. So liegen bei den genannten Studien, die Effekte der ganztägigen Organisation auf Schulebene untersuchen, nur zwei (quasi-)längsschnittliche Studien vor, die kausale Inferenzen erlauben (Linberg et al., 2018; Strietholt et al., 2015). Beide wiesen hierbei nur einmalige Messzeitpunkte der Organisationsform auf, betrachteten potenziell also auch Schulen, die ihren Ganztagsbetrieb gerade erst aufgenommen hatten als Ganztagsschulen und Schulen, die kurz davor standen einen Ganztagsbetrieb aufzunehmen, als

Halbtagschulen. Darüber hinaus lieferte nur die Studie von Strietholt und Kolleg/-innen (2015) auch Befunde zu Effekten auf Bildungsungleichheitsmaße. Die Befunde basierten hierbei auf repräsentativen Stichproben zwischen der vierten und zehnten Jahrgangsstufe und deckten die Kompetenzbereiche Mathematik, Lesen und Naturwissenschaften ab. Fünf der oben geschilderten Studien untersuchten Teilnahmeeffekte auf Individualebene längsschnittlich und ließen somit kausale Schlüsse zu (Bellin & Tamke, 2010; Fischer et al., 2016; Linberg et al., 2018; Lossen et al., 2016; Tillmann et al., 2018). Zusammenfassend betrachtet untersuchten diese Studien Effekte der Teilnahme an Ganztagsangeboten im Allgemeinen, an lese- und naturwissenschaftsbezogenen Angeboten, an Hausaufgabenhilfe sowie an Förderunterricht. Verglichen wurden an Angeboten teilnehmende und nichtteilnehmende Schüler/-innen innerhalb von Ganztagschulen oder an allen Schulformen. Untersuchte abhängige Maße umfassten über die Studien hinweg die Leistungsentwicklungen in Lesen, Mathematik und Naturwissenschaften; eine Studie untersuchte außerdem kompensatorische Effekte der Teilnahme für Schüler/-innen mit Migrationshintergrund. Die untersuchten Stichproben stammten aus der ersten bis siebten Jahrgangsstufe und nur Linberg und Kollegen (2018) nutzten einen für Deutschland repräsentativen Datensatz.

1.2.2 Effekte erweiterter Bildung und Betreuung im internationalen Raum

Im Gegensatz zu Ganztagschulen in Deutschland haben andere internationale Modelle erweiterter Bildung und Betreuung eine längere Tradition und sind weiter verbreitet als in Deutschland. Hierbei unterscheiden sich die Ausformungen zwischen den Ländern sehr deutlich (Organisation for Economic Co-operation and Development [OECD], 2011; Plantenga & Remery, 2013). Die Schulsysteme mancher Länder sind vollständig ganztägig organisiert, weswegen dort keine vergleichenden Evaluationsstudien vorliegen. Insbesondere aus Befunden aus Ländern, in denen ebenfalls Variation in Hinblick auf erweiterte Bildung Betreuung besteht, lassen sich Erkenntnisse für die Ganztagschule in Deutschland ableiten.

Am ehesten vergleichbar mit der Ganztagschule sind die sogenannten Tagesschulen in der deutschsprachigen Schweiz, die eine Alternative zu Halbtagsgrundschulen darstellen. Tagesschulen bieten überwiegend ein freiwillig zu wählendes Nachmittagsprogramm nach dem regulären Unterricht an, das nicht immer auf eine fachliche Förderung ausgerichtet ist, sondern typischerweise spiel- und freizeitbezogene Elemente enthält. In einer Evaluationsstudie zeigten intensiv an den Nachmittagsangeboten teilnehmende Kinder bessere Lese- und Mathematikentwicklungen zwischen erster und dritter Klasse, als

nichtteilnehmende. Gleichzeitig konnten keine kompensatorischen Effekte der Teilnahme auf soziale Bildungsungleichheiten in Lesen und Mathematik festgestellt werden (Schüpbach, 2014, 2015; Schüpbach, Herzog & Ignaczewska, 2013). In der Stichprobe war die Gruppe der intensiv Teilnehmenden sozioökonomisch privilegiert. Da dies in den Analysen nicht kontrolliert wurde, müssen die Ergebnisse mit Vorsicht interpretiert werden. Bei unabhängiger Untersuchung einer anderen Stichprobe zwischen erster und zweiter Klasse zeigte sich hingegen kein Effekt der dauerhaften Angebotsteilnahme auf die Mathematikleistung (Schüpbach, Nieuwenboom, Frei & Allmen, 2018). Es wurden ungleichheitsreduzierende Effekte der Teilnahme in Hinblick auf den sozialen Hintergrund der Schüler/-innen, aber ungleichheitssteigernde Effekte mit Blick auf den Migrationshintergrund gefunden. Die Befunde sind demnach nicht eindeutig.

Neben Tagesschulen in der Deutschschweiz gibt es sogenannte After-School-Programme in den USA. Unter After-School-Programmen versteht man eine Vielzahl verschiedener Angebote im Nachmittagsbereich, die typischerweise unabhängig von der Schule organisiert und in separaten und schulübergreifenden Einrichtungen durchgeführt werden. Die Angebote richten sich häufig an spezifische Gruppen (z.B. nur Jugendliche mit Anspruch auf finanzielle Unterstützung des Mittagessens) und basieren auf einer freiwilligen Teilnahme. Weitere Merkmale unterscheiden sich jedoch sehr deutlich, wie beispielsweise die Öffnungszeiten, Aktivitäten und Qualifikationen des Aufsichtspersonals. Aufgrund der Heterogenität der Programme und der umfangreichen Evaluationsforschungslage konnten verschiedene Metaanalysen und Literaturreviews nicht nur allgemeine Effekte der Teilnahme auf schulische Leistungen, sondern auch Mediatoren untersuchen. Diese Übersichtsarbeiten belegten, dass sich die Leistungen der Schüler/-innen durch die Teilnahme an After-School-Programmen durchschnittlich verbesserten (Apsler, 2009; Cooper, Charlton, Valentine, Muhlenbruck & Borman, 2000; Fashola, 1998; Feldman & Matjasko, 2005; Scott-Little, Hamann & Jurs, 2002). Dies galt insbesondere für bildungsbenachteiligte Schüler/-innengruppen. Es wurden höhere Effektstärken bei Programmen gefunden, die spezifische Ziele verfolgten (z.B. Leistungsförderung), eine enge Verbindung der Programm- und Lehrplaninhalte und geringe Arbeitsgruppengrößen aufwiesen, deren Teilnehmer/-innen regelmäßig und intensiv teilnahmen und die von gut qualifiziertem Fachpersonal geleitet wurden (ebd.). Bei Programmen, die eine niedrigere Qualität aufwiesen, also beispielsweise von pädagogischen Laien geleitet wurden, zeigten sich keine förderlichen Effekte der

Teilnahme (Apsler, 2009; Durlak, Weissberg & Pachan, 2010; Patall, Cooper & Allen, 2010; Roth, Malone & Brooks-Gunn, 2010).

Zusammenfassend scheinen insbesondere Programme leistungsförderlich zu wirken, die möglichst vergleichbar mit regulärem Unterricht sind (z.B. inhaltliche Lernorientierung, Verbindung zu Lehrplan, qualifizierte Instruktion, regelmäßige Teilnahme). Diese After-School-Programme wiesen einen förderlichen Effekt auf das Leistungsniveau auf, insbesondere für bildungsbenachteiligte Schüler/-innen. Angebote, bei denen die Betreuung der Schüler/-innen im Vordergrund stand, zeigten hingegen keine Effekte auf Leistungsmaße.

1.2.3 Effekte erweiterter Unterrichtszeit im internationalen Raum

Vor dem Hintergrund der vorausgehenden Schlussfolgerung, dass nur möglichst unterrichtsähnliche, nicht aber auf Betreuung ausgerichtete Nachmittagsangebote einen Effekt auf Schüler/-innenleistungen zu haben scheinen, sind zusätzlich Befunde zu Effekten regulärer Unterrichtszeit interessant. In Hinblick auf die jährliche Unterrichtszeit bestehen erhebliche Unterschiede zwischen den europäischen Ländern, sowohl in Hinblick auf die Anzahl der Schultage pro Jahr als auch die Unterrichtsstunden pro Woche und Fach (European Commission/EACEA/Eurydice, 2018). Im Folgenden werden Befunde von Studien zusammengefasst, die Zusammenhänge zwischen Unterrichtszeit und Leistungsbeziehungsweise Ungleichheitsmaßen untersuchten, indem sie international vergleichende Schulleistungsstudien analysierten. Vergleichbar mit Modellen erweiterter Bildung und Betreuung (s. Abschnitt 1.2.2) wird hierbei davon ausgegangen, dass ein Mehr an Unterrichtszeit grundsätzlich lernförderlich wirkt (vgl. Carroll, 1989; Scheerens, 2014). Für Schüler/-innen mit Migrationshintergrund und/oder einem wenig lernanregenden häuslichen Umfeld wird erwartet, dass sie von schulischem Unterricht stärker profitieren können, als privilegiertere Schüler/-innen, die ohnehin über bildungsförderliche Umgebungen verfügen (vgl. Boudon, 1974; Ditton, 2000; OECD, 2011; Strietholt et al., in Druck).

Eine frühe Studie im Bereich der Leistungseffekte untersuchte, wie Leistungsmaße international mit der allgemeinen beziehungsweise fachbezogenen Unterrichtszeit korrelierten (Baker, Fabrega, Galindo & Mishook, 2004). Die Studie fand auf Länderebene kaum Zusammenhänge zwischen der wöchentlichen Unterrichtszeit und dem Leistungsniveau der Länder. Auch wenn die Korrelation wöchentlicher Unterrichtszeit und der Leistung der Schüler/-innen innerhalb von Ländern untersucht wurde, schnitten Schüler/-innen mit mehr Unterrichtszeit nur geringfügig besser ab. Im Vergleich mit dieser nutzte eine spätere Studie

die methodisch robustere Identifikationsstrategie der Modellierung fixierter Effekte (Rivkin & Schiman, 2015). Es zeigte sich, dass ein Mehr an Unterrichtszeit innerhalb von Schulen mit einem höheren Leistungsniveau einherging. Dieser Mehrwert von Unterrichtszeit war geringer, wenn die Klassenumgebung problematisch war (z.B. viele Unterrichtsstörungen durch Schüler/-innen, Respektlosigkeit gegenüber Lehrkräften). Im Vergleich das stärkste Design für die Ableitung kausaler Effekte wies eine Studie von Lavy (2015) auf, die die Variation in der Fachleistung in Mathematik, Naturwissenschaften und Lesen innerhalb von Fünfzehnjährigen in einen Zusammenhang mit der individuellen Unterrichtszeit in diesen Fächern stellte und fixierte Effekte modellierte. Es fanden sich durchweg signifikante positive Effekte der wöchentlichen Unterrichtszeit auf die drei Leistungsmaße. Diese Studie zeigte weiterhin, dass die Effekte der Unterrichtszeit für Schüler/-innen mit weniger gebildeten Eltern und für Jugendliche mit Migrationshintergrund ausgeprägter waren. Dieser Befund weist darauf hin, dass ein Mehr an Unterrichtszeit tatsächlich zu einer Reduktion von Bildungsungleichheit beitragen kann. Diese Schlussfolgerung wird von einer Untersuchung auf Basis fixierter Effekte gestützt, bei der eine höhere durchschnittliche Unterrichtszeit pro Jahr mit einer besseren Integration von Schüler/-innen mit Migrationshintergrund einherging (Schneeweis, 2011). In Bezug auf den Zusammenhang zwischen elterlicher Bildung und der Leistungen der Schüler/-innen sind die Befunde zweier weiterer Studien jedoch heterogener. Burger (2016) fand, dass der Zusammenhang zwischen der elterlichen Bildung und der Leistung der fünfzehnjährigen Schüler/-innen in Ländern mit mehr Unterrichtszeit pro Jahr geringer war. Gleichzeitig zeigten sich widersprüchliche Befunde bei der Berücksichtigung der sozialen Schulkomposition. In einer anderen Untersuchung fand sich ein bildungsungleichheitsreduzierender Effekt der jährlichen Unterrichtszeit in westlichen Ländern, jedoch ein gegenteiliger Effekt in den postkommunistischen Ländern Europas (Schlicht, Stadelmann-Steffen & Freitag, 2010).

In der Zusammenschau weisen diese Befunde darauf hin, dass ein Mehr an regulärem Unterricht sowohl leistungssteigernde als auch bildungsungleichheitsreduzierende Effekte haben kann. Dies passt zu Befunden zu US-amerikanischen After-School-Programmen, die bei einer unterrichtsähnlichen Durchführung lernförderlicher wirkten (s. Abschnitt 1.2.2). In Bezug auf die Ganztagschule in Deutschland lässt dies vermuten, dass erweiterte zeitliche Spielräume, wie sie an der Ganztagschule bestehen, ein leistungssteigerndes und bildungsungleichheitsreduzierendes Potenzial haben können, wenn sie sich in ihrer Ausgestaltung an regulärem Unterricht orientieren (vgl. auch Abschnitt 1.1.2).

1.3 Forschungsanliegen und Forschungsfragen

Hinter dem Ganztagschulausbau in Deutschland standen unter anderem die politischen Annahmen, dass dieser zu einer Steigerung des Leistungsniveaus und zum Abbau von Bildungsungleichheiten beitragen könnte (s. Abschnitt 1.1.1). Die KMK-Ganztagsdefinition bildet jedoch eher strukturelle als pädagogische Merkmale ab und trägt so zu der sehr heterogenen Ausgestaltung des Ganztagskonzepts bei (s. Abschnitt 1.1.3). Pädagogische Wirksamkeitsmodelle gehen gleichzeitig davon aus, dass Wirkungen auf Merkmale von Schüler/-innen nur bei einer hohen Qualität des Ganztagsbetriebs zu erwarten sind (s. Abschnitt 1.1.2). Die politischen Erwartungen sind demnach umstritten; außerdem sind sie nur unzureichend empirisch geprüft. Einerseits deuten internationale Befunde darauf hin, dass Modelle erweiterter Bildung und Betreuung sowie mehr regulärer Unterricht lernförderlich wirken *können*, insbesondere für benachteiligte Gruppen (s. Abschnitte 1.2.2 und 1.2.3). Andererseits zeigten sich für Ganztagschulen in Deutschland bislang keine entsprechenden Effekte (s. Abschnitt 1.2.1). Die Dissertation knüpft an dieses Forschungsdesiderat an und untersucht die übergeordnete Frage, ob Ganztagschulen gegenüber Halbtagschulen einen Mehrwert für die Leistungsentwicklung von Schüler/-innen und den Abbau sozialer Bildungsungleichheiten bilden. Hierzu wird in separaten Beiträgen Folgendes untersucht:

- (1) Inwiefern stellen Ganztagschulen im Vergleich zu Halbtagschulen eine neue Organisationsform mit erweiterten Bildungspotenzialen dar und lassen demnach über halbtägige Beschulung hinausgehende Effekte erwarten?
- (2) Gab es zu Beginn des Ganztagschulausbaus einen Effekt des schulischen Anbietens nachmittäglicher Förderungsprogramme auf Leistungsentwicklungen von Schüler/-innen oder auf herkunftsbedingte Bildungsungleichheiten?
- (3) Gab es im späteren Verlauf des Ganztagschulausbaus einen Unterschied zwischen Halb- und Ganztagschulen in der Leistungsentwicklung von Schüler/-innen oder der sozialen Ungleichheit in diesen Leistungsentwicklungen?
- (4) Gab es Unterschiede in der Leistungsentwicklung von Schüler/-innen offener Ganztagschulen, die an lernförderlich ausgerichteten Ganztagsangeboten teilnahmen und solchen, die nicht daran teilnahmen?

Es folgt eine Auflistung der Einzelbeiträge (s. Abschnitt 2), die anschließend in einer Gesamtdiskussion zusammengefasst und hinsichtlich ihrer Reichweite und Implikationen eingeordnet werden (s. Abschnitt 3).

1.4 Literaturverzeichnis

- Apsler, R. (2009). After-school programs for adolescents: A review of evaluation research. *Adolescence, 44*, 1–19.
- Arnoldt, B. (2011). Kooperation zwischen Ganztagschule und außerschulischen Partnern: Entwicklung der Rahmenbedingungen. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG)* (S. 312–329). Weinheim, Basel: Beltz Juventa.
- Baker, D. P., Fabrega, R., Galindo, C., & Mishook, J. (2004). Instructional time and national achievement: Cross-national evidence. *Prospects, 34*(3), 311–334.
- Bargel, T., & Kuthe, M. (Hrsg.) (1991). *Ganztagschule. Untersuchungen zu Angebot und Nachfrage, Versorgung und Bedarf*. Bad Honnef: Bundesdruckerei.
- Bellin, N., & Tamke, F. (2010). Bessere Leistungen durch Teilnahme am offenen Ganztagsbetrieb? *Empirische Pädagogik, 24*(2), 93–112.
- Berkemeyer, N., Bos, W., Hermstein, B., Abendroth, S., & Semper, I. (2017). *Chancenspiegel – eine Zwischenbilanz: Zur Chancengerechtigkeit und Leistungsfähigkeit der deutschen Schulsysteme seit 2002*. Gütersloh: Verlag Bertelsmann Stiftung.
- BMBF (2003a). *Investitionsprogramm „Zukunft Bildung und Betreuung“: Ganztagschulen. Zeit für mehr*, Bundesministerium für Bildung und Forschung.
http://www.bmbf.de/pub/ganztagschulen-zeit_fuer_mehr.pdf.
- BMBF (2003b). *Verwaltungsvereinbarung Investitionsprogramm „Zukunft Bildung und Betreuung“ 2003-2007*.
https://www.ganztagschulen.org/_media/20030512_verwaltungsvereinbarung_zukunft_bildung_und_betreuung.pdf.
- BMBF (2005). *Ergänzende Information zur Verwaltungsvereinbarung Investitionsprogramm „Zukunft Bildung und Betreuung“: Kostenneutrale Verlängerung des Förderzeitraums*.
https://www.ganztagschulen.org/_media/izbb_ergaenzende_info.pdf.
- Böttcher, W. (2015). Ganze Tage in der Schule: Politik und Wissenschaft zwischen Anspruch und Wirklichkeit. In T. Hascher, T. S. Idel, S. Reh & W. Thole (Hrsg.), *Bildung über den ganzen Tag. Forschungs- und Theorieperspektiven der Erziehungswissenschaft* (S. 39–53). Opladen, Berlin, Toronto: Barbara Budrich Publishers.

- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in Western society*. New York: Wiley.
- Burger, K. (2016). Intergenerational transmission of education in Europe: Do more comprehensive education systems reduce social gradients in student achievement? *Research in Social Stratification and Mobility*, 44, 54–67.
- Carroll, J. B. (1989). The Carroll Model: A 25-year retrospective and prospective view. *Educational Researcher*, 18(1), 26–31.
- Cooper, H., Charlton, K., Valentine, J. C., Muhlenbruck, L., & Borman, G. D. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development*, 65(1), 1–127.
- Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. In A. Helmke, W. Hornstein & E. Terhart (Hrsg.), *Qualität und Qualitätssicherung im Bildungsbereich Schule, Sozialpädagogik, Hochschule* (S. 73–92). Weinheim: Beltz.
- Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American journal of community psychology*, 45(3-4), 294–309.
- European Commission/EACEA/Eurydice (2018). *Recommended annual instruction time in full-time compulsory education in Europe: 2017/18*. Eurydice – Facts and Figures. Brussels: European Commission.
- Fashola, O. S. (1998). *Review of extended-day and after-school programs and their effectiveness*. Report: Bd. 24. Baltimore, MD: Center for Research on the Education of Students Placed at Risk.
- Feldman, A. F., & Matjasko, J. L. (2005). The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. *Review of Educational Research*, 75(2), 159–210.
- Fischer, N., & Klieme, E. (2013). Quality and effectiveness of German all-day schools. Results of the study on the development of all-day schools in Germany. In J. Ecarius, E. Klieme, L. Stecher & J. Woods (Hrsg.), *Extended education – an international perspective. Proceedings of the international conference on extracurricular and out-of-school time educational research* (S. 27–52). Opladen, Berlin, Toronto: Barbara Budrich Publishers.

- Fischer, N., Sauerwein, M. N., Theis, D., & Wolgast, A. (2016). Vom Lesenlernen in der Ganztagschule: Leisten Ganztagsangebote einen Beitrag zur Leseförderung am Beginn der Sekundarstufe I? *Zeitschrift für Pädagogik*, 62(6), 780–796.
- Hartig, J., & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 128–143). Berlin: Springer.
- Hertel, S., Klieme, E., Radisch, F., & Steinert, B. (2008). Nachmittagsangebote im Sekundarbereich und ihre Nutzung durch die Schülerinnen und Schüler. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme et al. (Hrsg.), *PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (S. 297–318). Münster, New York, NY, München, Berlin: Waxmann.
- Höhm, K., Bergmann, K., & Gebauer, M. (2008). Das Personal. In H. G. Holtappels, E. Klieme, T. Rauschenbach & L. Stecher (Hrsg.), *Ganztagsschule in Deutschland. Ergebnisse der Ausgangserhebung der „Studie zur Entwicklung von Ganztagsschulen“ (StEG)* (2. korrigierte Auflage) (S. 77–85). Weinheim und München: Juventa.
- Höhm, K., Holtappels, H. G., & Schnetzer, T. (2004). Ganztagschule: Konzeptionen, Forschungsbefunde, aktuelle Entwicklungen. In H. G. Holtappels, K. Klemm, H. Pfeiffer, H.-G. Rolff & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung. Daten, Beispiele und Perspektiven* (S. 253–290). Weinheim: Juventa.
- Holtappels, H. G. (2006). Stichwort: Ganztagschule. *Zeitschrift für Erziehungswissenschaft*, 9(1), 5–29.
- Holtappels, H. G. (2008). Angebotsstruktur, Schülerteilnahme und Ausbaugrad ganztägiger Schulen. In H. G. Holtappels, E. Klieme, T. Rauschenbach & L. Stecher (Hrsg.), *Ganztagsschule in Deutschland. Ergebnisse der Ausgangserhebung der „Studie zur Entwicklung von Ganztagsschulen“ (StEG)* (2. korrigierte Auflage) (S. 186–206). Weinheim und München: Juventa.
- Holtappels, H. G., Jarsinski, S., & Rollett, W. (2011). Teilnahme als Qualitätsmerkmal für Ganztagschulen: Entwicklung von Schülerteilnahmequoten auf Schulebene. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagsschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagsschulen (StEG)* (S. 97–119). Weinheim, Basel: Beltz Juventa.

- Holtappels, H. G., Radisch, F., Rollett, W., & Kowoll, M. E. (2010). Bildungsangebot und Schülerkompetenzen in Ganztagsgrundschulen. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried & E.-M. Lankes (Hrsg.), *IGLU 2006 – die Grundschule auf dem Prüfstand* (S. 165–198). Münster: Waxmann.
- Klieme, E., & Hartig, J. (2008). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. *Zeitschrift für Erziehungswissenschaft, Sonderheft 8*, 11–29.
- KMK (2002). *PISA 2000 – Zentrale Handlungsfelder: Zusammenfassende Darstellung der laufenden und geplanten Maßnahmen in den Ländern (Stand 07.10.2002)*.
https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2002/2002_10_07-Pisa-2000-Zentrale-Handlungsfelder.pdf.
- KMK (2008). *Allgemein bildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: Statistik 2002 bis 2006*.
http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2008/2008_03_04-Allgem-Schulen-Ganztagsform-02-06.pdf.
- KMK (2015). *Ganztagsschulen in Deutschland: Bericht der Kultusministerkonferenz vom 03.12.2015*. Berlin, Bonn: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland.
- KMK (2018). *Allgemein bildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: Statistik 2012 bis 2016*.
https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/GTS_2016_Bericht.pdf.
- Lavy, V. (2015). Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries. *The Economic Journal*, 125(588), 397-424.
- Linberg, T., Struck, O., & Bäumer, T. (2018). Vorzug Ganztagschule? Zusammenhänge mit der Kompetenzentwicklung im Bereich Lesen und Mathematik. *Zeitschrift für Erziehungswissenschaft*, 14(2), 1–23.
- Lossen, K., Tillmann, K., Holtappels, H. G., Rollett, W., & Hannemann, J. (2016). Entwicklung der naturwissenschaftlichen Kompetenzen und des sachunterrichtsbezogenen Selbstkonzepts bei Schüler/-innen in Ganztagsgrundschulen: Ergebnisse der Längsschnittstudie StEG-P zu Effekten der

- Schülerteilnahme und der Angebotsqualität. *Zeitschrift für Pädagogik*, 62(6), 760–779.
- Manitius, V., Hermstein, B., Berkemeyer, N., & Bos, W. (Hrsg.) (2015). *Zur Gerechtigkeit von Schule: Theorien, Konzepte, Analysen*. Münster, New York: Waxmann.
- Miller, B. M., & Truong, K. A. (2009). The role of afterschool and summer in achievement. The untapped power of afterschool and summer to advance student achievement. *Zeitschrift für Pädagogik*, 55(54. Beiheft), 124–142.
- OECD (2011). *Quality time for students: Learning in and out of school*. https://www.oecd-ilibrary.org/quality-time-for-students-learning-in-and-out-of-school_5kmdfm0f17ln.pdf?itemId=%2Fcontent%2Fpublication%2F9789264087057-en&mimeType=pdf.
- Ottweiler, O. (2005). Die Positionen von Parteien, Verbänden und Kirchen zur Ganztagschule. In V. Ladenthin (Hrsg.), *Die Ganztagschule* (S. 177–198). Weinheim: Juventa.
- Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the school day or school year: A systematic review of research (1985-2009). *Review of Educational Research*, 80(3), 401–436.
- Plantenga, J., & Remery, C. (2013). *Childcare services for school age children: A comparative review of 33 countries*. Luxembourg: Publications Office of the European Union.
- Radisch, F. (2009). *Qualität und Wirkung ganztägiger Schulorganisation: Theoretische und empirische Befunde*. Studien zur ganztägigen Bildung. Weinheim, München: Juventa.
- Radisch, F., Klieme, E., & Bos, W. (2006). Gestaltungsmerkmale und Effekte ganztägiger Angebote im Grundschulbereich. *Zeitschrift für Erziehungswissenschaft*, 9(1), 30–50.
- Rivkin, S. G., & Schiman, J. C. (2015). Instruction time, classroom quality, and academic achievement. *The Economic Journal*, 125(588), 425–448.
- Rollett, W., Lossen, K., Jarsinski, S., Lüpschen, N., & Holtappels, H. G. (2011). Außerunterrichtliche Angebotsstruktur an Ganztagschulen: Entwicklungstrends und Entwicklungsbedingungen. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG)* (S. 76–96). Weinheim, Basel: Beltz Juventa.

- Roth, J. L., Malone, L. M., & Brooks-Gunn, J. (2010). Does the amount of participation in afterschool programs relate to developmental outcomes? A review of the literature. *American journal of community psychology*, 45(3-4), 310–324.
- Scheerens, J. (Hrsg.) (2014). *Effectiveness of time investments in education*. SpringerBriefs in Education. Cham: Springer International Publishing.
- Schlicht, R., Stadelmann-Steffen, I., & Freitag, M. (2010). Educational inequality in the EU. *European Union Politics*, 11(1), 29–59.
- Schneeweis, N. (2011). Educational institutions and the integration of migrants. *Journal of Population Economics*, 24(4), 1281–1308.
- Schüpbach, M. (2014). Extended education and social inequality in Switzerland: Compensatory effects? *Journal for Educational Research Online*, 6(3), 95–114.
- Schüpbach, M. (2015). Effects of extracurricular activities and their quality on primary school-age students' achievement in mathematics in Switzerland. *School Effectiveness and School Improvement*, 26(2), 279–295.
- Schüpbach, M., Herzog, W., & Ignaczewska, J. (2013). Entwicklung der Mathematikleistung von Ganztagsschulkindern: kompensatorische Wirkung der Ganztagsschule? *Zeitschrift für Pädagogische Psychologie*, 27(3), 157–167.
- Schüpbach, M., Nieuwenboom, W., Frei, L., & Allmen, B. von (2018). Offene Tagesschulen als Mittel um Mathematikleistungen zu fördern und sozial oder kulturell bedingte Disparitäten zu verringern? *Journal for Educational Research Online*, 10(2), 93–116.
- Scott-Little, C., Hamann, M. S., & Jurs, S. G. (2002). Evaluations of after-school programs: A meta-evaluation of methodologies and narrative synthesis of findings. *American Journal of Evaluation*, 23(4), 387–419.
- Stecher, L., Radisch, F., Fischer, N., & Klieme, E. (2007). Bildungsqualität außerunterrichtlicher Angebote in der Ganztagsschule. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 27(4), 346–366.
- StEG-Konsortium (2013). *Ganztagsschule 2012/2013: Deskriptive Befunde einer bundesweiten Befragung*. http://www.projekt-steg.de/sites/default/files/Bundesbericht_Schulleiterbefragung_2012_13.pdf.
- StEG-Konsortium (2015). *Ganztagsschule 2014/2015: Deskriptive Befunde einer bundesweiten Befragung*. http://www.projekt-steg.de/sites/default/files/StEG_Bundesbericht%202015_final_0.pdf.

- StEG-Konsortium (2016). *Ganztagsschule: Bildungsqualität und Wirkungen außerunterrichtlicher Angebote: Ergebnisse der Studie zur Entwicklung von Ganztagsschulen 2012-2015*. http://www.projekt-steg.de/sites/default/files/StEG_Brosch_FINAL.pdf.
- Steiner, C. (2009). Mehr Chancengleichheit durch die Ganztagsschule? *Zeitschrift für Pädagogik*, 55(54. Beiheft), 81–105.
- Steiner, C. (2011). Teilnahme am Ganztagsbetrieb: Zeitliche Entwicklung und mögliche Selektionseffekte. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagsschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagsschulen (StEG)* (S. 57–75). Weinheim, Basel: Beltz Juventa.
- Strietholt, R., Gustafsson, J.-E., Hogrebe, N., Rolfe, V., Rosén, M., Steinmann, I. et al. (in Druck). The impact of education policies on socioeconomic inequality in student achievement: A review of comparative studies. In L. Volante, S. Schnepf, J. Jerrim & D. Klinger (Hrsg.), *Socioeconomic inequality and student outcomes: National trends, policies, and practices*. New York, NY: Springer.
- Strietholt, R., Manitius, V., Berkemeyer, N., & Bos, W. (2015). Bildung und Bildungsungleichheit an Halb- und Ganztagsschulen. *Zeitschrift für Erziehungswissenschaft*, 18(4), 737–761.
- Tillmann, K., Sauerwein, M., Hannemann, J., Decristan, J., Lossen, K., & Holtappels, H. G. (2018). Förderung der Lesekompetenz durch Teilnahme an Ganztagsangeboten? – Ergebnisse der Studie zur Entwicklung von Ganztagsschulen (StEG). In M. Schüpbach, L. Frei & W. Nieuwenboom (Hrsg.), *Tagesschulen. Ein Überblick* (S. 289–307). Wiesbaden: Springer Fachmedien.
- Wendt, H., Goy, M., Walzebug, A., & Valtin, R. (2016). Bildungsangebote an Ganz- und Halbtagsgrundschulen in Deutschland. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 225–246). Münster, New York: Waxmann Verlag.
- Willems, A. S., Wendt, H., & Radisch, F. (2015). Domänenspezifische Kompetenzen und Chancengerechtigkeit im Vergleich von Ganz- und Halbtagsgrundschultypen: Zur Rolle individueller Herkunftsmerkmale und der Komposition der Schülerschaft. In H. Wendt, T. C. Stubbe, K. Schwippert & W. Bos (Hrsg.), *10 Jahre international*

vergleichende Schulleistungsforschung in der Grundschule. Vertiefende Analysen zu IGLU und TIMSS 2001 bis 2011 (S. 219–238). Münster, New York: Waxmann.

Züchner, I., & Fischer, N. (2014). Kompensatorische Wirkungen von Ganztagschulen – Ist die Ganztagschule ein Instrument zur Entkopplung des Zusammenhangs von sozialer Herkunft und Bildungserfolg? *Zeitschrift für Erziehungswissenschaft*, 17, 349–367.

2 Beiträge der Arbeit

2.1 Beitrag I: Ganztagschulen als neue Organisationsformen – Entwicklungen und Evaluation

Isa Steinmann

Die im Folgenden abgedruckte Version des Manuskripts ist das Preprint des in der Enzyklopädie Erziehungswissenschaft Online erschienenen Beitrags.

Abstract

Im vorliegenden Handbuchartikel werden Ganztagschulen zunächst mit Blick auf ihre Bedeutung im Schulsystem in Deutschland verortet, in dem sie sich in den letzten Jahrzehnten von der Ausnahme zum Regelfall entwickelt haben. Anschließend wird zusammengefasst, welche politischen Reformen und Motivlagen historisch zu diesem massiven Ganztagschulenausbau beigetragen haben. Ausgehend von den damit verbundenen Wirkungsannahmen wird als Kern des Beitrags untersucht, wie sich Ganztags in der Praxis auswirkt. Es werden Befunde dazu zusammengeführt, wie Ganztagschulen organisiert und ausgestaltet werden und welche Konsequenzen für die Eltern, das Familienleben und die Schüler/-innen bestehen, auch im Vergleich mit Halbtagschulen. Auf Ebene der Schüler/-innen werden hierbei Merkmale der Teilnahme und Effekte auf die Individuen differenziert. In einem Fazit wird diskutiert, inwiefern sich Ganztagschulen vor dem Hintergrund dieser Befunde als Organisationsformen von Halbtagschulen abgrenzen lassen.

Ganztagschulen als neue Organisationsformen – Entwicklungen und Evaluation

1. Bedeutung des Ganztags im Schulsystem

Im deutschen Schulsystem finden sich Hort- und Ganztagsschulangebote, die den schulischen Vormittagsunterricht um institutionalisierte Betreuungs- und Bildungszeit am Nachmittag erweitern. Während sich das traditionelle Halbtagsschulsystem in Deutschland erst in diesem Jahrhundert zunehmend in ein Ganztagsschulsystem wandelt, sind halbtägige Systeme im internationalen Vergleich mittlerweile eher die Ausnahme. In den skandinavischen Ländern sind die Kommunen beispielsweise seit langem verpflichtet, nachmittägliche Bildungs- und Betreuungsoptionen zu gewährleisten. Generell ist in vielen europäischen Ländern ein Trend zu einer immer flächendeckenderen Versorgung mit Ganztagsbetreuungsangeboten erkennbar. Die Länder begründen Ausbauinvestitionen typischerweise mit steigenden Betreuungsbedarfen durch veränderte Familien- und Erwerbsstrukturen und mit pädagogischen Zielsetzungen. International unterscheiden sich die Definitionen und Ausgestaltungsformen erweiterter Bildung und Betreuung jedoch sehr (Allemann-Ghionda 2009; OECD 2011; Plantenga/Remery 2013).

In Deutschland definiert die Kultusministerkonferenz (KMK) Ganztagschulen bundesländerübergreifend als Schulen, „bei denen im Primar- und Sekundarbereich I

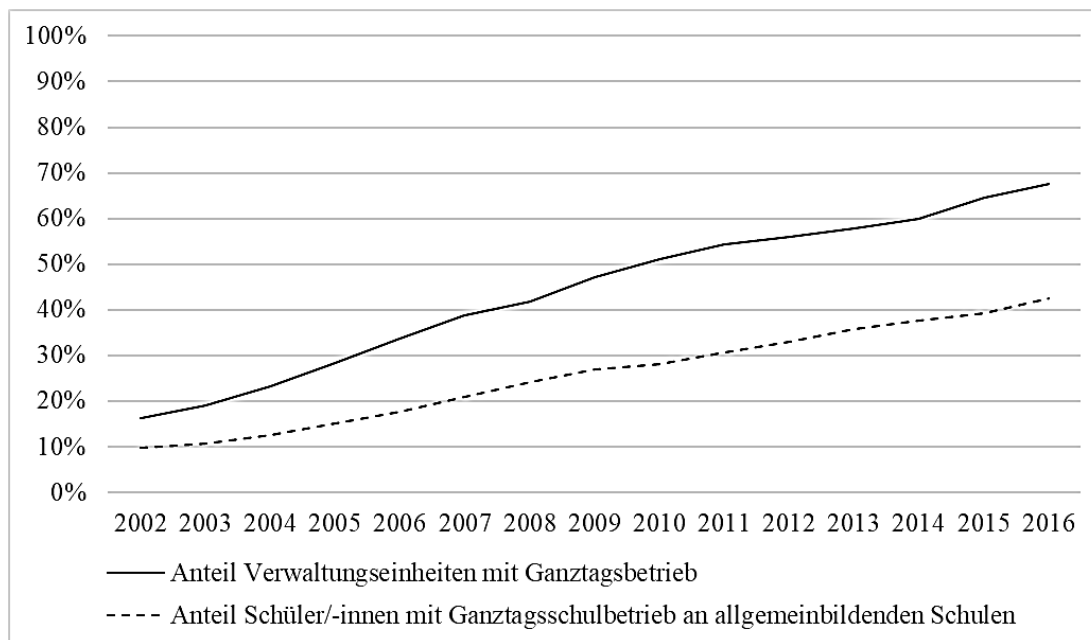
- an mindestens drei Tagen in der Woche ein ganztägiges Angebot für die Schülerinnen und Schüler bereitgestellt wird, das täglich mindestens sieben Zeitstunden umfasst,
- an allen Tagen des Ganztagsschulbetriebs den teilnehmenden Schülerinnen und Schülern ein Mittagessen bereitgestellt wird;
- die Ganztagsangebote unter der Aufsicht und Verantwortung der Schulleitung organisiert und in enger Kooperation mit der Schulleitung durchgeführt werden sowie in einem konzeptionellen Zusammenhang mit dem Unterricht stehen“ (2018, S. 4 f.).

Weiterhin werden bei der KMK drei Formen unterschieden: In 2016 waren etwa zwei Drittel aller Ganztagschulen sogenannte offene Ganztagschulen, ein Sechstel teilweise gebundene und ein Sechstel voll gebundene Ganztagschulen (ebd.). An offenen Ganztagschulen entscheiden die Schüler/-innen beziehungsweise Eltern für mindestens ein Schulhalbjahr, ob, und wenn ja, welches Ganztagsangebot besucht wird. An teilweise gebundenen

Ganztagsschulen nimmt ein Teil der Schülerschaft (z.B. nur die Orientierungsstufe) an mindestens drei siebenstündigen (oder längeren) Tagen teil. An voll gebundenen Ganztagsschulen gilt dies für die gesamte Schülerschaft. Seit dem Schuljahr 2016/17 fallen auch offene Ganztagsangebote in die KMK-Definition, bei denen „die Schulleitung auf der Basis eines gemeinsamen pädagogischen Konzeptes mit einem außerschulischen Träger kooperiert und eine Mitverantwortung der Schulleitung für das Angebot besteht“ (ebd., S. 6). Die Ganztagsschuldefinition umfasst nun also zum Beispiel auch Hortangebote, bei denen die Hauptverantwortung und Organisationsarbeit bei einem außerschulischen Träger wie der Jugendhilfe liegt.

Diese Definition liegt einer Statistik der KMK zugrunde, der zufolge die Anteile an Ganztagsschulen und an Schüler/-innen im Ganztagsbetrieb in diesem Jahrhundert drastisch gestiegen sind (vgl. Abb. 1). Mittlerweile sind demnach zwei von drei Schulen Ganztagsschulen. Da insbesondere der Anteil an offenen und nicht an gebundenen Ganztagsschulen gestiegen ist, war die absolute Zunahme an Schüler/-innen im Ganztagsbetrieb nicht so ausgeprägt, wie die der Schulen; 43% aller Schüler/-innen allgemeinbildender Schulen besuchten in 2016 ein Ganztagsangebot (ebd.).

Abbildung 1: Prozentuale Anteile an schulischen Verwaltungseinheiten in öffentlicher und privater Trägerschaft und an Schüler/-innen mit Ganztagsbetrieb zwischen 2002 und 2016³



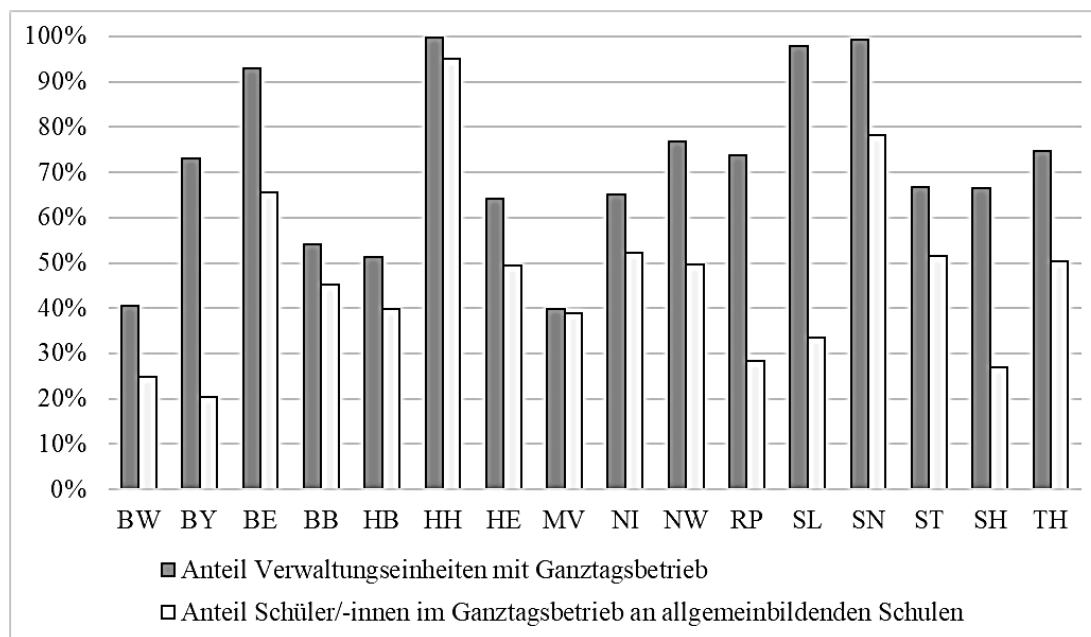
(Quelle: eigene Darstellung, basierend auf KMK 2008; KMK 2013; KMK 2018)

Diese Übersicht illustriert, dass die Bedeutung der Ganztagschule für das Schulsystem aus quantitativer Sicht stark zugenommen hat. Aufgrund der föderalen Struktur des Bildungssystems lohnt sich darüber hinaus ein Blick in die Länder (vgl. Abb. 2). Hierbei zeigt sich, dass sich der Ganztagsausbaustand deutlich unterscheidet. Während beispielsweise Hamburg eine nahezu vollständige Ganztagschullandschaft aufgebaut hat, sind in Mecklenburg-Vorpommern nur knapp 40% aller Schulen ganztägig organisiert. Die Anteile an Schüler/-innen im Ganztagsbetrieb variieren mit 20% bis 95% noch deutlicher zwischen den Ländern. Darüber hinaus illustriert die Abbildung, dass es in manchen Bundesländern wenig teilnehmende Schüler/-innen im Verhältnis zum Ganztagschulausbaustand gibt (z.B. Saarland, Bayern). In anderen Bundesländern ist der Anteil an Schulen und der Anteil an Schüler/-innen im Ganztagsbetrieb hingegen eher ausgeglichen (z.B. Mecklenburg-Vorpommern, Hamburg). Dies liegt darin begründet, dass die Bundesländer offene und

³ In dieser Statistik werden Ganztagsangebote unter nicht-schulischer Verantwortung, beispielsweise Horte, die von der Jugendhilfe ausgerichtet werden, für die Jahre 2002-2015 nicht erfasst. Ab dem Schuljahr 2016/17 werden diese miterfasst. Einrichtungen mit verschiedenen Schularten (z.B. Schulzentrum mit Haupt- und Realschulzweig) werden als eine schulische Verwaltungseinheit gezählt.

gebundene Formen des Ganztags verschieden akzentuiert einsetzen und unterschiedliche Regularien etabliert haben (KMK 2015). Darüber hinaus setzen vor allem die neuen Bundesländer traditionell nicht nur auf Ganztags, sondern auch auf Hortangebote der Kinder- und Jugendhilfe (Autorengruppe Bildungsberichterstattung 2016).

Abbildung 2: Prozentuale Anteile an schulischen Verwaltungseinheiten in öffentlicher Trägerschaft und an Schüler/-innen mit Ganztagsbetrieb in den Bundesländern im Jahr 2016



(Quelle: eigene Darstellung, basierend auf KMK 2018)

Obwohl die Bundesländer Ganztags verschieden stark ausbauen und ausgestalten, wird dem Thema flächendeckend eine große Bedeutung zugeschrieben, wie aus einer Länderbefragung der KMK hervorgeht. Nordrhein-Westfalen beispielsweise setzt sich den „Ausbau von Ganztagschulen und außerunterrichtlichen Ganztags- und Betreuungsangeboten zu einem attraktiven, qualitativ hochwertigen und umfassenden örtlichen Bildungs-, Erziehungs- und Betreuungsangebot, das sich an dem jeweiligen Bedarf der Kinder und Jugendlichen sowie der Eltern orientiert“ (KMK 2015, S. 20) zum erklärten Ziel. Über die Länder hinweg werden verschiedene „bildungs-, familien-, frauen-, sozial- und arbeitsmarktpolitische“ Motive dafür genannt, dass eine Weiterentwicklung und ein weiterer Ausbau der Ganztagsbeschulung flächendeckend angestrebt wird (ebd., S. 4).

Somit lässt sich zunächst aufgrund der starken Verbreitung und hohen bildungspolitischen Relevanz festhalten, dass Ganztagschulen aus der deutschen Schullandschaft nicht mehr wegzudenken sind. Im vorliegenden Beitrag wird hiervon

ausgehend zunächst beleuchtet, welche politischen Entwicklungen und Zielsetzungen hinter dem massiven Ganztagsausbau standen. Im Zusammenspiel mit den an Ganzttag gerichteten Wirkungserwartungen wird anschließend beleuchtet, welche Konsequenzen Ganzttag auf den Ebenen der Schulen, Familien und Schüler/-innen im Vergleich zu Halbtagschulen hat. Darauf aufbauend wird abschließend diskutiert, ob und wie sich Ganzttagsschulen faktisch als Organisationsformen von Halbtagschulen abgrenzen lassen.

2. Zielsetzungen und Entwicklung der Ganztagschule

Die Entwicklung der Ganztagschule in Deutschland und deren zugrundeliegenden Zielsetzungen und Motivlagen werden im Folgenden entlang zentraler Reformprogramme nachgezeichnet: Resümiert werden erstens die Ganztagsreformen der 1960er bzw. 1970er Jahre in der alten Bundesrepublik Deutschland (BRD) und der Deutschen Demokratischen Republik (DDR) und zweitens das Investitionsprogramm Zukunft Bildung Betreuung (IZBB) der frühen 2000er Jahre. Drittens werden angekündigte Neuerungen durch den Koalitionsvertrag der Regierungsparteien im Jahr 2018 benannt, der eine weitere Reform verspricht.

2.1 Ganztagsprogramme in BRD und DDR

Zum Zeitpunkt der Gründung der BRD im Jahr 1949 trugen erhebliche personelle, räumlich-infrastrukturelle und finanzielle Engpässe dazu bei, dass das mehrgliedrige Halbtagschulsystem des 19. Jahrhunderts übernommen und kein Ganztagsystem etabliert wurde.⁴ Schulen waren zum überwiegenden Teil nur bis mittags geöffnet, sodass das Mittagessen und die Nachmittagsbetreuung der Kinder und Jugendlichen weitestgehend privat organisiert werden mussten. Institutionalisierte Betreuungsmöglichkeiten wie in Horten waren rar (Hagemann 2009; Holtappels 2006; Ipfling/Lorenz 1979).

Ende der 1960er Jahre gewann ganztägige Beschulung in der BRD an politischer Bedeutung. Die Gründe für Reformbestrebungen lagen in postulierten Missständen des Schulsystems, das insbesondere im internationalen Vergleich als nicht zukunftsfruchtig erachtet wurde. Beispielsweise wurde die geringe Anzahl an Abiturient/-innen, aber auch die soziale Ungleichheit in Bildungschancen problematisiert. Gleichzeitig sollte das Problem

⁴ Ein Überblick der Ganztagschulgeschichte vor 1949 findet sich zum Beispiel bei Hagemann (2009) oder Ipfling und Lorenz (1979).

fehlender nachmittäglicher Betreuung sogenannter Schlüsselkinder angegangen werden, die nach der Schule nicht durch Erwachsene betreut wurden (Dahrendorf 1966; Picht 1964). Im Jahr 1968 empfahl der Deutsche Bildungsrat das erste systematische Ganztagsmodellprogramm. Zu diesem Zeitpunkt gab es in Deutschland etwa 40 Ganztagsversuchsschulen, die auf reformpädagogischer Initiative einzelner Personengruppen gegründet und nicht wissenschaftlich begleitet wurden. Die Ganztagschulen des anvisierten systematischen und wissenschaftlich begleiteten Modellversuchs sollten laut Deutschem Bildungsrat verschiedene Zielsetzungen verfolgen (1968): Im Gegensatz zu Halbtagschulen sollten Ganztagschulen besser auf ein komplexer werdendes Leben mit neuen Kommunikationsanforderungen und kooperativen Arbeitsformen vorbereiten und soziale Ungleichheiten in Bildungschancen reduzieren. Darüber hinaus sollten qualitativ hochwertigere Betreuungsoptionen im Nachmittagsbereich bereitstehen und der Wegfall des sechsten Wochenschultags durch die größeren zeitlichen Kapazitäten besser kompensiert werden. Die Organisationsform der Ganztagschule grenzte sich laut Bildungsrat insbesondere dadurch von Halbtagschulen ab, dass Lern- und Entspannungsphasen über den Schultag lernpsychologisch sinnvoll abgewechselt, neben Lehrkräften vermehrt auch andere Berufsgruppen involviert (z.B. Erzieher/-innen, Schulpsycholog/-innen) und ein Mittagessen bereitgestellt würde.

Im Jahr 1969 empfahl der Deutsche Bildungsrat zusätzlich das Erproben von Gesamtschulen, die das Ganztagschulkonzept erweiterten. Integrierte Gesamtschulen sollten Ganztagschulen sein, die außerdem eine gemeinsame Beschulung aller Fähigkeitsgruppen in der Mittelstufe sowie eine berufsorientierte Oberstufe mit dem Abitur als Schulabschluss boten (Deutscher Bildungsrat 1969). Obwohl es auch außerhalb der Gesamtschulprogrammatisierung Ganztagschulen gab, wurden die beiden Reformlinien im öffentlichen und wissenschaftlichen Diskurs fast synonym diskutiert. Der sogenannte Bildungsgesamtplan der Bund-Länder Kommission für Bildungsplanung erneuerte 1973 die Forderung nach einem flächendeckenden Ausbau ganztägiger Beschulung (BLK 1973). Trotz dieser Reformbestrebungen blieb die Realität hinter diesen Zielsetzungen zurück. Die zentralsten Gründe dafür, dass der Ganztags in der BRD eine Ausnahme blieb, waren der finanzielle Mehraufwand, Ressentiments gegen längere institutionalisierte Betreuungszeiten aus Politik und Bevölkerung, aber auch die ambivalenten Befunde zu Effekten der Versuchsschulen (Bargel 1991; Ipfling/Lorenz 1979).

Im Vergleich zur BRD gewann der Ganztags in der DDR früher, nämlich bereits Ende der 1950er Jahre an Bedeutung. Der Steigerung der Erwerbstätigkeit von Müttern, einem Kernpunkt der politischen Agenda, standen insbesondere für Grundschulkindern nur sehr begrenzte öffentliche Nachmittagsbetreuungsmöglichkeiten entgegen. Neben Mehrwerten für die Berufstätigkeit von Müttern versprach man sich auch Vorteile für Lernerfolge insbesondere von benachteiligten Schüler/-innen. Darüber hinaus sollte durch das Nachmittagsprogramm die Möglichkeit der moralischen und politischen Erziehung von Schüler/-innen entsprechend der Leitideen der DDR ausgeweitet werden. Es wurden zum einen sogenannte Tagesschulen etabliert und zum anderen Hortangebote an Halbtagschulen massiv ausgebaut, die unter schulischer Trägerschaft standen. Im Jahr 1960 wurde das Ziel ausgesprochen, bis 1975 alle Schulen schrittweise in Tagesschulen umzustrukturieren, langfristig sogar in Internate. Diese Zielsetzung wurde in den kommenden Jahren jedoch aufgrund des zu hohen finanziellen Aufwands und Vorbehalten aus der Bevölkerung reduziert. Es setzte sich in der Folge ein nahezu flächendeckendes Hortangebot mit vereinzelt Tagesschulen durch. 1989 hatten vier Fünftel der Schüler/-innen zwischen sechs und vierzehn Jahren einen Ganztagsplatz (Mattes 2009). Die gesellschaftliche Akzeptanz der ganztägigen Betreuung nahm hierbei nur langsam zu. Teile der Bevölkerung befürchteten, dass sich durch den Ganztagsausbau der Einfluss der Eltern aber auch der Kirche reduzieren und dass die Kinder überfordert werden könnten. Auch von Seiten der Lehrkräfte gab es Ressentiments (DDR-bezogene Überblicke bei Gebhardt 1993; Hagemann 2009; Mattes 2009).

Nach der Wiedervereinigung wurde das Schulsystem der BRD inklusive der primär halbtägigen Schulorganisation in den neuen Bundesländern etabliert, wo jedoch ein umfangreicheres Hortangebot unter verschiedenen Trägerschaften erhalten blieb. Bis zur Jahrhundertwende änderte sich an der halbtägigen Ausrichtung des Schulsystems der neuen BRD wenig.

2.2 Investitionsprogramm Zukunft Bildung und Betreuung

Im Jahr 2001 sorgte die Veröffentlichung der Ergebnisse der PISA-Studie (Programme for International Student Assessment) für breite Diskussionen um die Qualität des deutschen Bildungssystems. Im internationalen Vergleich schnitten die fünfzehnjährigen Schüler/-innen aus Deutschland sowohl in Lesen und Mathematik als auch den Naturwissenschaften deutlich unter dem Durchschnitt der teilnehmenden OECD-Länder ab und zeigten außerdem große

Leistungsstreuungen und erhebliche soziale Bildungsungleichheiten (KMK 2002). Die KMK identifizierte in der Folge dieser PISA-Befunde sieben Strategien, mit denen diese Problembereiche angegangen werden sollten. Eine betraf „Maßnahmen zum Ausbau von schulischen und außerschulischen Ganztagsangeboten mit dem Ziel erweiterter Bildungs- und Fördermöglichkeiten, insbesondere für Schülerinnen und Schüler mit Bildungsdefiziten und besonderen Begabungen“ (ebd., S. 7). Während eine allgemeine Steigerung der Leistung und der Förderung von Bildungseliten anvisiert wurde, lag ein weiterer Fokus auf der Förderung von Schüler/-innen mit sozial benachteiligten Hintergründen und aus dem unteren Leistungsbereich. Die Integration und gezielte Förderung dieser Schüler/-innen gehörte zu den Kernzielen des Ganztagsausbaus. Ganztagschulen sollten außerunterrichtliche Bildungsgelegenheiten für Kinder aus finanziell schlechter gestellten Familien öffnen, die ihnen ansonsten vorenthalten blieben (Holtappels 2006; Steiner 2009). Neben solchen bildungs- und sozialpolitischen Zielen war für den Ganztagschulausbau auch die Betreuungsdimension weiterhin von großer Bedeutung (KMK 2002; KMK 2015). Beispielsweise steigen die (Teilzeit-)Erwerbstätigkeit von Frauen und der Anteil an Ein-Eltern-Familien seit Jahrzehnten an (BPB 2016). Die politischen Parteien, Elternverbände, aber auch Gewerkschaften sprachen sich insbesondere aufgrund der Betreuungskomponente von Ganztag für einen Ausbau aus (Ottweiler 2005).

In der Folge wurde 2003 das Investitionsprogramm Zukunft Bildung und Betreuung (IZBB) von Bund und Ländern mit dem Ziel verabschiedet, „zusätzliche Ganztagschulen zu schaffen und bestehende Ganztagschulen qualitativ weiterzuentwickeln“ (BMBF 2003, S. 2). Schulen konnten diese Mittel beantragen und für bauliche Zwecke und Ausstattungsinvestitionen verwenden, nicht jedoch für zum Beispiel pädagogisches Personal. Ursprünglich war der Abrufzeitraum der vier Milliarden Euro Bundesmittel für 2003 bis 2008 festgesetzt, wurde aber bis 2009 verlängert, da Schulen in vielen Ländern mehr Zeit für das Abrufen benötigten (BMBF 2005). Inklusive Ländermitteln beliefen sich die IZBB-Investitionen auf knapp viereinhalb Milliarden Euro. Dieser finanzielle Aufwand ermöglichte den oben geschilderten umfangreichen Ganztagsausbau (vgl. Abb. 1).

2.3 Aktuelle Entwicklungen

Seit Ablauf des IZBB werden der Erhalt und weitere Ausbau von Ganztag, den alle Bundesländer verfolgen, durch die Länder selbst finanziert. Diese entwickelten hierzu verschiedene Investitionskostenförderprogramme (KMK 2015). Auf diese Weise sind in

finanzschwächeren Ländern geringere Investitionen zu erwarten. Zu Beginn des Jahres 2018 wurde durch die Regierung bekanntgegeben, dass es in diesem Zusammenhang zu Neuerungen und einem weiteren Investitionsprogramm durch den Bund kommen soll, um die Vereinbarkeit von Familie und Beruf und erweiterte Bildungsmöglichkeiten weiter zu verbessern. In dem Koalitionsvertrag der regierenden Parteien wurde eine „Anpassung der Rechtsgrundlage im Grundgesetz als Voraussetzung, um Länder bei Investitionen in die Bildungsinfrastruktur unterstützen zu können“ (Koalitionsparteien 2018, S. 11 f.), angekündigt. Die strikte Finanzhoheit der Länder in Bildungsfragen soll demnach gelockert werden, um eine direkte Bundesinvestition in Ganztags- und Betreuungsangebote zu ermöglichen. Diese soll zwei Milliarden Euro betragen. Bis 2025 soll ein flächendeckender Rechtsanspruch auf Ganztagsbetreuung für alle Grundschulkindern geschaffen werden. Wie schon bei der Einführung von Rechtsansprüchen auf einen Kindergartenplatz für Drei- bis Sechsjährige im Jahr 1996 und für Ein- bis Dreijährige im Jahr 2013 ist im Fall des Eintretens der neuen Reform mit weiteren massiven Ausbauentwicklungen zu rechnen (vgl. auch Autorengruppe Bildungsberichterstattung 2018). Inwiefern diese Wahlversprechen jedoch eingehalten und welche Auswirkungen sie haben werden, muss sich zukünftig zeigen.

3. Annahmen und Befunde zu Konsequenzen des Ganztags

Während die Schwerpunkte der politischen Zielsetzungen über die Zeit leicht verschieden gelagert waren, lässt sich doch zusammenfassen, dass zum einen steigende nachmittägliche Betreuungsbedarfe und zum anderen Bildungsziele ursächlich für den forcierten Ganztagschulausbau waren und sind. Wie beispielsweise von Steiner (2009) zusammengefasst besteht jedoch „gerade beim Thema Ganztagschule eine auffällige Diskrepanz zwischen zum Teil weit reichenden politischen Reformhoffnungen und eher verhaltenen Erfolgserwartungen gerade auch auf Seiten von Experten/innen“ (ebd., S. 81), die die Hoffnungen für nicht hinreichend pädagogisch untermauert erachten. Wie sehr sich die politischen Erwartungen von Bildungs- und Betreuungszielen also mit tatsächlichen Wirkungen decken, ist eine empirische Frage (vgl. auch Böttcher 2015; Rauschenbach et al. 2012; Steinmann/Strietholt im Druck). Im Folgenden werden daher wissenschaftliche Annahmen und empirische Befunde zu Wirkungen ganztägiger Bildung und Betreuung überblickt.

Die zentralste Quelle empirischer Befunde stellt hierbei die Begleitevaluation des Ganztagschulausbaus dar, die längsschnittliche Studie zur Entwicklung von Ganztagschulen

(StEG). StEG fokussierte in der ersten Projektphase (2005-2011) Rahmenbedingungen der Ganztagsarbeit und in der zweiten Phase (2012-2015) deren Wirkbedingungen. Aktuell (2016-2019) werden Untersuchungen zu Gestaltungseffekten vertieft und erstmals auch andere Settings als Ganztagschulen vergleichend betrachtet (DIPF 2018). In den Jahren 2012, 2015 und 2018 gab es außerdem querschnittliche repräsentative Ganztagschulleitungsbefragungen (StEG 2013; StEG 2015).⁵ Neben den StEG-Teilstudien sind weitere Einzelevaluationen mit alternativen Datengrundlagen von Interesse, insbesondere, wenn Vergleiche zwischen Halb- und Ganztagschulen angestellt werden.⁶ Dieser Vergleich stellt einen zentralen Anspruch des folgenden Literaturüberblicks dar. Es wird herausgearbeitet, welche Konsequenzen das Ganztagskonzept in der Praxis auf Ebene der Schulen, Familien und Schüler/-innen hat, die über Halbtagsschulen hinausgehen. Auf diese Weise soll beleuchtet werden, inwiefern die Ganztagszielsetzungen erreicht werden und ganztägige Beschulung Mehrwerte schafft.

3.1 Organisation und Ausgestaltung von Ganztagschulen

Um Ganztags zu etablieren, müssen über Halbtagsschulen hinausgehende, umfangreiche Herausforderungen der Schulorganisation und -entwicklung bewältigt werden. Diese liegen unter anderem in dem Bereitstellen räumlich-struktureller Voraussetzungen für das zusätzliche Ganztagsangebot und Mittagessen, in der Erarbeitung eines Ganztagskonzepts und in der Organisation geeigneten Personals und externer Kooperationspartner.

Zunächst einmal interessant sind die verschiedenen Motive, sich als Ganztagschule neu zu gründen oder von einer Halb- zu einer Ganztagschule zu werden. Verschiedene Befragungen zeigen, dass hierbei insbesondere die Elternnachfrage nach verlässlicher Nachmittagsbetreuung ausschlaggebend ist. Auch die Ziele einer Schulöffnung und holistischeren Ausrichtung (z.B. soziales Lernen, Persönlichkeitsentwicklung) stehen für die Schulen häufiger im Vordergrund als Motive im Bereich gesteigerter Kompetenz- und Begabungsförderung. Das Ziel einer erweiterten fachlichen Lernkultur findet sich eher an weiterführenden Schulen als im Primarbereich und eher an gebundenen als an offenen

⁵ Zu der StEG-Projektphase 2016-2019 sowie der Schulleitungsbefragung 2018 liegen bislang keine publizierten Befunde vor.

⁶ Forschungsüberblicke außerhalb der StEG-Studien bieten beispielsweise BMBF (2012) oder Züchner und Fischer (2011).

Ganztagsschulen (Höhmnn/Holtappels/Schnetzer 2004; Holtappels 2008b; Spillebeen/Holtappels/Rollett 2011; StEG 2013; StEG 2015).

Den Schulen obliegt primär selbst, welche Strategien sie in Bezug auf Schulentwicklungsaufgaben bei der Ganztagsschulgründung verfolgen. Das IZBB förderte formell nur Schulen, die ein pädagogisches Konzept hatten. Gleichzeitig zeigen verschiedene Studien, dass diese Ganztagskonzepte häufig nicht schriftlich festgehalten sind. Durchschnittlich sind systematische Maßnahmen zur (Weiter-)Entwicklung von Ganztags an gebundenen Ganztagsschulen und in den neuen Bundesländern weiter fortgeschritten. In Bezug auf wahrgenommene Herausforderungen bewerten Schulleitungen vor allem die Finanzierung von qualifiziertem zusätzlichem Personal als schwierig. Das Organisieren von Kooperationen wird insbesondere im ländlichen Bereich und die räumlich-materielle Ausstattung im städtischen Raum als herausfordernd erachtet (BMBF 2003; BMBF 2012; Höhmnn/Holtappels/Schnetzer 2004; Höhmnn/Grewe/Strietholt 2008; Holtappels/Rollett 2008; Spillebeen/Holtappels/Rollett 2011; StEG 2013; StEG 2015).

Im Folgenden wird erstens beleuchtet, welche Ausgestaltungsformen Ganztagsschulen in der Praxis annehmen, und zweitens wird dargelegt, was für Angebote mit welcher personellen Ausstattung bereitgestellt werden. Diese Befunde werden in einem dritten Abschnitt zusammengefasst und eingeordnet.

3.1.1 Strukturelle Formen und organisatorische Verantwortlichkeiten

Aufgrund der breiten Ganztagsschuldefinition der KMK und der unterschiedlichen Regelungen in den Bundesländern (KMK 2015; KMK 2018) haben sich sehr verschiedene strukturelle Formen von Ganztagsschulen herausgebildet. Neben dem Verbindlichkeitsgrad der Teilnahme der Schüler/-innen (offen, teilgebunden, vollgebunden) unterscheiden sich beispielsweise auch die zeitlichen Dimensionen (Stunden pro Tag, Tage pro Woche) und die täglichen Zeitstrukturen (Rhythmisierung von Unterricht und Angeboten, offener Anfang oder Schulschluss). Generell zeigt sich, dass Organisationsmodelle mit einer freiwilligen Ganztagsbeteiligung deutlich überwiegen; so waren in 2016 zwei Drittel aller Ganztagsschulen offen organisiert (KMK 2018). An diesen ist das Ganztagsangebot typischerweise additiv zum Vormittagsunterricht organisiert. Insbesondere an Gesamtschulen und Schulen mit mehreren Bildungsgängen finden sich gebundene Ganztagsformen und rhythmisierte Unterrichts- und Ganztagsselemente. Generell sind rhythmisierte Modelle selten, an denen also auch im Vormittagsbereich außerunterrichtliche Aktivitäten vorgesehen sind. Nur wenig Schulen

bieten ihr Ganztagsangebot an fünf Tagen pro Woche an. In Bezug auf die zeitlichen Ausdehnungen des Ganztags finden sich lediglich graduelle Unterschiede zwischen offenen, teilgebundenen und gebundenen Organisationsformen. In Grundschulen gibt es durchschnittlich längere Öffnungszeiten und vereinzelt auch Programme in den Schulferien (Holtappels 2008a; KMK 2018; StEG 2013; StEG 2015; Züchner/Fischer 2014).

Weitere zentrale Unterschiede zwischen Ganztagsmodellen bestehen in den organisatorischen Verantwortlichkeiten. Dies spiegelt sich zum einen in verschiedenen Trägerschaften des Ganztagsbetriebs, der operativen Verantwortlichkeit und dem Grad der Öffnung von Schule wider. Das IZBB sah lediglich eine Förderung von Schulen mit Hort- oder Jugendhilfeträgern vor, wenn „die Weiterentwicklung zu einem in die Schule fachlich integrierten Ganztagsangebot angestrebt“ (BMBF 2003, S. 3) wurde. Der Ganzttag wird jedoch oft durch externe Träger oder in Kooperation mit externen Trägern organisiert. Dies ist insbesondere an Schulen im Primarbereich und mit offener Organisationsform der Fall. Insgesamt wird der Ganzttag häufig nicht von den (stellvertretenden) Schulleitungen gesteuert. Konform mit den Zielen einer breiten Öffnung von Schule arbeitet der überwiegende Teil der Ganzttagsschulen eng mit mehreren außerschulischen Partnern zusammen. Sowohl der Fall, dass das Ganztagsprogramm ausschließlich durch außerschulische Partner bereitgestellt wird, als auch der, dass es ausschließlich ohne Partner organisiert wird, bilden die Ausnahme. In den neuen Bundesländern sind die Schulen typischerweise breiter vernetzt. Die meisten Ganzttagsschulen kooperieren mit Partnern aus den Bereichen Sport, Kunst, Musik, mit Bibliotheken und Kirchengemeinden. Insbesondere im Grundschulbereich spielen Horte und Jugendzentren eine große Rolle. Insgesamt seltener finden sich zum Beispiel Kooperationen mit Einrichtungen der Jugendsozialarbeit oder Betrieben (Arnoldt 2008; Arnoldt 2011; BMBF 2012; StEG 2013; StEG 2015).

3.1.2 Ganztagsangebote und eingebundenes Personal

Konform mit dem Ziel eines breiten Angebotsspektrums an Ganzttagsschulen werden flächendeckend facettenreiche Angebote bereitgestellt. Die drei Bereiche sportbezogene Angebote (z.B. Fußballclub), musisch-künstlerische Angebote (z.B. Chor, Kunstarbeitsgemeinschaft) sowie lernunterstützende und fachliche Angebote (z.B. Hausaufgabenbetreuung bzw. Lernzeiten, Förderunterricht, Informatikkurse) werden von nahezu allen Ganzttagsschulen abgedeckt. Generell wird der zusätzliche Zeitrahmen an Ganzttagsschulen häufig für das Bearbeiten von Lerndefiziten und für eine individuelle

Förderung genutzt. Seltener, aber auch weitverbreitet sind beispielsweise Angebote in den Bereichen des sozialen Lernens, neuer Medien oder der Berufsorientierung. Vertiefende fachliche Angebote finden sich häufiger an Gymnasien, während Berufsorientierungs- und Handwerksangebote stärker an anderen weiterführenden Schulformen vertreten sind. Tendenziell sind die Inhalte der Angebote an Schulen mit einer längeren Ganztagsgeschichte und gebundener Organisationsform enger mit den Unterrichtsinhalten verzahnt als an jüngeren oder offenen Ganztagschulen. Diese Verknüpfung von Angeboten mit dem regulären Unterricht, die Bestandteil der KMK-Ganztagschuldefinition ist, ist jedoch an etwa der Hälfte der Ganztagschulen wenig gegeben (StEG 2015). Für die am Ganztag teilnehmenden Schüler/-innen wird überall ein Mittagessen angeboten, das in den meisten Fällen mit Kosten für die Eltern verbunden ist. Die Teilnahme an den Ganztagsprogrammen ist überwiegend kostenfrei, vor allem im Sekundarbereich (BMFSFJ 2017; Holtappels 2008a; Holtappels/Rollett 2008; Rollett et al. 2011; StEG 2013; StEG 2015). Auch an vielen Halbtagschulen finden sich vielfältige außerschulische Angebote (z.B. Hausaufgabenbetreuung, Arbeitsgemeinschaften, Förderangebote, Freizeitangebote), für die sich die Schüler/-innen anmelden und an denen sie teilweise auch vier bis fünfmal wöchentlich teilnehmen können. An Ganztagschulen steht jedoch eine deutlich größere Angebotsbreite zur Verfügung (Drossel et al. 2012; Hertel et al. 2008; Wendt et al. 2016).

In der Praxis zeigt sich, dass im Gegensatz zu Halbtagschulen typischerweise sehr verschiedene Personalgruppen in den Ganztagsbetrieb eingebunden werden. Dieses zusätzliche Personal bildet eine sehr heterogene Beschäftigtengruppe mit unterschiedlichen Qualifikationen (z.B. pädagogische Laien ohne akademischen Abschluss, Sozialarbeiter/-innen, Pädagog/-innen) und Beschäftigungsformen (z.B. ehrenamtlich, bei Kooperationspartnern, Schul- oder anderem Träger angestellt, Stundenumfang, Dauer der Beschäftigung). Das weitere Personal fluktuiert stark und ist gerade im Grundschulbereich häufig nicht beim Schulträger angestellt. Auch arbeiten durchschnittlich nur zwischen einem Drittel und der Hälfte der Lehrkräfte selbst am Ganztag mit (StEG 2015). Die Lehrkraftbeteiligung ist bei Schulen in gebundener Form und mit längerer Ganztagstradition höher. An vielen vor allem weiterführenden Schulen unterscheiden sich die Kooperationen stark und es gibt keine gemeinsamen Fortbildungsveranstaltungen von Lehrkräften und weiterem Personal. Diese tauschen sich zwar häufig über Lernentwicklungen aus, jedoch werden nur an circa der Hälfte der Schulen gemeinsame Lerninhalte für Unterricht und Angebote erarbeitet (ebd.). Häufig und insbesondere an Gymnasien sind für solche

Kooperationen keine festen Zeiträume eingeplant. An Schulen der neuen Bundesländer, mit gebundener Form und längerer Ganztagsstradition besteht durchschnittlich mehr Raum für Austausch zwischen den Personalgruppen (Autorengruppe Bildungsberichterstattung 2018; BMBF 2012; Höhmann/Bergmann/Gebauer 2008; K. Tillmann 2011; StEG 2013; StEG 2015).

3.1.3 Zwischenfazit – Heterogenität an Organisationsformen

Zusammenfassend lässt sich sagen, dass Ganztagschulen in der Praxis sehr heterogene Ausgestaltungsformen aufweisen, die durch die Unterteilung in offene, teilgebundene und vollgebundene Ganztagschulen nur bedingt widerspiegelt werden. Darüber hinaus grenzen sich Ganztagschulen in verschiedener Hinsicht nur eingeschränkt von Halbtagschulen ab, die auch ein Nachmittagsprogramm anbieten können.

Über die beschriebenen Ausgestaltungsfacetten hinweg lassen sich trotz der schwierigen Kategorisierbarkeit zwei Extremformen von Ganztagschulen kontrastieren. Einen Prototyp bilden Schulen mit einem Ganztagsprogramm, das an vier bis fünf Wochentagen angeboten wird und bei dem die Teilnahme verpflichtend ist. Diese Schulen richten ihr Ganztagskonzept typischerweise auf eine verbesserte Lernkultur aus und halten es schriftlich fest. Unterrichtliche und außerunterrichtliche Elemente werden über den Tag rhythmisiert angeboten und die Inhalte miteinander verzahnt. Die Lehrkräfte sind an den Ganztagsangeboten beteiligt und kooperieren mit dem weiteren Personal. Der Ganztag ist üblicherweise eng in den Schulalltag integriert, wird von der Schulleitung konzipiert und gesteuert sowie vom Schulträger getragen. Diese Form von Ganztagschulen findet sich insbesondere in den neuen Bundesländern, an Ganztagschulen mit längerer Tradition und an Gesamtschulen beziehungsweise Schulen mit mehreren Bildungsgängen.

Die oben beschriebenen Befunde suggerieren, dass diesem Prototyp eine (größere) Gruppe von Schulen gegenübersteht, bei denen der Ganztag vor allem die Ziele erweiterter Betreuung und Öffnung von Schule verfolgt. Hierbei ist der Ganztag nur lose mit der regulären Beschulung verknüpft. So bestehen kein schriftliches Ganztagskonzept, keine verbindliche Teilnahme und eine rein additive Zeitstruktur von Unterricht und Ganztagsangeboten. Der Ganztag unterliegt einer außerschulischen Trägerschaft und Steuerung, die Lehrkräfte und das weitere Personal haben wenig Berührungspunkte und das Ganztagsangebot ist weitestgehend unabhängig von den Unterrichtsinhalten. Diese Form

findet sich vor allem in den alten Bundesländern, im Grundschulbereich und an jungen Ganztagschulen.

3.2 Konsequenzen des Ganztags für Eltern und Familienleben

Ein Ziel von Ganztag ist, dass die verlängerte Betreuung von Ganztagschüler/-innen insbesondere im Grundschulbereich die Eltern entlasten und zu einer besseren Vereinbarkeit von Familie und Beruf beitragen soll. Wie Abbildung 1 illustriert ist die Verfügbarkeit von Betreuungsplätzen durch Ganztagsangebote in den vergangenen Jahren erheblich gestiegen. Befragungen belegen darüber hinaus, dass die Nachfrage von Ganztagsplätzen unter Eltern seit Jahren hoch ist und weiter steigt, insbesondere unter Alleinerziehenden und Familien mit zwei berufstätigen Elternteilen. Die Eltern bevorzugen hierbei Angebote mit freiwilliger Teilnahme. Einerseits liegen Einschätzungen von Eltern vor, nach denen Familie und Beruf durch den Ganztagsbesuch besser vereinbar sind (Börner 2011; Höhmann/Holtappels/Schnetzler 2004; Züchner 2011). Andererseits zeigen auch empirische Evaluationen, dass die durch Ganztag breiter verfügbaren Betreuungsmöglichkeiten im Grundschulbereich tatsächlich zu einer gesteigerten Arbeitsmarktbeteiligung von Eltern, insbesondere Müttern, führen (Bauernschuster/Schlotter 2015; Gambaro/Marcus/Peter 2016).

Neben Betreuungserleichterungen werden verschiedene positive und negative Effekte auf das Familien- und Freizeitleben antizipiert. Einerseits erhoffen sich manche Familien durch die zeitlichen Entlastungen und das Verlagern der Hausaufgaben und interessensnaher Nachmittagsbeschäftigungen in die Schule positive, konfliktreduzierende Effekte auf das Familienleben. Andererseits wird teilweise befürchtet, dass Ganztagschüler/-innen zu viel Zeit in der Schule verbringen, sich das Familienleben einschränken und die Freizeit der Kinder und Jugendlichen verschulen könnte. Fasst man Befunde verschiedener Befragungen von Schüler/-innen und Eltern aus den letzten Jahren zu diesen Erwartungen zusammen, dann zeigt sich zunächst, dass die Akzeptanz von Ganztag in den Familien steigt. Befragungen von Eltern zeigen einerseits, dass diese die Ganztagsangebote grundsätzlich positiv bewerten, sich jedoch eine bessere fachliche Lernunterstützung wünschen. Insbesondere wenn eine hochwertige Hausaufgabenbetreuung in der Schule geboten wird, schätzen Eltern dies sehr, da diese Lernunterstützung eine Überforderung darstellen und zu alltäglichen Konflikten mit ihren Kindern führen kann. Auch im Bereich erzieherischer Probleme fühlen sich Eltern von Ganztagschüler/-innen entlastet. Aus Sicht von Eltern, insbesondere solchen mit einem niedrigen sozioökonomischen Status, fördert die Ganztagsteilnahme familiäre Beziehungen.

Befragungen von Schüler/-innen zeigen, dass sie die Ganztagsbeteiligung besonders dann positiv bewerten, wenn sie eine gute Beziehung zu dem verantwortlichen Personal haben, die Angebote ihren Interessen entsprechen, sie diese mitgestalten und sich selbst zur Teilnahme entscheiden können. Generell bewerten ältere Schüler/-innen die Prozessqualität des Ganztags negativer als jüngere. In Bezug auf eine Verschulung von Freizeit zeigt sich lediglich im Bereich von Sportvereinen, dass am Ganztag Teilnehmende weniger involviert sind als Halbtagschüler/-innen. In anderen Bereichen unterscheiden sich Teilnehmende und Nichtteilnehmende kaum in ihren außerschulischen oder familiären Aktivitäten (BMBF 2012; Börner 2011; Brümmer/Rollett/Fischer 2011; Höhmann/Holtappels/Schnetzer 2004; StEG 2016; Wissenschaftlicher Beirat für Familienfragen 2006; Züchner 2011; Züchner/Arnoldt 2011).

Neben den genannten Punkten sollen Ganztagschulen bessere Möglichkeiten haben, Eltern in das Schulgeschehen einzubeziehen, als Halbtagschulen. Ganztagschulen sollen aufgrund ihrer zeitlichen und personellen Ressourcen und breiteren Öffnung mehr Chancen bieten, Erziehungspartnerschaften und aktive Kooperationen zwischen Eltern und Schule zu ermöglichen. So können Eltern beispielsweise in die Ganztagsplanung und -ausführung einbezogen werden oder von einem direkteren Kontakt zur Jugendhilfe profitieren. Diesbezüglich weisen Studien darauf hin, dass insbesondere das Erreichen schuldistanzierter Eltern oder Eltern mit Migrationshintergrund an Ganztagschulen gelingen kann, wenn dies als zentrales Ziel wahrgenommen und sorgfältig verfolgt wird (BMBF 2012; Lokhande/Hoefl/Wendt 2014; Wissenschaftlicher Beirat für Familienfragen 2006).

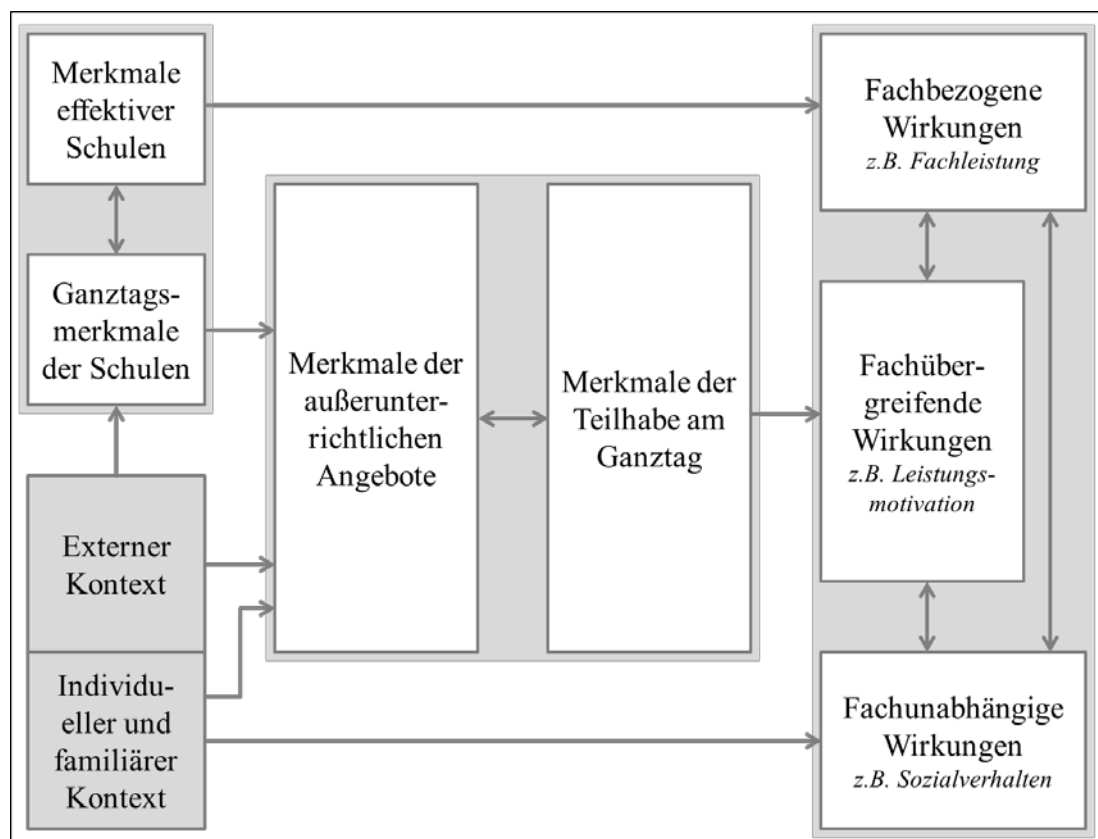
Zusammenfassend lässt sich sagen, dass Ganztag erwartungsgemäß die Vereinbarkeit von Familie und Beruf erleichtert und entgegen mancher Befürchtungen eher förderliche als abträgliche Effekte auf das Familienleben aufweist. Darüber hinaus scheinen Ganztagschulen ein höheres Potenzial für den Einbezug von Eltern zu haben.

3.3 Konsequenzen des Ganztags für Schüler/-innen – Teilnahmemerkmale und Effekte

Eine zentrale Wirkebene von Ganztagschulen stellt die der Schüler/-innen dar. Radisch (2009) illustriert in seinem Wirkmodell, dass der individuelle, familiäre und externe Bezugsrahmen sowie der allgemeine und ganztagsbezogene Kontext der Schulen den Rahmen für Effekte auf Schüler/-innen bilden. Diese Wirkungen werden sowohl durch Merkmale der außerunterrichtlichen Angebote als auch der Ganztagsbeteiligung beeinflusst. Bei Effekten auf Schüler/-innen werden fachbezogene (z.B. Fachleistung), fachübergreifende (z.B.

Leistungsmotivation) und fachunabhängige (z.B. Sozialverhalten) Wirkungen unterschieden (vgl. Abb. 3). Das Modell nimmt an, dass Ganztagschulen Schüler/-innen potenziell besser als Halbtagschulen fördern können, wenn die Schulen geeignete Merkmale aufweisen; beispielsweise hinsichtlich ihrer Personalversorgung, Kooperationspartner und Ganztagsangebote. Selbst wenn eine Ganztagschule ein hochwertiges Angebot bereitstellt, sind nur Effekte für Schüler/-innen zu erwarten, die ausreichend intensiv und dauerhaft an den entsprechenden Angeboten teilnehmen (vgl. auch Miller/Truong 2009; Stecher et al. 2007). Der Grundlogik dieses Modells folgend werden in den folgenden drei Abschnitten Befunde zu Teilnahmecharakteristika, zu nichtfachbezogenen sowie zu fachbezogenen Wirkungen auf Schüler/-innen illustriert.

Abbildung 3: Modell der Bildungsqualität außerunterrichtlicher Angebote in der Ganztagschule



(Quelle: Darstellung adaptiert nach Radisch 2009, S. 87, Abb. 4-3)

3.3.1 Teilhabe am Ganzttag

Im Vergleich zur traditionellen halbtägigen Beschulung sollen Ganzttagsschüler/-innen die Möglichkeit haben an reichhaltigen Angeboten teilzunehmen. Vergleicht man Angebot und Nachfrage, so findet sich im Grundschulbereich an jeder vierten Schule ein Überhang von Anmeldungen im Vergleich zum Platzangebot (StEG 2015). In weiterführenden Schulen decken sich Angebot und Nachfrage eher. Im Vergleich mit früheren Befunden scheint der Bedarf mittlerweile insgesamt besser gedeckt zu sein (Bargel 1991; Höhmann/Holtappels/Schnetzler 2004; StEG 2015; StEG 2016).

Verschiedene Angebote werden mehr oder weniger intensiv besucht, was auch mit dem Grad der Teilnahmeverpflichtung zusammenhängt. Daher sind an offenen Ganzttagsschulen niedrigere Teilnahmequoten zu finden als an gebundenen Formen. In den neuen Bundesländern finden sich höhere Beteiligungsquoten als in den alten. Generell werden Sport-, Kunst-, Musik- und Freizeitangebote häufiger und zeitlich umfangreicher besucht, als fachliche Angebote wie zum Beispiel Arbeitsgemeinschaften oder Förderangebote, deren Beteiligungsquoten häufig unter 20% liegen. Hausaufgabenbetreuungsangebote erreichen im Vergleich mit anderen fachlichen Programmen durchweg hohe Beteiligungsquoten, in der fünften Klasse beispielsweise um die 50% (StEG 2016). Jugendliche nehmen mit zunehmendem Alter insgesamt seltener teil, gleichzeitig aber stärker an fachlichen im Vergleich zu anderen Angeboten. Darüber hinaus fluktuiert die Teilhabe über die Zeit und nur wenige Schüler/-innen nutzen dasselbe Angebot über mehrere Jahre (Autorengruppe Bildungsberichterstattung 2016; Autorengruppe Bildungsberichterstattung 2018; Holtappels/Jarsinski/Rollett 2011; KMK 2018; StEG 2013; StEG 2015; StEG 2016; Steiner 2011b). Hierbei muss betont werden, dass Kinder und Jugendliche auch an Halbtagschulen an freizeitbezogenen und fachlichen Nachmittagsangeboten teilnehmen und sich die Beteiligungsquoten insbesondere im Vergleich mit offenen Ganzttagsschulen kaum unterscheiden (Hertel et al. 2008; Wendt et al. 2016; Willems et al. 2014).

Ein weiteres Ziel der Ganzttagsschule liegt darin, auch Schüler/-innen eine Teilnahme an Lern- und Freizeitangeboten zu ermöglichen, die ansonsten keinen Zugang dazu hätten (Steiner 2011b; Züchner/Fischer 2014). In Bezug auf diese chancengerechtigkeitssteigernde Funktion von Ganzttag muss berücksichtigt werden, dass Gebühren für das Mittagessen und die Angebotsteilnahme erhoben werden können und es Bundesländer gibt, in denen Schüler/-innen mit berufstätigen Eltern bei der Ganztagsplatzvergabe bevorzugt werden. Solche Regelungen benachteiligen tendenziell Schüler/-innen aus finanzschwachen Elternhäusern.

Evaluationen der sozialen Selektivität der Teilnahme zeichnen ein heterogenes Bild. Abhängig von Stichprobe und untersuchtem Angebot zeigen manche Studien, dass sozioökonomisch schlechter gestellte Schüler/-innen gleich häufig oder sogar häufiger an außerunterrichtlichen Angeboten an Ganz- und Halbtagschulen teilnehmen als besser gestellte. Andere legen hingegen nahe, dass sie seltener teilnehmen. In Bezug auf weitere Schüler/-innenmerkmale sind die Befunde einheitlicher. So nehmen Kinder mit berufstätigen Eltern häufiger teil als von Nichterwerbstätigen. Es finden sich konsistent kaum Teilnahmeunterschiede zwischen Mädchen und Jungen, Schüler/-innen mit und ohne sonderpädagogischem Förderbedarf oder Schüler/-innen mit und ohne Migrationshintergrund. Generell begünstigt eine frühere Ganztagsbeteiligung die zukünftige Beteiligung (BMBF 2012; BMFSFJ 2017; Prein/Rauschenbach/Züchner 2009; StEG 2015; StEG 2016; Steiner 2011b; Wendt et al. 2016; Willems et al. 2014).

Zusammenfassend lässt sich sagen, dass die Teilnahme an Ganztagsangeboten nur wenig sozial selektiv verläuft. In diesem geringen Ausmaß sozialer Selektivität unterscheiden sich Ganz- und Halbtagschulen mit Nachmittagsangebot kaum (Wendt et al. 2016; Willems et al. 2014). Diese Befunde sind insbesondere interessant, wenn man sie in den Kontext alternativer Maßnahmen setzt. So zeigt sich, dass der Zugang außerhalb der Schule zu privater Nachhilfe, Musikschulen oder Sportvereinen primär wegen höherer Kosten deutlich stärker sozioökonomisch selektiv ist (Autorengruppe Bildungsberichterstattung 2014; BMBF 2012; Stecher/Maschke 2013).

3.3.2 Fachübergreifende und fachunabhängige Wirkungen

Aufgrund der breiten, über klassischen Halbtagsunterricht hinausgehenden Angebote und der Einbindung multiprofessionellen Personals wird von Ganztagschulen eine holistischere Betreuung, Erziehung und Bildung erwartet. Im Vergleich mit herkömmlicher halbtägiger Schule sollen an Ganztagschulen mehr Möglichkeiten für die Schüler/-innen bestehen, Programme interessensgeleitet zu wählen und sich selbst einzubringen. Von Ganztagsangeboten, die Fächer vertiefen oder mit Freizeitelementen verknüpfen, werden beispielsweise ein verstärktes fachliches Interesse oder höhere Leistungsmotivationen erhofft. Angebote aus dem Bereich des sozialen Lernens sollen soziale Kompetenzen stärken, problematisches Verhalten reduzieren und das Schulklima verbessern (Holtappels 2006; Radisch 2009; Stecher et al. 2007).

Längsschnittliche Evaluationsstudien deuten darauf hin, dass Ganzttag in diesen fachübergreifenden und fachunabhängigen Bereichen tatsächlich positive Effekte aufweist, wenn bestimmte Voraussetzungen erfüllt sind oder einzelne Schüler/-innengruppen betrachtet werden. Grundsätzlich scheinen hochwertige Programme zu günstigen motivationalen Entwicklungen und einer verbesserten Schulverbundenheit beitragen zu können. Durch das Besuchen fachbezogener Angebote in den Bereichen Lesen, Medien und Naturwissenschaften zeigen sich beispielsweise förderliche Effekte auf entsprechend fachliche Leistungsmotivationen, Selbstkonzepte und Lernzielorientierungen; insbesondere, wenn die Teilnehmenden die Angebotsqualität positiv bewerten und sich freiwillig angemeldet haben. Darüber hinaus weisen Schüler/-innen mit Migrationshintergrund oder mit Eltern mit niedrigen Bildungsabschlüssen stärkere Vorteile auf. In Bezug auf soziales Lernen zeigen sich keine allgemeinen Effekte, allerdings scheint eine kontinuierliche Teilnahme an qualitativ hochwertigen Angeboten in den Bereichen des sozialen Lernens und Teamsports zu einer günstigen Entwicklung prosozialen Verhaltens zu führen, ebenfalls insbesondere bei Schüler/-innen mit Migrationshintergrund (Fischer/Brümmer/Kuhn 2011; Fischer/Kuhn/Züchner 2011; Fischer et al. 2016; Fischer/Theis 2014; Lossen et al. 2016; StEG 2016). Im Vergleich mit Halbtagschüler/-innen scheinen sich prosoziales und Beziehungsverhalten sowie sozial-emotionale Kompetenzen von Ganztagschüler/-innen tendenziell positiver zu entwickeln (BMBF 2012).

3.3.3 Fachbezogene Wirkungen

Ein wichtiges Ziel von Ganztagschulen ist es, die zusätzlichen zeitlichen und personellen Kapazitäten für eine verbesserte Lernunterstützung zu nutzen. Beispielsweise sollen in Phasen der Hausaufgabenbetreuung oder individuellen Förderung Lernstärken und -schwächen besser erkannt werden, sodass (insbesondere sozial schwache) Schüler/-innen gezielter gefördert werden können. Durch das Einbinden von Entspannungs- und Bewegungsphasen in den Nachmittag oder sogar Vormittag sollen außerdem Leistungstiefs aller Schüler/-innen vermieden und die Lernmotivation aufrechterhalten werden (Holtappels 2006; KMK 2002; Radisch 2009; Steiner 2009). In empirischen Evaluationen wurden einerseits Ganztageeffekte auf Schulnoten und -laufbahn sowie andererseits auf mit Kompetenztests gemessenen Leistungen untersucht. In Längsschnittstudien konnte gezeigt werden, dass eine generelle Teilnahme noch keine notenförderlichen Effekte hat, aber eine intensive Beteiligung insbesondere an Ganzttag mit einer hohen wahrgenommenen Prozessqualität zu besseren

Notenentwicklungen in Mathematik und Deutsch führt. Schüler/-innen, die Angebote insbesondere leistungsorientiert auswählen, scheinen außerdem bessere Notenentwicklungen aufzuweisen, als solche mit eher freizeitbezogenen Nutzungsprofilen. Darüber hinaus reduziert eine intensive und kontinuierliche Ganztagsbeteiligung das Risiko einer Klassenwiederholung (in der fünften Klasse um etwa 50%) und scheint mit einem geringeren Abschulungsrisiko einherzugehen (Kuhn/Fischer 2011; StEG 2016; Steiner 2011a).

Darüber hinaus sind Wirkungen auf standardisierte Testleistungen von Interesse. Studien, die Kompetenzentwicklungen untersuchen und relevante Variablen kontrollieren, finden keine signifikanten Unterschiede zwischen teilnehmenden und nichtteilnehmenden Ganztagschüler/-innen. Es finden sich jedoch vorsichtig zu interpretierende Hinweise darauf, dass eine freiwillige Teilnahme und eine hohe Angebotsqualität bedeutsame Mediatoren für Lernentwicklungen sein könnten. Studien, die Schulen mit und ohne Ganztagsangebot vergleichen, finden keine signifikanten Unterschiede in den Leistungsniveaus. Auch zeigen sich keine signifikanten Effekte auf das Ausmaß sozialer Ungleichheit in den Kompetenzen (Bellin/Tamke 2010; Fischer et al. 2016; Linberg/Struck/Bäumer 2018; Lossen et al. 2016; Steinmann/Strietholt im Druck; Steinmann/Strietholt/Caro im Druck; Strietholt et al. 2015). Demnach lassen sich die teils positiven Effekte auf Schulnoten nicht mit standardisierten Kompetenzmessungen replizieren. Eine mögliche Erklärung ist, dass durch Lehrkräfte vergebene Schulnoten und damit zusammenhängende Laufbahnentscheidungen (z.B. Nichtversetzung, Abschlussschulung) im Gegensatz zu Kompetenztests auch Verhaltensaspekte widerspiegeln (z.B. Erledigen der Hausaufgaben, Sozialverhalten). Diese Überlegung ist konform mit den zuvor geschilderten Befunden zu nichtfachbezogenen Wirkungen des Ganztags.

4. Fazit – Ganztagschule als neue Schulorganisationsform?

Zunächst lässt sich festhalten, dass der durch das IZBB angestoßene Ganztagschulausbau zu den ambitioniertesten Reformen des Bildungswesens gehört. Während sich der enorme finanzielle Aufwand sehr deutlich in der quantitativen Zunahme von Ganztagsbetreuung widerspiegelt und aus wissenschaftlicher Sicht ein hohes Potenzial erweiterter Bildung besteht, schöpft die derzeitige Ausgestaltung des Ganztags dieses Bildungspotenzial zumindest in der Breite nicht aus. Aus den Evaluationsbefunden zu Konsequenzen des Ganztagskonzepts für Schulen, Familien und Schüler/-innen geht hervor, dass Ganztags in der Praxis eine „heterogene Intervention“ (Züchner/Fischer 2011, S. 15) darstellt und sich nur

bedingt von traditioneller Halbtagschule abgrenzen lässt (StEG 2015; Wendt et al. 2016).

Die größten schulorganisatorischen Unterschiede zwischen Halb- und Ganztagschulen kommen durch das Sicherstellen räumlicher, struktureller und personeller Voraussetzungen zustande, wobei die Schulleitungen in sehr unterschiedlichem Maß involviert sind. Insbesondere wenn Ganztagschulen ein additives Nachmittagsprogramm bereitstellen, das nicht mit dem Unterricht verknüpft und nicht durch Lehrkräfte geleitet wird, sind die Abgrenzungen zu Halbtagschulen mit Nachmittagsangebot graduell. Die definitorische Unterscheidung offener, teilgebundener und gebundener Ganztagschulen erweist sich in der Praxis als wenig differenzierend. Auf Ebene der Familien zeigt sich, dass der Ganztags zu einer besseren Vereinbarkeit von Familie und Beruf führt und insgesamt breit akzeptiert wird. Entsprechende Angebote scheinen eine entlastende und bereichernde Funktion für das Familienleben zu haben. Auf Ebene der Schüler/-innen ist zunächst festzustellen, dass im Ganztags tatsächlich breite Bildungs- und Freizeitaktivitäten zur Verfügung stehen. Der Zugang ist an Ganztagschulen im Vergleich mit privaten Angeboten wenig sozial selektiv und trägt damit zu einer erhöhten Chancengleichheit bei. Gleichzeitig sind auch hier die Abgrenzungen zu Halbtagschulen mit Nachmittagsangebot fraglich. Darüber hinaus lassen sich bei intensiver Teilnahme an qualitativ hochwertigen Programmen positive Effekte auf Entwicklungen in den Bereichen der Motivation, des Sozialverhaltens, der Schulnoten und Kontinuität der Schullaufbahn nachweisen. In Hinblick auf standardisiert gemessene Schulleistungen und soziale Bildungsungleichheiten in diesen Schulleistungen zeigen sich jedoch keine Unterschiede zwischen Halb- und Ganztagschulen.

Die große Heterogenität in der Ausgestaltung von Ganztagschulen hängt mit der recht zeitstrukturell gehaltenen KMK-Definition zusammen, die als einzige pädagogische Qualitätsaspekte des Ganztags die Verantwortlichkeit und Beteiligung der Schulleitung sowie einen „konzeptionellen Zusammenhang mit dem Unterricht“ (KMK 2018, S. 5) nennt. Darüber hinaus wird die KMK-Definition in den Bundesländern verschieden ausgestaltet (vgl. auch Böttcher 2015; StEG 2015). Holtappels (2006) bemängelte zum Beispiel schon früh, dass die KMK-Definition keine Standards beinhalte, und auch Böttcher (2015) kommt zu dem Schluss, dass eine „strategische Regie, die die operative Arbeit vor Ort strukturieren könnte“ (ebd., S. 51) fehle. Interessant ist in diesem Zusammenhang, dass die KMK-Definition mittlerweile noch weiter geöffnet wurde und nun auch die Kombination von Schule und kooperierendem Hort umfasst, bei der die Schulleitung lediglich mitverantwortlich für den Ganztagsbetrieb ist (KMK 2018). Dadurch scheint aus pädagogischer Sicht mittlerweile umso

mehr zu gelten, dass „die Kriterien der KMK zur Ganztagschule an strukturbildender Kraft verlieren. Ganztagschulen könnten auf ihre zusätzlichen Betreuungszeiten reduziert werden und pädagogisch motivierte Ziele, die mit der Einführung von Ganztagschule verbunden waren [...] aus dem Blick geraten“ (StEG 2015, S. 106).

Wenn die Ganztagschule in Deutschland zukünftig in Hinblick auf ihre Bildungspotenziale weiterentwickelt werden soll, dann scheint es einen Bedarf nach einer klareren definitorischen Abgrenzung von Ganz- und Halbtagsschulmodellen zu geben. Diese müsste über die bestehende Unterscheidung offener, teilgebundener und gebundener Modelle hinausgehen und über die Länder hinweg vereinheitlicht werden. Es müssten solche Merkmale abgebildet werden, die einer Ganztagschule erst den potenziellen pädagogischen Mehrwert gegenüber Halbtagsschulen verleihen und sie zu einer neuen Schulorganisationsform machen (Radisch 2009; StEG 2015; StEG 2016). Diese tatsächlich über die Halbtagsschule hinausgehende Ganztagsorganisationsform mit erweiterter Bildungsfunktion ist eher bei (bislang seltenen) Ganztagschulen zu erwarten, die ein zeitlich intensives und verpflichtendes Ganztagsprogramm bereitstellen, das pädagogisch und personell hochwertig ausgestattet, mit dem regulären Unterricht verbunden und zeitlich rhythmisiert ist (Holtappels 2006; StEG 2015; Steiner 2009).

Denkt man über mögliche künftige Entwicklungen der Ganztagschule nach, erscheinen die von den Regierungsparteien angekündigten neuen Ganztagsinvestitionen und gesetzlichen Änderungen besonders interessant (Koalitionsparteien 2018). Diese neue Finanzierung wäre nicht nur für den weiteren Ausbau, sondern auch die Weiterentwicklung von Ganztags bedeutsam, da Schulleitungen insbesondere das Finanzieren des Ganztagsbetriebs als problematisch für ihre Schulentwicklungsarbeit wahrnehmen (StEG 2013). Darüber hinaus bietet die angekündigte Aufweichung der Bundesländerzuständigkeit für Bildungsfragen im Ganztagsbereich das Potenzial, bundesweit geltende Kriterien an die Investitionen zu knüpfen. Die Sicherung der personellen und strukturellen Ausstattung der Ganztagschulen, aber auch die Entwicklung bindender Definitions- und Qualitätsstandards über die Bundesländer hinweg werden als zentral für die erfolgreiche Weiterentwicklung des Ganztagskonzepts in der Praxis diskutiert (Böttcher 2015; K.-J. Tillmann 2011; Züchner/Fischer 2011). Solche Maßnahmen könnten dazu beitragen, die Schulorganisationsformen, die unter dem Begriff *Ganztagschule* zusammengefasst werden, künftig trennschärfer und pädagogisch gehaltvoller von Halbtagsschulen abzugrenzen.

Literaturverzeichnis

- Allemann-Ghionda, C. (2009): Ganztagschule im europäischen Vergleich: Zeitpolitiken modernisieren – durch Vergleich Standards setzen? In: Zeitschrift für Pädagogik 54, S. 190–208.
- Arnoldt, B. (2008): Öffnung von Ganztagschule. In: Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L. (Hrsg.): Ganztagschule in Deutschland. Ergebnisse der Ausgangserhebung der „Studie zur Entwicklung von Ganztagschulen“ (StEG). 2. korrigierte Auflage. Weinheim und München: Juventa, S. 86–105.
- Arnoldt, B. (2011): Kooperation zwischen Ganztagschule und außerschulischen Partnern: Entwicklung der Rahmenbedingungen. In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 312–329.
- Autorengruppe Bildungsberichterstattung (2014): *Bildung in Deutschland 2014: Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: W. Bertelsmann Verlag.
- Autorengruppe Bildungsberichterstattung (2016): *Bildung in Deutschland 2016: Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: W. Bertelsmann Verlag.
- Autorengruppe Bildungsberichterstattung (2018): *Bildung in Deutschland 2018: Ein indikatorengestützter Bericht mit einer Analyse zu Wirkungen und Erträgen von Bildung*. Bielefeld: W. Bertelsmann Verlag.
- Bargel, T. (1991): Bestands- und Bedarfsanalysen zu Ganztagschulen und Ganztagsangeboten. In Bargel, T./Kluth, M. (Hrsg.): Ganztagschule. Untersuchungen zu Angebot und Nachfrage, Versorgung und Bedarf. Bad Honnef: Bundesdruckerei, S. 67–85.
- Bauernschuster, S./Schlotter, M. (2015): Public child care and mothers' labor supply— Evidence from two quasi-experiments. In: Journal of Public Economics 123, S. 1–16.
- Bellin, N./Tamke, F. (2010): Bessere Leistungen durch Teilnahme am offenen Ganztagsbetrieb? In: Empirische Pädagogik 24, S. 93–112.
- BLK–Bund-Länder Kommission für Bildungsplanung (1973): Bildungsgesamtplan. Bonn: Bund-Länder Kommission für Bildungsplanung.

- BMBF–Bundesministerium für Bildung und Forschung (2003): Verwaltungsvereinbarung Investitionsprogramm „Zukunft Bildung und Betreuung“ 2003-2007. Verfügbar unter www.ganztagsschulen.org/_media/20030512_verwaltungsvereinbarung_zukunft_bildung_und_betreuung.pdf (Abruf 12.10.2018).
- BMBF–Bundesministerium für Bildung und Forschung (2005): Ergänzende Information zur Verwaltungsvereinbarung Investitionsprogramm „Zukunft Bildung und Betreuung“: Kostenneutrale Verlängerung des Förderzeitraums. Verfügbar unter www.ganztagsschulen.org/_media/izbb_ergaenzende_info.pdf (Abruf 12.10.2018).
- BMBF–Bundesministerium für Bildung und Forschung (2012): Ganztägig bilden: Eine Forschungsbilanz. Verfügbar unter www.ganztagsschulen.org/_media/121206_BMBF_GTS-Forschungsbilanz_bf_df.pdf (Abruf 12.10.2018).
- BMFSFJ–Bundesministerium für Familie, Senioren, Frauen und Jugend (2017): 15. Kinder- und Jugendbericht: Bericht über die Lebenssituation junger Menschen und die Leistungen der Kinder- und Jugendhilfe in Deutschland. Verfügbar unter www.bmfsfj.de/blob/115438/d7ed644e1b7fac4f9266191459903c62/15-kinder-und-jugendbericht-bundestagsdrucksache-data.pdf (Abruf 12.10.2018).
- Börner, N. (2011): Ganztagschule und ihre Auswirkungen auf Familien. In: Zeitschrift für Erziehungswissenschaft 14, S. 221–236.
- Böttcher, W. (2015): Ganze Tage in der Schule: Politik und Wissenschaft zwischen Anspruch und Wirklichkeit. In: Hascher, T./Idel, S./Reh, S./Thole, W. (Hrsg.): Bildung über den ganzen Tag. Forschungs- und Theorieperspektiven der Erziehungswissenschaft. Opladen, Berlin und Toronto: Barbara Budrich Publishers, S. 39–53.
- BPB–Bundeszentrale für politische Bildung (2016): Datenreport 2012: Ein Sozialbericht für die Bundesrepublik Deutschland. Bonn: Bundeszentrale für politische Bildung.
- Brümmer, F./Rollett, W./Fischer, N. (2011): Prozessqualität der Ganztagsangebote aus Schülersicht: Zusammenhänge mit Angebots- und Schulmerkmalen. In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 162–186.
- Dahrendorf, R. (1966): Bildung ist Bürgerrecht: Plädoyer für eine aktive Bildungspolitik. Hamburg: Nannen.

- Deutscher Bildungsrat (1968): Einrichtung von Schulversuchen mit Ganztagschulen: Empfehlungen der Bildungskommission. Bonn: Deutscher Bildungsrat.
- Deutscher Bildungsrat (1969): Einrichtung von Schulversuchen mit Gesamtschulen: Empfehlungen der Bildungskommission. Stuttgart: Klett.
- DIPF–Deutsches Institut für Internationale Pädagogische Forschung (2018): StEG: Die Teilprojekte der Studie. www.projekt-steg.de/content/steg-die-teilprojekte-der-studie (Abruf 12.10.2018).
- Drossel, K./Wendt, H./Schmitz, S./Eickelmann, B. (2012): Merkmale der Lehr- und Lernbedingungen im Primarbereich. In: Bos, W./Wendt, H./Köller, O./Selter, C. (Hrsg.): TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster, New York, München und Berlin: Waxmann, S. 171–202.
- Fischer, N./Brümmer, F./Kuhn, H. P. (2011): Entwicklung von Wohlbefinden und motivationalen Orientierungen in der Ganztagschule: Zusammenhänge mit der Prozess- und Beziehungsqualität in den Angeboten. In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 227–245.
- Fischer, N./Kuhn, H. P./Züchner, I. (2011): Entwicklung von Sozialverhalten in der Ganztagschule: Wirkungen der Ganztagsbeteiligung und der Angebotsqualität In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 246–266.
- Fischer, N./Sauerwein, M. N./Theis, D./Wolgast, A. (2016): Vom Lesenlernen in der Ganztagschule: Leisten Ganztagsangebote einen Beitrag zur Leseförderung am Beginn der Sekundarstufe I? In: Zeitschrift für Pädagogik 62, S. 780–796.
- Fischer, N./Theis, D. (2014): Quality of extracurricular activities – Considering developmental changes in the impact on school attachment and achievement. In: Journal for Educational Research Online 6, S. 54–75.
- Gambaro, L./Marcus, J./Peter, F. (2016): School entry, afternoon care and mother's labour supply. DIW Discussion Papers No. 1622.

- Gebhardt, B. (1993): Die Tagesschule der DDR. Betrachtungen zum sozialistischen Konzept der Ganztagserschulung. In: Zeitschrift für Pädagogik 39, S. 991–1006.
- Hagemann, K. (2009): Die Ganztagschule als Politikum. Die bundesdeutsche Entwicklung in gesellschafts- und geschlechtergeschichtlicher Perspektive. In: Zeitschrift für Pädagogik 54, S. 209–229.
- Hertel, S./Klieme, E./Radisch, F./Steinert, B. (2008): Nachmittagsangebote im Sekundarbereich und ihre Nutzung durch die Schülerinnen und Schüler. In: PISA-Konsortium Deutschland (Hrsg.): PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich. Münster, New York, München und Berlin: Waxmann, S. 297–318.
- Höhmann, K./Bergmann, K./Gebauer, M. (2008): Das Personal. In: Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L. (Hrsg.): Ganztagschule in Deutschland. Ergebnisse der Ausgangserhebung der „Studie zur Entwicklung von Ganztagschulen“ (StEG). 2. korrigierte Auflage. Weinheim und München: Juventa, S. 77–85.
- Höhmann, K./Grewe, M./Strietholt, R. (2008): Gründung und Ausstattung. In: Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L. (Hrsg.): Ganztagschule in Deutschland. Ergebnisse der Ausgangserhebung der „Studie zur Entwicklung von Ganztagschulen“ (StEG). 2. korrigierte Auflage. Weinheim und München: Juventa, S. 70–76.
- Höhmann, K./Holtappels, H. G./Schnitzer, T. (2004): Ganztagschule: Konzeptionen, Forschungsbefunde, aktuelle Entwicklungen. In: Holtappels, H. G./Klemm, K./Pfeiffer, H./Rolff, H.-G./Schulz-Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung. Daten, Beispiele und Perspektiven. Band 13. Weinheim: Juventa, S. 253–290.
- Holtappels, H. G. (2006): Stichwort: Ganztagschule. In: Zeitschrift für Erziehungswissenschaft 9, S. 5–29.
- Holtappels, H. G. (2008a): Angebotsstruktur, Schülerteilnahme und Ausbaugrad ganztägiger Schulen. In: Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L. (Hrsg.): Ganztagschule in Deutschland. Ergebnisse der Ausgangserhebung der „Studie zur Entwicklung von Ganztagschulen“ (StEG). 2. korrigierte Auflage. Weinheim und München: Juventa, S. 186–206.
- Holtappels, H. G. (2008b): Ziele, Konzepte, Entwicklungsprozesse. In: Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L. (Hrsg.): Ganztagschule in Deutschland.

- Ergebnisse der Ausgangserhebung der „Studie zur Entwicklung von Ganztagschulen“ (StEG). 2. korrigierte Auflage. Weinheim und München: Juventa, S. 139–163.
- Holtappels, H. G./Jarsinski, S./Rollett, W. (2011): Teilnahme als Qualitätsmerkmal für Ganztagschulen: Entwicklung von Schülerbeteiligungquoten auf Schulebene. In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 97–119.
- Holtappels, H. G./Rollett, W. (2008): Organisationskultur, Entwicklung und Ganztagschulausbau. In: Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L. (Hrsg.): Ganztagschule in Deutschland. Ergebnisse der Ausgangserhebung der „Studie zur Entwicklung von Ganztagschulen“ (StEG). 2. korrigierte Auflage. Weinheim und München: Juventa, S. 209–226.
- Ipfling, H.-J./Lorenz, U. (1979): Schulversuche mit Ganztagschulen. Schulversuche und Bildungsforschung Bd. 28. Mainz: v. Hase & Koehler.
- KMK–Sekretariat der Ständigen Konferenz der Kultusminister der Länder (2002): PISA 2000 – Zentrale Handlungsfelder: Zusammenfassende Darstellung der laufenden und geplanten Maßnahmen in den Ländern (Stand 07.10.2002). Verfügbar unter www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2002/2002_10_07-Pisa-2000-Zentrale-Handlungsfelder.pdf (Abruf 12.10.2018).
- KMK–Sekretariat der Ständigen Konferenz der Kultusminister der Länder (2008): Allgemein bildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: Statistik 2002 bis 2006. Bonn: Sekretariat der Ständigen Konferenz der Kultusminister der Länder.
- KMK–Sekretariat der Ständigen Konferenz der Kultusminister der Länder (2013): Allgemein bildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: Statistik 2007 bis 2011. Berlin: Sekretariat der Ständigen Konferenz der Kultusminister der Länder.
- KMK–Sekretariat der Ständigen Konferenz der Kultusminister der Länder (2015): Ganztagschulen in Deutschland: Bericht der Kultusministerkonferenz vom 03.12.2015. Berlin, Bonn: Sekretariat der Ständigen Konferenz der Kultusminister der Länder.

- KMK–Sekretariat der Ständigen Konferenz der Kultusminister der Länder (2018): Allgemein bildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: Statistik 2012 bis 2016. Berlin: Sekretariat der Ständigen Konferenz der Kultusminister der Länder.
- Koalitionsparteien (2018): Ein neuer Aufbruch für Europa. Eine neue Dynamik für Deutschland. Ein neuer Zusammenhalt für unser Land: Koalitionsvertrag zwischen CDU, CSU und SPD. Verfügbar unter www.cdu.de/system/tdf/media/dokumente/koalitionsvertrag_2018.pdf?file=1 (Abruf 12.10.2018).
- Kuhn, H. P./Fischer, N. (2011): Entwicklung der Schulnoten in der Ganztagschule: Einflüsse der Ganztagsbeteiligung und der Angebotsqualität. In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 207–226.
- Linberg, T./Struck, O./Bäumer, T. (2018): Vorzug Ganztagschule? Zusammenhänge mit der Kompetenzentwicklung im Bereich Lesen und Mathematik. In: Zeitschrift für Erziehungswissenschaft 14, S. 1–23.
- Lokhande, M./Hoeft, M./Wendt, H. (2014): Eltern als Bildungspartner: Wie Beteiligung an Grundschulen gelingen kann. Berlin: Sachverständigenrat deutscher Stiftungen für Integration und Migration.
- Lossen, K./Tillmann, K./Holtappels, H. G./Rollett, W./Hannemann, J. (2016): Entwicklung der naturwissenschaftlichen Kompetenzen und des sachunterrichtsbezogenen Selbstkonzepts bei Schüler/-innen in Ganztagsgrundschulen: Ergebnisse der Längsschnittstudie StEG-P zu Effekten der Schülerteilnahme und der Angebotsqualität. In: Zeitschrift für Pädagogik 62, S. 760–779.
- Mattes, M. (2009): Ganztagsbildung in der DDR. „Tagesschule“ und Hort in den Politiken und Diskursen der 1950er bis 1970er Jahre. In: Zeitschrift für Pädagogik 54, S. 230–246.
- Miller, B. M./Truong, K. A. (2009): The role of afterschool and summer in achievement. The untapped power of afterschool and summer to advance student achievement. In: Zeitschrift für Pädagogik 54, S. 124–142.

- OECD–Organisation for Economic Co-operation and Development (2011): Quality time for students: Learning in and out of school: OECD Publishing.
- Ottweiler, O. (2005): Die Positionen von Parteien, Verbänden und Kirchen zur Ganztagschule. In: Ladenthin, V. (Hrsg.): Die Ganztagschule. Weinheim: Juventa, S. 177–198.
- Picht, G. (1964): Die deutsche Bildungskatastrophe: Analyse und Dokumentation. Olten: Walter-Verlag.
- Plantenga, J./Remery, C. (2013): Childcare services for school age children: A comparative review of 33 countries. Luxembourg: Publications Office of the European Union.
- Prein, G./Rauschenbach, T./Züchner, I. (2009): Analysen zur Selektivität von offenen Ganztagschulen. In: Prüß, F./Kortas, S./Schöpa, M. (Hrsg.): Die Ganztagschule: von der Theorie zur Praxis. Anforderungen und Perspektiven für Erziehungswissenschaften und Schulentwicklung. Weinheim: Juventa, S. 81–97.
- Radisch, F. (2009): Qualität und Wirkung ganztägiger Schulorganisation: Theoretische und empirische Befunde. Studien zur ganztägigen Bildung. Weinheim: Juventa.
- Rauschenbach, T./Arnoldt, B./Steiner, C./Stolz, H.-J. (2012): Ganztagschule als Hoffnungsträger für die Zukunft? Ein Reformprojekt auf dem Prüfstand. 2. Auflage. Gütersloh: Verlag Bertelsmann Stiftung.
- Rollett, W./Lossen, K./Jarsinski, S./Lüpschen, N./Holtappels, H. G. (2011): Außerunterrichtliche Angebotsstruktur an Ganztagschulen: Entwicklungstrends und Entwicklungsbedingungen. In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 76–96.
- Spillebeen, L./Holtappels, H. G./Rollett, W. (2011): Schulentwicklungsprozesse an Ganztagschulen: Effekte schulischer Entwicklungsarbeit im Längsschnitt. In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 120–138.
- Stecher, L./Maschke, S. (2013): Research on extended education in Germany - A general model with all-day schooling and private tutoring as two examples. In: International Journal for Research on Extended Education, IJREE 1, S. 31–52.

- Stecher, L./Radisch, F./Fischer, N./Klieme, E. (2007): Bildungsqualität außerunterrichtlicher Angebote in der Ganztagsschule. In: Zeitschrift für Soziologie der Erziehung und Sozialisation 27, S. 346–366.
- StEG–Studie zur Entwicklung von Ganztagsschulen (2013): Ganztagsschule 2012/2013: Deskriptive Befunde einer bundesweiten Befragung. Verfügbar unter www.projekt-steg.de/sites/default/files/Bundesbericht_Schulleiterbefragung_2012_13.pdf (Abruf 12.10.2018).
- StEG–Studie zur Entwicklung von Ganztagsschulen (2015): Ganztagsschule 2014/2015: Deskriptive Befunde einer bundesweiten Befragung. Verfügbar unter www.projekt-steg.de/sites/default/files/StEG_Bundesbericht%202015_final_0.pdf (Abruf 12.10.2018).
- StEG–Studie zur Entwicklung von Ganztagsschulen (2016): Ganztagsschule: Bildungsqualität und Wirkungen außerunterrichtlicher Angebote: Ergebnisse der Studie zur Entwicklung von Ganztagsschulen 2012-2015. Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung.
- Steiner, C. (2009): Mehr Chancengleichheit durch die Ganztagsschule? In: Zeitschrift für Pädagogik 54, S. 81–105.
- Steiner, C. (2011a): Ganztagsteilnahme und Klassenwiederholung. In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagsschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagsschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 187–206.
- Steiner, C. (2011b): Teilnahme am Ganztagsbetrieb: Zeitliche Entwicklung und mögliche Selektionseffekte. In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagsschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagsschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 57–75.
- Steinmann, I./Strietholt, R. (im Druck): Effekte nachmittäglicher Bildungsangebote auf sprachliche Kompetenzentwicklungen. In: Zeitschrift für Pädagogik.
- Steinmann, I./Strietholt, R./Caro, D. (im Druck): Participation in extracurricular activities and student achievement: Evidence from German all-day schools. In: School Effectiveness and School Improvement.

- Strietholt, R./Manitius, V./Berkemeyer, N./Bos, W. (2015): Bildung und Bildungsungleichheit an Halb- und Ganztagschulen. In: Zeitschrift für Erziehungswissenschaft 18, S. 737–761.
- Tillmann, K. (2011): Innerschulische Kooperation und Schulprogramm: Zur Bedeutung des Schulprogramms als Schulentwicklungsinstrument. In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 139–161.
- Tillmann, K.-J. (2011): Die Steuerung von Ganztagschulen. In: Zeitschrift für Erziehungswissenschaft 14, S. 11–24.
- Wendt, H./Goy, M./Walzebug, A./Valtin, R. (2016): Bildungsangebote an Ganz- und Halbtagsgrundschulen in Deutschland. In: Wendt, H./Bos, W./Selter, C./Köller, O./Schwippert, K./Kasper, D. (Hrsg.): TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster, New York: Waxmann Verlag, S. 225–246.
- Willems, A. S./Wendt, H./Gröhlich, C./Walzebug, A./Bos, W. (2014): Mehr Chancengerechtigkeit durch die Ganztagschule? Ein Vergleich selektionsbedingter Ungleichheiten in der Bildungsteilhabe bei Schülerinnen und Schülern an Ganz- und Halbtagsgrundschulen auf Basis der IGLU und TIMSS 2011 Daten. In: Holtappels, H. G./ Willems, A. S./Pfeifer, M./Bos, W./McElvany, N. (Hrsg.): Jahrbuch der Schulentwicklung. Daten, Beispiele und Perspektiven. Band 18. Weinheim und Basel: Beltz Juventa, S. 62–100.
- Wissenschaftlicher Beirat für Familienfragen (2006): Ganztagschule: Eine Chance für Familien: Gutachten für das Bundesministerium für Familie, Senioren, Frauen und Jugend. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Züchner, I. (2011): Ganztagschulen und Familienleben. In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 291–311.
- Züchner, I./Arnoldt, B. (2011): Schulische und außerschulische Freizeit- und Bildungsaktivitäten: Teilhabe und Wechselwirkungen. In: Fischer, N./Holtappels, H.

G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 267–290.

Züchner, I./Fischer, N. (2011): Ganztagschulentwicklung und Ganztagschulforschung: Eine Einleitung. In: Fischer, N./Holtappels, H. G./Klieme, E./Rauschenbach, T./Stecher, L./Züchner, I. (Hrsg.): Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG). Weinheim und Basel: Beltz Juventa, S. 9–17.

Züchner, I./Fischer, N. (2014): Kompensatorische Wirkungen von Ganztagschulen – Ist die Ganztagschule ein Instrument zur Entkopplung des Zusammenhangs von sozialer Herkunft und Bildungserfolg? In: Zeitschrift für Erziehungswissenschaft 17, S. 349–367.

2.2 Beitrag II: Effekte nachmittäglicher Bildungsangebote auf sprachliche Kompetenzentwicklungen

Isa Steinmann & Rolf Strietholt

Die im Folgenden abgedruckte Version des Manuskripts ist das Preprint des Beitrags in der Zeitschrift für Pädagogik, der im zweiten Heft in 2019 erscheinen wird.

Deutsche Zusammenfassung

Der flächendeckende Ganztagschulausbau wurde Anfang der 2000er Jahre initiiert, um Missständen im deutschen Schulsystem entgegenzuwirken. Zu diesem Zeitpunkt gab es jedoch keine empirische Evidenz zu Effekten auf Leistungsniveaus oder Bildungsungleichheiten. Daher untersucht die vorliegende Studie, ob damals bestehende, lernförderlich ausgerichtete Nachmittagsangebote einen Einfluss auf Leistungsmaße hatten. Hierzu nutzen wir Daten der repräsentativen Längsschnittstudie DESI, in der die Entwicklung sprachlicher Kompetenzen in Deutsch und Englisch im Verlauf der neunten Jahrgangsstufe untersucht wurden. Unsere Ergebnisse bieten keine Evidenz dafür, dass zu Beginn des Angebotsausbaus davon ausgegangen werden konnte, dass Nachmittagsangebote das Leistungsniveau oder soziale Bildungsungleichheiten beeinflusst hätten.

Englische Zusammenfassung

The massive expansion of all-day schooling was initiated in the early 2000s to improve educational outcomes in the German school system. Empirical evidence on effects of all-day schooling on achievement levels or educational inequalities did not exist at this time. Against this background, the present study uses data from the early 2000s to investigate effects of all-day schooling on educational outcomes. We use data from the representative longitudinal DESI study that assessed student performance in German and English as a second language in the course of the ninth grade. Our results do not support that afternoon programs affected achievement levels or educational inequalities at the beginning of the expansion of all-day schooling in Germany.

Effekte nachmittäglicher Bildungsangebote auf sprachliche Kompetenzentwicklungen

1. Einleitung

Die Forderung nach einer evidenzbasierten Bildungs- und Sozialpolitik ist nicht neu, ihr wird aber in vielen Bereichen kaum oder verspätet nachgegangen. Das trifft in der Bildungspolitik auch auf Reformen der Ganztagsbeschulung zu. Bereits Ende der 1960er Jahre empfahl der Deutsche Bildungsrat (1968) Ganztagschulen in Modellversuchen zu erproben und wissenschaftlich zu evaluieren, um wahrgenommenen Missständen im Bildungswesen entgegenzuwirken (z.B. Dahrendorf, 1966; Picht, 1964). Während in den Folgejahren tatsächlich einige Schulen als Ganztagschulen geführt wurden, wurde die Forderung nach einer aussagekräftigen und ergebnisbezogenen Evaluation nicht oder nur in unzureichendem Maß erfüllt. Die Aussagekraft der wenigen Untersuchungen der 1970er Jahre ist designbedingt beschränkt, da die Reformen untrennbar mit Gesamtschulreformen (Deutscher Bildungsrat, 1969) verknüpft waren. Darüber hinaus wurden keine Kompetenzmaße erhoben und keine Kontrollgruppen von Halbtagschulen beobachtet (Fend, 1982; Ipfling & Lorenz, 1979).

Anfang der 2000er Jahre kam es zu einer weiteren, umfangreicheren Ausbaureform. Zwischen 2003 und 2009 wurde der Ausbau ganztägiger schulischer Bildung und Betreuung durch das mehr als vier Milliarden schwere Investitionsprogramm Zukunft Bildung und Betreuung (IZBB) bundesweit vorangetrieben (BMBF, 2003a, 2003b). War zu Beginn des Ausbaus nur jede sechste Schule in Deutschland als Ganztagschule organisiert, so sind es heute mehr als 60% (KMK, 2006, 2016). Der Ganztagschulausbau wurde unter anderem mit den in PISA 2000 (Programme for International Student Assessment) konstatierten niedrigen Leistungsniveaus und der ausgeprägten sozialen Bildungsungleichheit begründet. Die Ergebnisse der PISA-Studie wurden sogar dahingehend interpretiert, dass sie Ganztagsreformen nahegelegt hätten (KMK, 2002), wohingegen die Autor/-innen der Studie unterstreichen, dass Schlüsse über die Effektivität institutioneller Determinanten von Schulleistung designbedingt nicht abzuleiten sind (Baumert & Schümer, 2001; Schümer, 2001). Die Forderung nach einer ergebnisbezogenen Evaluation ganztägiger Bildung wurde auch im Rahmen der wissenschaftlichen Begleitevaluation des IZBB kaum erfüllt. Die erste Förderphase der Studie zur Entwicklung von Ganztagschulen (StEG) beleuchtete zwar Bedingungen und Prozesse an Ganztagschulen, es wurden jedoch keine standardisierten Leistungsmaße erfasst und eine Vergleichsgruppe von Halbtagschulen fehlte.

Wenngleich vor kurzem einige wenige Studien zur Effektivität von Ganztagschulen für Lernergebnisse von Schüler/-innen vorgelegt wurden, so mangelt es weiterhin an belastbaren Forschungsergebnissen zur Effektivität. Vor diesem Hintergrund betrachten wir anhand von Daten der Längsschnittstudie DESI (Deutsch-Englisch-Schülerleistungen-International) die Effekte von Nachmittagsangeboten auf die Entwicklung sprachlicher Kompetenzen im Verlauf der neunten Jahrgangsstufe zu dem Zeitpunkt, als der flächendeckende Ausbau begann. Im Folgenden skizzieren wir zunächst Annahmen zur Wirksamkeit schulischer Nachmittagsangebote und zur Ausgestaltungssituation in Deutschland. Danach resümieren wir empirische Studien, die Effekte ganztägiger Beschulung auf das Lernen von Schüler/-innen thematisieren, um im Anschluss die Daten, Methoden und Ergebnisse unserer Analysen vorzustellen und zu diskutieren.

1.1 Annahmen zur Wirksamkeit nachmittäglicher schulischer Bildungsangebote

An Ganztagschulen stehen im Vergleich zu Halbtagschulen mehr zeitliche und personelle Spielräume zur Verfügung, die für Zugewinne an Lernzeit (beispielsweise in Form fachlicher Zusatzangebote) oder an Freizeitelementen genutzt werden können. Solche zeitlich erweiterten und inhaltlich breiteren Angebotsstrukturen werden als unter anderem für die Lernmotivation von Schüler/-innen bedeutsam diskutiert. Lehrkräfte und das weitere im Ganztags pädagogisch tätige Personal sollen durch intensivere Einblicke Förderbedarfe besser aufdecken oder Talente fördern können. Insbesondere für Schüler/-innen aus Haushalten mit geringeren Unterstützungsmöglichkeiten soll eine schulische Förderung fehlende familiäre Ressourcen, wie eine intensive elterliche Unterstützung oder private Nachhilfe, kompensieren. Des Weiteren müssen nach dem Vormittagsunterricht stattfindende Angebote nicht nur additiv sein; vielmehr kann der gesamte Schultag aus alternierenden Lern- und Freizeitphasen bestehen, was als Rhythmisierung bezeichnet wird. Eine rhythmisierte Ganztagsbeschulung wird häufig als Idealmodell diskutiert (Überblicke bieten: BMBF, 2003a; Höhmann, Holtappels & Schnetzer, 2004; Holtappels, 2006; Züchner & Fischer, 2014).

Gleichzeitig kann nicht erwartet werden, dass sich solche förderlichen Effekte von Nachmittagsangeboten automatisch einstellen. Es wird vielmehr angenommen, dass die Wirksamkeit von weitreichenden Bedingungsgefügen abhängt, die zumindest implizit die Grundlage für bildungsbezogene Annahmen darstellen. Wenn durch Nachmittagsangebote Leistung gefördert und soziale Bildungsungleichheit kompensiert werden sollen, müssen die

Angebote auch hierauf ausgerichtet sein. In der wissenschaftlichen Debatte werden verschiedene Prozessmerkmale ganztägiger Beschulung diskutiert, zu denen allgemeine Charakteristika der Ganztagschulen (z.B. Angebotsarten, Rhythmisierung des Schultags), aber auch weitere Faktoren zählen (z.B. tatsächliche Nutzung des Angebots in Dauer und Umfang). Lernförderliche und kompensatorische Effekte auf Ebene der Schüler/-innen sind nur zu erwarten, wenn ein günstiges Zusammenspiel solcher Qualitätsmerkmale besteht (vgl. Fischer & Klieme, 2013; Holtappels & Rollett, 2008; Radisch, 2009).

Inwiefern Ganztagsmodelle in Bezug auf Bildung einen Mehrwert gegenüber Halbtagsmodellen bilden, hängt von der Ausgestaltung ab. In diesem Kontext sei darauf hingewiesen, dass die Kultusministerkonferenz (KMK, 2006) und das IZBB-Förderprogramm wenig darüber aussagen, wie Ganztag qualitativ ausgestaltet werden soll. Inwiefern die Begründungs- und Handlungsebene des Ganztagschulenausbaus übereinstimmen, ist demnach nicht gesichert, sondern eine empirische Frage.

1.2 Ausgestaltung des Ganztagsbetriebs

Frühe Einblicke zur Ganztagsausgestaltung bietet eine Befragung von 663 Ganztagschulleitungen in 15 Bundesländern, in der Durchführungsvariationen des Ganztagsbetriebs für das Schuljahr 2003/2004 erfragt wurden. Dabei waren 45% der Schulen voll gebundene Ganztagschulen (alle Schüler/-innen zur Ganztagesteilnahme verpflichtet), 17% teilweise gebunden (ein Teil des Ganztagsangebots für Gruppen von Schüler/-innen verpflichtend) und 38% offen (freiwillige Teilnahme). Sowohl innerhalb als auch zwischen diesen Organisationsformen variierten Faktoren, die als Bedingungen für eine Leistungsförderlichkeit ganztägiger Beschulung diskutiert werden. So boten die Schulen vielfältige Angebote an, die in verschiedenem Maß leistungsförderlich ausgerichtet waren. Solche auf eine Leistungsförderung ausgelegten Programme waren eher breit vertreten, so wiesen z.B. 75% der Ganztagschulen Arbeitsgemeinschaften und 77% eine Hausaufgabenbetreuung auf. Während erstere primär von Lehrkräften geleitet wurden, waren an letzteren vermehrt auch Honorarkräfte beteiligt (Höhmman et al., 2004).

Detailliertere, mehrperspektivische und zeitlich nachgelagerte Befunde gehen aus der StEG-Studie hervor. In den Erhebungen zwischen 2005 und 2009 zeigte sich einerseits ebenfalls, dass Qualitätsmerkmale der prozessualen Ausgestaltung des nachmittäglichen Lernraums zwischen den Ganztagschulen variierten (Fischer et al., 2011). An die Annahmen zu qualitativen Bedingungen eines Förderungsmehrwerts von Ganztagschulen anschließend

sind andererseits Befunde zur Angebotsinanspruchnahme von Interesse. Beispielsweise nahmen in 2005 lediglich 32% der Ganztagschüler/-innen der Sekundarstufe I an einer Hausaufgabenbetreuung und 25% an fachbezogenen Angeboten teil, was sich auch bis 2009 nur geringfügig veränderte (Holtappels, Jarsinski & Rollett, 2011).

1.3 Forschungsstand zu Effekten auf Leistungs- und Bildungsungleichheitsmaße

Wenn die Einführung eines Reformprojekts damit begründet wird, dass auf Outcome-Ebene bestimmte Effekte erwartet werden, wird häufig gefordert, dass sich diese Erwartungen erstens auf Evidenz stützen und zweitens eine Outcome-Evaluation stattfindet (vgl. Bellmann, 2016; Böttcher, 2015). Wie schon erwähnt wurde diese Forderung auch in Bezug auf den Ganztagsausbau bereits 1968 durch den Deutschen Bildungsrat formuliert. Umso erstaunlicher scheint, dass es im deutschsprachigen Raum bislang kaum empirische Evidenz für Effekte ganztägiger Beschulung auf Lernergebnisse gibt und die wenigen belastbaren Studien hierzu *nach* Implementation des IZBB vorgelegt wurden (vgl. Radisch, 2009; Radisch & Klieme, 2003). Die Erwartung einer Leistungsförderlichkeit von Ganztagschulen war demnach zumindest zum Zeitpunkt des IZBB-Beschlusses nicht empirisch gestützt. In Hinblick auf die Outcome-Evaluation des Reformprojekts wurde zwar die wissenschaftliche Begleitstudie StEG durchgeführt, jedoch konnten in der ersten Projektphase designbedingt nur Leistungseffekte auf Schulnoten untersucht werden, nicht auf standardisierte Maße. Darüber hinaus wurden keine Halbtagsschulen untersucht, sodass kein Organisationsformvergleich möglich war.

Nichtsdestotrotz wurden seit Beginn des Ausbaus im Jahr 2003 einige Studien zum Vergleich halb- und ganztägiger Schulen publiziert. In diesen Studien wurden sowohl das Leistungsniveau als auch herkunftsbedingte Bildungsungleichheiten querschnittlich untersucht. Hierzu wurden Daten der repräsentativen Schulleistungsstudien IGLU (Internationale Grundschul-Lese-Untersuchung), TIMSS (Trends in International Mathematics and Science Study) und PISA analysiert (Hertel, Klieme, Radisch & Steinert, 2008; Holtappels, Radisch, Rollett & Kowoll, 2010; Radisch, Klieme & Bos, 2006; Willems, Wendt & Radisch, 2015). Diese Studien fanden entweder keine Unterschiede zwischen Halb- und Ganztagschulen oder geringere Leistungen an Ganztagschulen. Auf Ebene der Schüler/-innen war die Teilnahme an Ganztagsangeboten teilweise mit niedrigeren Testergebnissen assoziiert. Es ist schwierig diese Befunde zu interpretieren, da aufgrund der Querschnittsdesigns unklar ist, inwiefern Selektionseffekte auf Individual- und Schulebene die

Korrelationen beeinflussen. Um derartige Selektionseffekte statistisch zu kontrollieren, haben Strietholt und Kolleg/-innen (2015) eine umfangreiche Analyse von Daten aus IGLU 2006, PISA 2006 und TIMSS 2007 durchgeführt und mithilfe von Matchingverfahren die Schulkomposition, die Schulform, das Bundesland sowie Approximationen für die Vorleistung kontrolliert (in TIMSS und IGLU Elterneinschätzungen zu vorschulischen Fähigkeiten und in PISA der Anteil an Jugendlichen mit Gymnasialempfehlung). Insgesamt zeigten sich weder in der Primar- noch in der Sekundarstufe Unterschiede im Leistungsniveau oder im Ausmaß der Bildungsungleichheiten an Halb- und Ganztagschulen.

Eine andere Gruppe an Studien nutzt längsschnittliche Designs, um kausale Effekte zu schätzen und vergleicht teilnehmende und nichtteilnehmende Schüler/-innen innerhalb von Ganztagschulen. Bellin und Tamke (2010) haben in Berlin die Leseleistungen von 604 Schüler/-innen in der ersten und zweiten Klasse getestet. Innerhalb von 39 Klassen an 23 offenen Ganztagsgrundschulen wurden am Ganzttag teilnehmende und nichtteilnehmende Kinder untersucht. Im Ergebnis zeigten sich nach Kontrolle der Vorleistung und kognitiven Grundfähigkeit (beides in der ersten Klasse erfasst) keine signifikanten Unterschiede in der Leistung in der zweiten Klasse. Des Weiteren wurden keine kompensatorischen Effekte für Kinder nichtdeutscher Herkunftssprache gefunden.

In der deutschsprachigen Schweiz wurden 295 Kinder an 35 Grundschulen von der ersten bis zur dritten Klasse begleitet (Schübach, 2015). Die Studie berichtet einen statistisch signifikanten Effekt der Teilnahme am Ganzttag (standardisiertes Regressionsgewicht von 0.18) auf die Leistungsentwicklung in Mathematik. Gleichzeitig muss darauf hingewiesen werden, dass nur jedes zehnte Kind der Stichprobe an Ganztagsangeboten teilnahm und diese Gruppe einen um ca. eine halbe Standardabweichung höheren sozioökonomischen Status aufgewiesen hat. Da dieser Unterschied in den Analysen nicht kontrolliert wurde, ist es möglich, dass die positiven Effekte der Ganzttagsteilnahme auf Selektionseffekte zurückzuführen sind.

Die beiden in Bezug auf das Studiendesign und die Stichprobengröße belastbarsten Studien nutzen Daten der StEG-Nachfolgestudien, die in den Jahren 2012-2015 durchgeführt wurden. Die erste Studie untersucht die Leistungsentwicklung im Leseverstehen zwischen Beginn des ersten und zweiten Schulhalbjahres in der fünften Jahrgangsstufe (2105 Schüler/-innen an 127 Schulen). In der Studie finden sich keine allgemeinen leseleistungssteigernden Effekte der Teilnahme (Fischer, Sauerwein, Theis & Wolgast, 2016). In weiteren Analysen zeigte die *freiwillige* Teilnahme an den Ganztagsangeboten jedoch einen signifikant positiven

Zusammenhang mit der Leistungsentwicklung. Die zweite StEG-Nachfolgestudie betrachtet die Entwicklung naturwissenschaftlicher Kompetenzen bei 2149 Kindern aus 67 Ganztagsgrundschulen vom Ende der dritten bis zum Ende der vierten Klasse. Es werden statistisch signifikante Teilnahmeeffekte domänenspezifischer Lernangebote auf die Entwicklung der Naturwissenschaftskompetenz berichtet, wobei der Effekt sehr gering ist (standardisiertes Regressionsgewicht von 0.05; Lossen, Tillmann, Holtappels, Rollett & Hannemann, 2016).

Anders als in Deutschland gibt es in den USA eine ganze Reihe von Untersuchungen, die sich mit der Effektivität von nachmittäglichen Bildungsangeboten beschäftigen. Evaluationen von staatenübergreifenden Initiativen und Ferienprogrammen zeigen insgesamt leistungsförderliche und bildungsungleichheitsabbauende Effekte (Afterschool Alliance, 2014; Riley & Peterson, 2013). Auch in Metaanalysen wurden förderliche Effekte über Jahrgangsstufen und Programmtypen hinweg gefunden, insbesondere für leistungsschwache Schüler/-innen (Durlak, Weissberg & Pachan, 2010; Feldman & Matjasko, 2005). Allerdings sind die Programme, die unter dem Begriff „extended education“ zusammengefasst werden, nicht mit ganztägiger Beschulung in Deutschland gleichzusetzen. Sie umfassen vielmehr sehr intensive und sorgfältig implementierte Maßnahmen, in denen zumeist explizit bildungsbenachteiligte Zielgruppen gefördert werden.

Bei einer Gesamtbetrachtung des Forschungsstands erscheint zum einen zentral, dass zu Beginn des Ganztagschulausbaus keine belastbaren empirischen Belege für eine Leistungsförderlichkeit von Nachmittagsangeboten vorlagen. In den Folgejahren anschließende Befunde aus dem deutschsprachigen Raum sind außerdem weniger vielversprechend als US-amerikanische. Hierbei zeigen Studien zu den Effekten der Angebotsteilnahme auf Individualebene weitestgehend neutrale oder kleine leistungsförderliche Effekte. Bei Vergleichen von Halb- und Ganztagschulen lassen sich keine Hinweise auf eine Leistungsförderlichkeit oder den Abbau von Bildungsungleichheit nachweisen. Darüber hinaus lässt sich konstatieren, dass die Forschung zu Effekten ganztägiger Beschulung im deutschsprachigen Raum wenig ausgeprägt und von methodischen Einschränkungen betroffen ist. Während die Sekundäranalysen von Daten aus IGLU, TIMSS und PISA repräsentativ für deutsche Schüler/-innen am Ende der Primar- bzw. Sekundarstufe sind und die Kompetenzbereiche Mathematik, Lesen und Naturwissenschaften abdecken, so sind die Ergebnisse aufgrund des querschnittlichen Designs mit Vorsicht zu interpretieren. Demgegenüber beruhen die referierten Längsschnittstudien nicht auf repräsentativen

Stichproben und decken die genannten Kompetenzbereiche für die unterschiedlichen Altersstufen nicht umfassend ab. Bislang hat nur eine Studie (Bellin & Tamke, 2010) längsschnittlich untersucht, ob Nachmittagsangebote soziale Bildungsungleichheit (in Hinblick auf die Herkunftssprache) beeinflussen.

1.4 Forschungsfragen

Die vorliegende Studie untersucht Effekte von Nachmittagsangeboten auf sprachliche Kompetenzen im Verlauf der neunten Jahrgangsstufe. Hierbei betrachten wir die beiden folgenden Fragestellungen: Hat das schulische Anbieten von nachmittäglicher Förderung im Rahmen von Hausaufgabenbetreuung, fachlichen Arbeitsgemeinschaften und Ganztagsunterricht einen Effekt auf Deutsch- und Englischleistungen von Neuntklässler/-innen? Hat das schulische Anbieten dieser nachmittäglichen Förderung einen Effekt auf herkunftsbedingte Bildungsungleichheiten in den sprachlichen Leistungsmaßen? Für unsere Analysen nutzen wir für Deutschland repräsentative längsschnittliche Daten aus dem Schuljahr 2003/2004, das den Beginn des Ganztagschulbaus darstellt.

2. Methode

2.1 Stichprobe

Die Studie basiert auf Daten der Schulleistungsstudie DESI, in der sprachliche Kompetenzentwicklungen im Verlauf der neunten Jahrgangsstufe im Schuljahr 2003/2004 längsschnittlich untersucht wurden (Beck & Klieme, 2007; Klieme, 2012). Die Zufallsstichprobe ist repräsentativ für die Schüler/-innen aller Regelschultypen in Deutschland. Neben Kompetenztests wurden Hintergrundinformationen zu Lebens- und Lernbedingungen bei Schüler/-innen, Eltern, Lehrkräften und Schulleitungen erhoben. Die Datenerhebung fand an zwei Messzeitpunkten zu Beginn (September bis Oktober 2003) und Ende (Mai bis Juli 2004) des neunten Schuljahres statt. Pro Schule wurden zufällig zwei Klassen ausgewählt. Die Datensätze enthalten Informationen von $N=10543$ Schüler/-innen in 219 Schulen. Von 35 Schulen liegen keinerlei Daten zu Nachmittagsangeboten vor. Nach Ausschluss dieser Schulen reduziert sich die effektive Stichprobengröße auf $N=8958$ Neuntklässler/-innen in 184 Schulen. Fehlende Werte bei einzelnen Nachmittagsangebots- sowie weiteren Variablen wurden imputiert (s. Abschnitt 2.2.4).

2.2 Variablen

2.2.1 Abhängige Variablen: Leistung der Schüler/-innen

Testleistungen in der Sprachbewusstheit in Deutsch und der Textrekonstruktion in Englisch als Fremdsprache (C-Test) bilden die abhängigen Variablen. Beide Kompetenzbereiche wurden zu zwei Messzeitpunkten erfasst, um den Leistungszuwachs innerhalb der neunten Jahrgangsstufe abzubilden. Es wurde ein Design mit rotierten Testheften angewandt, wobei die Schüler/-innen zu beiden Zeitpunkten verschiedene Items bearbeiteten (Beck, Bundt & Gomolka, 2008).

Sprachbewusstheit wird in DESI als grammatisch korrekter und stilistisch hochwertiger Umgang mit der deutschen Sprache definiert. Es geht hierbei um die Kenntnis grammatischer Regeln sowie das ständige Abgleichen eigener und fremder Sprachnutzung mit stilistischen und grammatischen Normen. In den entsprechenden Aufgaben sollen Fehler innerhalb eines Texts identifiziert, eigene kurze Textbestandteile formuliert oder grammatisch korrekte Items identifiziert werden. Der Sprachbewusstheitstest umfasst 34 Items mit variierender Schwierigkeit und hat eine EAP/PV-Reliabilität von 0.79 (Eichler, 2008).

Der C-Test ist ein etabliertes Testinstrument, das den allgemeinen Sprachstand in der Fremdsprache Englisch erfasst. Dabei müssen in einem Lückentext einzelne oder mehrere Buchstaben rekonstruiert und ergänzt werden. In der verwendeten Fassung besteht der Test aus 29 Items mit einer EAP/PV-Reliabilität von 0.90 (Harsch & Schröder, 2008).

Für den Sprachbewusstheits- und den C-Test liegen jeweils fünf Plausible Values vor. Plausible Values sind Leistungsmaße mit denen Populationsparameter (z.B. Regressionsparameter) unverzerrt geschätzt werden können. Des Weiteren wird auf diese Weise der Messfehler der Leistungsmessung bei der Standardfehlerschätzung berücksichtigt (von Davier, Gonzalez & Mislevy, 2009). Statistiken der Testwerte sind in Tabelle 1 zusammengefasst und weisen in beiden Kompetenzbereichen auf einen substantiellen Leistungsanstieg hin. Die hohen Standardabweichungen in den Zuwachsraten sind ein Hinweis auf unterschiedliche Entwicklungsbedingungen der Schüler/-innen. Die Intra-Klassen-Korrelationen der Leistungen zum Ende der neunten Klasse liegen sowohl in Deutsch als auch Englisch bei $ICC=0.63$, was auf substantielle Leistungsunterschiede zwischen den Schulen hinweist (vgl. z.B. Schulformdifferenzen in Tabelle 1). Die Zusammenhänge zwischen den Testwerten in Deutsch und Englisch sind zu Beginn ($r=0.77$, $p<0.05$) und Ende der neunten Jahrgangsstufe ($r=0.81$, $p<0.05$) hoch, wohingegen der Zusammenhang der

entsprechenden Zuwachswerte deutlich geringer ausfällt ($r=0.19$, $p<0.05$) (Hartig, Jude & Wagner, 2008).

Tabelle 1 Deskriptive Statistiken der Schüler/-innen und Schulen

	Alle Schulen <i>M(SD)</i> oder %	Haupt- schule <i>M(SD)</i> oder %	Sonstige Schule <i>M(SD)</i> oder %	Real- schule <i>M(SD)</i> oder %	Gesamt- schule <i>M(SD)</i> oder %	Gym- nasium <i>M(SD)</i> oder %
<i>Schüler/-innen</i>	<i>N=8958</i>	<i>N=1257</i>	<i>N=805</i>	<i>N=2718</i>	<i>N=615</i>	<i>N=3563</i>
Deutsch Messzeitpunkt 1	469.22 (90.08)	391.29 (67.52)	429.87 (59.59)	479.96 (60.80)	404.23 (74.22)	555.14 (62.58)
Deutsch Messzeitpunkt 2	503.37 (95.51)	417.09 (67.34)	459.81 (68.10)	516.21 (61.57)	434.41 (87.24)	596.53 (59.24)
Deutsch Zuwachs	34.15 43.61)	25.80 43.51)	29.95 (49.46)	36.26 (41.54)	30.18 (46.70)	41.39 (40.49)
Englisch Messzeitpunkt 1	479.87 (92.93)	410.20 (63.32)	413.24 (64.60)	493.34 (63.12)	407.24 (70.84)	571.16 (63.73)
Englisch Messzeitpunkt 2	503.31 (97.71)	425.63 (69.91)	435.27 (68.38)	517.69 (65.31)	436.87 (75.91)	598.13 (67.15)
Englisch Zuwachs	23.43 (38.16)	15.43 (40.27)	22.03 (37.12)	24.36 (36.39)	29.62 (43.53)	26.97 (35.74)
Bücher	3.55 (1.46)	2.95 (1.39)	3.05 (1.33)	3.54 (1.39)	3.21 (1.46)	4.33 (1.27)
<i>Schulen</i>	<i>N=184</i>	<i>N=33</i>	<i>N=19</i>	<i>N=49</i>	<i>N=15</i>	<i>N=68</i>
Hausaufgabenbetreuung	48.9%	53.4%	41.7%	45.1%	74.6%	40.3%
AG Deutsch	60.9%	32.3%	49.7%	60.4%	70.8%	94.2%
AG Englisch	22.6%	6.2%	17.3%	22.1%	31.3%	40.3%
Ganztagsunterricht	15.0%	4.7%	8.3%	5.5%	61.5%	19.2%
L2Bücher	3.55 (0.64)	2.95 (0.41)	3.05 (0.26)	3.54 (0.28)	3.21 (0.56)	4.33 (0.32)

Anmerkung. Basierend auf imputierten Daten. Sonstige Schulen sind Mittelschulen, Sekundarschulen oder Regelschulen. L2Bücher in unstandardisierter Form angegeben.

2.2.2 Unabhängige Variablen: Nachmittagsangebot

Die Schulleitungen wurden am Ende der neunten Jahrgangsstufe gefragt, wie häufig ihre Schule verschiedene Nachmittagsangebote bereitstellt („Inwieweit gibt es an Ihrer Schule die folgenden Ganztagsangebote, die mit dem Unterricht am Vormittag konzeptionell verbunden sind und unter der Verantwortung der Schulleitung stehen?“; vgl. Steinert, 2005). Dabei wurde unter anderem nach den folgenden vier Angebotstypen gefragt:

- Hausaufgabenbetreuung
- Freiwillige fachbezogene Arbeitsgemeinschaften am Nachmittag in Deutsch (z.B. Theater, Schreibwerkstatt)
- Freiwillige fachbezogene Arbeitsgemeinschaften am Nachmittag in Englisch
- Unterricht am Vor- und Nachmittag in rhythmisierter Form (Ganztagsunterricht).

Die vier Antwortalternativen waren „gar nicht“, „ein- bis zweimal pro Woche“, „mindestens dreimal pro Woche“ und „an jedem Schultag“. Die Verteilungen dieser ordinalen Skalen unterschieden sich zwischen den Angebotsvariablen deutlich. Hausaufgabenbetreuung wurde von 51% der Schulen nicht angeboten, von 27% an ein bis zwei Tagen, von 18% an drei bis vier Tagen pro Woche und von 5% an allen Schultagen. AGs in Deutsch wurden von 39% der Schulen nicht und von 53% an ein bis zwei Tagen pro Woche angeboten (5% 3-4 Tage, 2% 5 Tage pro Woche). AGs in Englisch wurden an den meisten Schulen gar nicht angeboten (77%, 20% 1-2 Tage, 1% 3-4 Tage und an 1% 5 Tage die Woche). Auch Ganztagsunterricht wurde von den meisten Schulen nicht angeboten (85%, 6% 1-2 Tage, 2% 3-4 Tage und 7% 5 Tage pro Woche). Aufgrund dieser unausgeglichene Verteilung der Antwortkategorien wurden die letzten drei in den Hauptanalysen zusammengefasst (s. auch Abschnitt 3.2.1). Auf diese Weise unterscheiden wir für jeden Angebotstypen zwischen Schulen, die das entsprechende Angebot aufweisen und nicht aufweisen. Tabelle 1 zeigt zum einen diese prozentuale Häufigkeit des mindestens wöchentlichen Angebots für alle Schulen, sowie im Schulformvergleich. In den Angebotshäufigkeiten sind teils deutliche Schulformunterschiede festzustellen.

2.2.3 Weitere Variablen

In der vorliegenden Studie ist der soziale Hintergrund der Schüler/-innen auf Individual- sowie Schulebene bedeutsam. Auf Individualebene betrachten wir den Zusammenhang zwischen Hintergrund und Leistung als Maß für die soziale Ungleichheit. Auf Schulebene aggregieren wir den auf der Individualebene beobachteten Hintergrund und verwenden dieses Kompositionsmaß als Kontrollvariable. In den Hauptanalysen verwenden wir die Anzahl der zuhause vorhandenen Bücher als Indikator für den sozioökonomischen Status (Antwortmöglichkeiten: „keine“, „1-10“, „11-50“, „51-100“, „101-250“, „251-500“, „mehr als 500“; s.a. Abschnitt 3.2.3). Nach Aggregation auf Schulebene werden die Werte der Büchervariable z-standardisiert (L2Bücher). Die Variable L2Bücher dient als Maß für die sozioökonomische Schulkomposition, d.h. inwiefern die Schule von benachteiligten oder

privilegierten Schüler/-innen besucht wird.

Während ein großer Anteil von Gesamtschulen Nachmittagsangebote bereithält, fällt der Anteil an anderen Schulformen geringer aus (s. Kennwerte der weiteren Variablen in Tabelle 1). Da sich auch bei anderen Angebotsvariablen Schulformunterschiede finden, nehmen wir diese als Kontrollvariable in die Analysen auf. Die Schulform geht in Form von vier Dummy-Variablen ein, wobei Hauptschulen als Referenzkategorie verwendet werden.

2.2.4 Fehlende Werte

Aus dem Originaldatensatz werden Schulen ausgeschlossen, deren Schulleitungen keinerlei Angaben zur nachmittäglichen Förderung gemacht haben, da für deren Imputation keine sinnvollen Prädiktoren vorliegen. Im Rohdatensatz unterscheiden sich die 35 ausgeschlossenen Schulen in Hinblick auf aggregierte Plausible Values der Deutsch- und Englischleistungen zu beiden Messzeitpunkten, der Verteilung der Schulformen, sowie der sozioökonomischen Komposition nicht statistisch signifikant von den 184 eingeschlossenen Schulen. Daher gehen wir davon aus, dass unsere Analysen zumindest hierauf bezogen weiterhin repräsentativ für Deutschland sind.

Während auf Individualebene für alle Schüler/-innen Testwerte in beiden Kompetenzbereichen vorliegen, fehlen bei 4% die Angaben zur Büchervariablen. Auf Schulebene haben die vier Variablen zum Nachmittagsangebot nach dem zuvor beschriebenen Schulausschluss zwischen 2% und 6% fehlende Werte. Diese werden mithilfe von Predictive Mean-Matching separat auf Individual- und Schulebene fünffach imputiert.

Auf Schulebene werden die genannten und weitere Angebotsvariablen (Mittagessen, Nachmittagsangebote in Sport und Freizeit, Angebote im musisch-künstlerischen Bereich), die Schulform und L2Bücher in das Imputationsmodell einbezogen. Auf Individualebene berücksichtigen wir außerdem die Variablen Geschlecht, Geburtsjahr, Geburtsländer der Eltern und des/der Schüler/-in, höchstes Beschäftigungsverhältnis der Eltern (HISEI, Highest International Socio-Economic Index of Occupational Status), höchster Bildungsabschluss der Eltern, kulturelle Ressourcen im Elternhaus, kognitive Grundfähigkeit, Leistungen zu beiden Messzeitpunkten sowie deren Zuwachs in weiteren Leistungsdomänen (Lesekompetenz und Textproduktion in Deutsch). Auf Individualebene wurden alle Informationen außer HISEI (bei den Eltern erfragt) bei den Schüler/-innen erhoben.

2.3 Analyseverfahren

Um die Forschungsfragen zu untersuchen, schätzen wir Mehrebenenregressionsmodelle. Dabei ist y_{m2ij} die Leistung von Schüler/-in i in Schule j zum zweiten Messzeitpunkt (m2) am Ende der neunten Jahrgangsstufe. Diese Variable wird auf Individualebene auf die familiäre Hintergrundvariable $Bücher_{ij}$ sowie die Leistung zum ersten Messzeitpunkt (m1) y_{m1ij} zurückgeführt:

$$\gamma_{m2ij} = \beta_{0j} + \beta_{1j}Bücher_{ij} + \beta_{2sj}\gamma_{m1ij} + e_{ij} \quad (1)$$

Der Achsenabschnitt β_{0j} beschreibt das Leistungsniveau in Schule j zum Ende der neunten Jahrgangsstufe nach Kontrolle der sozialen Herkunft sowie der Leistung zum ersten Messzeitpunkt. Um den Einfluss der Nachmittagsangebote auf das Leistungsniveau zu schätzen, sagen wir diesen Parameter auf Schulebene durch die Angebotsvariablen vorher. Da die Angebotsvariablen nicht randomisiert wurden, werden die Dummy-Variablen für die Schulform sowie die standardisierte aggregierte Büchervariable als Maß für die soziale Schulkomposition zusätzlich als Kontrollvariablen in das Regressionsmodell aufgenommen:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Angebote_j + \gamma_{02}Schulform_j + \gamma_{03}L2Bücher_j + u_0 \quad (2)$$

Der Steigungsparameter β_{1s} beschreibt den Zusammenhang zwischen der Herkunft der Schüler/-innen und deren Leistung (sozialer Gradient). Um den Einfluss der Nachmittagsangebote auf soziale Bildungsungleichheit zu schätzen, modellieren wir diesen Steigungsparameter als Zufallsvariable und führen ihn auf die Angebotsvariablen zurück. Wiederum werden die Schulform und die aggregierte Büchervariable auf Schulebene kontrolliert:

$$\beta_{1s} = \gamma_{10} + \gamma_{11}Angebote_j + \gamma_{12}Schulform_j + \gamma_{13}L2Bücher_j + u_1 \quad (3)$$

Alle Analysen werden separat für die Leistungsdomänen und Angebotsvariablen sowie unter Berücksichtigung der Surveygewichte umgesetzt. Die Schätzungen erfolgen hierbei separat für die fünf Datensätze mit imputierten Werten; die Ergebnisse werden abschließend nach Rubins Regeln verknüpft (Rubin, 1987).

3. Ergebnisse

3.1 Ergebnisse der Hauptanalysen

Die Ergebnisse für den Kompetenzbereich Deutsch sind in Tabelle 2 zusammengefasst. In der ersten Spalte findet sich zunächst ein Basismodell, in dem weder die Deutschleistung zu Messzeitpunkt 1, noch die Schulform oder -komposition als Kontrollvariablen aufgenommen wurden. In diesem querschnittlichen Modell zeigte sich, dass die Deutschleistung am Ende der neunten Jahrgangsstufe an Schulen die Hausaufgabenbetreuung anbieten niedriger und an Schulen mit AG Deutsch höher ausfiel. Dass diese korrelativen Befunde auf Selektionsmechanismen zurückgehen wird bei Betrachtung der zweiten Spalte deutlich. In Modell 2 wurden sowohl die Deutschleistung zu Messzeitpunkt 1 als auch die Schulform und -komposition kontrolliert. In diesem Modell waren weder die Effekte von Hausaufgabenbetreuung, AG Deutsch, noch Ganztagsunterricht auf die Deutschleistung signifikant von null verschieden. Im Gegensatz zu den Angebotsvariablen waren die Schulkomposition (L2Bücher) sowie die Schulform statistisch signifikant mit dem Leistungszuwachs in der neunten Jahrgangsstufe assoziiert. Die höchsten Lernzuwächse wurden an Schulen mit einer positiv selektierten Schülerschaft sowie an Gymnasien und Realschulen beobachtet.

Tabelle 2 Ergebnisse der Mehrebenenmodelle für die Sprachbewusstheit in Deutsch

	Modell 1 <i>b(SE)</i>	Modell 2 <i>b(SE)</i>	Modell 3 <i>b(SE)</i>	Modell 4 <i>b(SE)</i>	Modell 5 <i>b(SE)</i>
<i>Feste Effekte</i>					
Achsenabschnitt	466.04* (9.84)	165.27* (10.60)	164.43* (10.47)	166.89* (10.47)	165.81* (10.08)
Deutsch Messzeitpunkt 1		0.66* (0.02)	0.66* (0.02)	0.66* (0.02)	0.66* (0.02)
Bücher	6.84* (1.16)	2.96 (1.92)	3.14 (1.86)	3.00 (1.90)	2.99 (1.80)
Hausaufgabenbetreuung	-33.36* (12.03)	3.73 (4.77)	2.97 (4.76)		
Hausaufgabenbetr.*Bücher	0.78 (1.22)	-0.18 (1.07)	0.04 (1.06)		
AG Deutsch	63.30* (12.46)	-4.10 (5.47)		-3.01 (5.22)	
AG Deutsch*Bücher	0.04 (1.26)	0.38 (1.06)		0.41 (1.07)	
Ganztagsunterricht	-8.28 (17.88)	0.66 (6.76)			1.14 (7.11)
Ganztagsunterricht*Bücher	1.65 (2.26)	1.13 (1.50)			1.14 (1.56)

Beiträge der Arbeit | Beitrag II

L2Bücher	14.80*	14.79*	14.51*	14.56*	
	(4.52)	(4.45)	(4.55)	(4.52)	
Sonstige Schule	5.73	4.93	5.28	4.57	
	(8.98)	(8.99)	(8.89)	(8.96)	
Realschule	30.88*	29.85*	30.56*	29.73*	
	(8.49)	(8.76)	(8.49)	(8.73)	
Gesamtschule	-4.19	-5.15	-3.64	-5.19	
	(9.49)	(8.60)	(8.50)	(9.92)	
Gymnasium	44.51*	42.47*	44.14*	42.11*	
	(11.20)	(11.21)	(11.23)	(11.27)	
Bücher*Sonstige Schule	4.08*	4.17*	4.07*	4.18*	
	(1.99)	(1.98)	(2.02)	(1.95)	
Bücher*Realschule	-0.71	-0.70	-0.79	-0.59	
	(1.90)	(1.96)	(1.91)	(1.93)	
Bücher*Gesamtschule	1.93	2.60	2.47	2.01	
	(2.43)	(2.44)	(2.42)	(2.45)	
Bücher*Gymnasium	-1.46	-1.27	-1.47	-1.23	
	(2.54)	(2.56)	(2.54)	(2.56)	
Bücher*L2Bücher	0.44	0.52	0.51	0.45	
	(1.17)	(1.13)	(1.14)	(1.14)	
<i>Zufallseffekte</i>					
σ^2_e	3679.40*	2009.94*	2009.46*	2009.69*	2010.29*
	(119.07)	(65.01)	(65.11)	(65.10)	(65.06)
σ^2_{u0}	4683.85*	304.62*	307.46*	309.60*	309.36*
	(507.79)	(46.34)	(47.58)	(47.18)	(47.56)
σ^2_{u1}	9.14	1.29	1.50	1.44	1.22
	(5.75)	(1.86)	(1.95)	(1.95)	(1.87)
<i>N</i> Schulen	184	184	184	184	184
<i>N</i> Schüler/-innen	8958	8958	8958	8958	8958

Anmerkung. Basierend auf imputierten Daten. Bei zweiseitigem Signifikanztest indiziert * $p < 0.05$. Abhängige Variablen sind die fünf Plausible Values der Sprachbewusstheit in Deutsch zum zweiten Messzeitpunkt. Sonstige Schulen sind Mittelschulen, Sekundarschulen oder Regelschulen. σ^2_e bezeichnet die Residualvarianz der individuellen Leistung, σ^2_{u0} die Residualvarianz des Achsenabschnitts und σ^2_{u1} die Residualvarianz des sozialen Gradienten.

Mit Blick auf die soziale Bildungsungleichheit untersuchten wir, ob die Stärke des Zusammenhangs zwischen der Leistung am Ende der neunten Jahrgangsstufe und der sozialen Herkunft durch die betrachteten Nachmittagsangebote beeinflusst wurde. Dabei betrachteten wir die kumulierte Ungleichheit am Ende der neunten Jahrgangsstufe nicht querschnittlich, sondern hielten die Leistung zum ersten Messzeitpunkt konstant, um bereits zu Beginn der neunten Jahrgangsstufe bestehende Ungleichheiten zu kontrollieren. In den Analysen waren Hauptschulen die Referenzkategorie und die Schulkomposition wurde standardisiert, sodass der Bücherparameter die erwartete soziale Bildungsungleichheit an Hauptschulen mit einer mittleren Schulkomposition und ohne Nachmittagsangebot schätzte. Zur Beantwortung der

Frage nach dem Effekt der Angebote auf soziale Bildungsungleichheit untersuchten wir deren Interaktion mit der Büchervariable. Die Interaktionsparameter unterschieden sich in keinem Modell statistisch signifikant von null.

Tabelle 3 Ergebnisse der Mehrebenenmodelle für den C-Test in Englisch

	Modell 1 <i>b(SE)</i>	Modell 2 <i>b(SE)</i>	Modell 3 <i>b(SE)</i>	Modell 4 <i>b(SE)</i>	Modell 5 <i>b(SE)</i>
<i>Feste Effekte</i>					
Achsenabschnitt	509.65* (8.92)	99.29* (9.34)	98.92* (9.21)	98.15* (9.29)	98.58* (9.40)
Englisch Messzeitpunkt 1		0.82* (0.01)	0.82* (0.01)	0.82* (0.01)	0.82* (0.01)
Bücher	3.32* (0.83)	-0.79 (1.42)	-0.66 (1.40)	-0.62 (1.41)	-0.74 (1.42)
Hausaufgabenbetreuung	-33.00* (12.13)	-0.94 (3.49)	-2.29 (3.21)		
Hausaufgabenbetr.*Bücher	1.11 (1.09)	-0.09 (0.76)	0.29 (0.70)		
AG Englisch	40.48* (14.43)	-1.92 (3.20)		-2.69 (3.18)	
AG Englisch*Bücher	1.09 (1.36)	0.75 (0.80)		0.85 (0.75)	
Ganztagsunterricht	-13.98 (18.01)	-7.90 (5.51)			-8.74 (5.11)
Ganztagsunterricht*Bücher	2.77 (1.67)	1.94 (1.29)			2.12 (1.17)
L2Bücher		8.80* (3.38)	8.45* (3.39)	8.71* (3.40)	8.79* (3.40)
Sonstige Schule		2.23 (7.15)	1.86 (7.25)	2.57 (7.10)	2.04 (7.25)
Realschule		16.21* (6.52)	16.50* (6.27)	16.87* (6.45)	16.12* (6.51)
Gesamtschule		9.60 (7.73)	5.10 (7.85)	5.41 (7.87)	9.52 (7.69)
Gymnasium		24.36* (7.95)	23.69* (7.95)	24.28* (8.07)	24.20* (7.99)
Bücher*Sonstige Schule		1.74 (1.79)	1.87 (1.79)	1.69 (1.76)	1.86 (1.78)
Bücher*Realschule		0.83 (1.63)	0.78 (1.56)	0.66 (1.60)	0.89 (1.61)
Bücher*Gesamtschule		1.00 (2.06)	2.16 (2.19)	1.99 (2.19)	1.05 (2.04)
Bücher*Gymnasium		1.79 (2.03)	2.02 (2.05)	1.79 (2.08)	1.92 (2.04)
Bücher*L2Bücher		-0.73 (0.85)	-0.63 (0.87)	-0.65 (0.86)	-0.71 (0.85)
<i>Zufallseffekte</i>					
σ^2_e	4015.38* (116.49)	1705.54* (51.83)	1796.62* (51.74)	1706.65* (51.74)	1705.59* (51.88)

Beiträge der Arbeit | Beitrag II

σ^2_{u0}	5812.32* (479.34)	110.33* (24.71)	108.75* (24.72)	109.97* (25.03)	110.73* (24.93)
σ^2_{u1}	7.24 (5.43)	0.47 (1.01)	0.54 (1.05)	0.44 (1.07)	0.51 (1.03)
<i>N</i> Schulen	184	184	184	184	184
<i>N</i> Schüler/-innen	8958	8958	8958	8958	8958

Anmerkung. Basierend auf imputierten Daten. Bei zweiseitigem Signifikanztest indiziert * $p < 0.05$. Abhängige Variablen sind die fünf Plausible Values des C-Tests in Englisch zum zweiten Messzeitpunkt. Sonstige Schulen sind Mittelschulen, Sekundarschulen oder Regelschulen. σ^2_e bezeichnet die Residualvarianz der individuellen Leistung, σ^2_{u0} die Residualvarianz des Achsenabschnitts und σ^2_{u1} die Residualvarianz des sozialen Gradienten.

Die Ergebnisse für die Englischleistung als abhängige Variable sind in Tabelle 3 abgetragen und ergaben in Bezug auf die Fragestellungen das gleiche Bild. Weder das Angebot von Hausaufgabenbetreuung, AG Englisch, noch Ganztagsunterricht zeigten signifikante Effekte auf die Englischleistung. Auch in Hinblick auf die Ungleichheitsfragestellungen fanden wir keine signifikanten Effekte der betrachteten Nachmittagsangebote.

In den Modellen 1 und 2 (Tabellen 2 und 3) wurden die Effekte der drei Angebotsvariablen in Gesamtmodellen gleichzeitig geschätzt, um etwaige Konfundierungen zwischen den Angeboten zu kontrollieren: 36% der Schulen boten sowohl Hausaufgabenbetreuung als auch AG Deutsch mindestens wöchentlich an (18% Hausaufgabenbetreuung und AG Englisch; 11% Hausaufgabenbetreuung und Ganztagsunterricht; 26% AG Deutsch und AG Englisch; 13% AG Deutsch und Ganztagsunterricht; 7% AG Englisch und Ganztagsunterricht). In den Modellen 3-5 betrachteten wir die Effekte der drei Angebotsvariablen zudem in separaten Modellen. Die Ergebnisse entsprachen inhaltlich den bereits berichteten Befunden: Die Angebotsvariablen zeigten in den längsschnittlichen Analysen weder Effekte auf die Leistungs-, noch auf die Ungleichheitsmaße in Deutsch oder Englisch.

3.2 Ergänzende Analysen

Eine Reihe weiterer Analysen bestätigen die Robustheit unserer Hauptergebnisse. Hierbei wurden die Angebotsvariablen alternativ kodiert, alternative Angebotsvariablen untersucht, der soziale Hintergrund mit anderen Indikatoren operationalisiert sowie eine empirische Spezifikation mit Zuwachswerten geschätzt.

3.2.1 Alternative Skalierung der Angebotsvariablen

Die Häufigkeit von Angeboten wurde in DESI mit einer ordinalen Antwortskala erfragt, wobei einzelne Antwortmöglichkeiten kaum genutzt wurden. Daher wurde in den Hauptanalysen zwischen Schulen unterschieden, die entsprechende Angebote gar nicht oder mindestens wöchentlich anboten. Etwaige Effekte könnten auch nur bei einem häufigeren Angebot beobachtbar sein. Daher replizierten wir die berichteten Analysen mit den Angebotsvariablen in Form von drei Dummy-Variablen mit der Kategorie „gar nicht“ als Referenz. Im Ergebnis fanden wir nahezu keine Effekte der Angebotsvariablen auf Leistungs- und Ungleichheitsmaße in Deutsch oder Englisch. Lediglich das Anbieten von Ganztagsunterricht an 1-2 Tagen pro Woche zeigte einen ungleichheitssteigernden Effekt von 2.92 Punkten in der Englischleistung ($p < 0.05$) (nicht aber das noch häufigere Angebot). Dieser Parameter unterschied sich zwar signifikant von null, entspricht im Verhältnis zur Standardabweichung des Englischtests zum zweiten Messzeitpunkt jedoch einer sehr kleinen Effektstärke von 0.03. Darüber hinaus muss der Parameter vorsichtig interpretiert werden, da lediglich 6% der Schulen an ein bis zwei Tagen pro Woche Ganztagsunterricht anboten.

3.2.2 Alternative Angebotsvariablen

Neben den in den Hauptanalysen untersuchten Angeboten wurden in DESI auch Nachmittagsangebote im Sport- und im Freizeitbereich sowie im musisch-künstlerischen Bereich (Chor, Orchester, Video) erfragt. Solche Angebote sind zwar nicht explizit auf eine Sprachförderung ausgelegt, können jedoch als implizite Lerngelegenheiten für die deutsche Sprache dienen. Obwohl geringere Effekte auf die Deutschleistung zu erwarten waren, replizierten wir die Hauptanalysen für Deutsch mit diesen beiden impliziten Angebotsvariablen. Die Antwortoptionen wurden ebenfalls in „gar nicht“ und „mindestens wöchentlich“ zusammengefasst. Im Ergebnis zeigten sich keine Effekte der Angebotsvariablen im Sport- und Freizeit- sowie im musisch-künstlerischen Bereich auf Leistungs- oder Ungleichheitsentwicklungen in der Sprachbewusstheit Deutsch.

3.2.3 Ungleichheitsindikatoren

Der familiäre Hintergrund kann vielfältig operationalisiert werden. Wie Jerrim und Micklewright (2014) herausstellen, kann sowohl die Wahl des Indikators als auch die Wahl der Informationsquelle (Schüler/-innen vs. Eltern) zu unterschiedlichen Befunden führen. Daher wurden die Analysen mit weiteren Variablen zum familiären Hintergrund repliziert.

Hierzu gehörte der HISEI (Elternfragebogen), wobei anzumerken ist, dass der Anteil fehlender Werte vor der Imputation bei 39% lag. Der durchschnittliche HISEI betrug 28.53 mit einer Standardabweichung von 14.08. Als weiteres Hintergrundmerkmal wurde die Muttersprache betrachtet, wobei wir zwischen Jugendlichen unterschieden, deren Muttersprache nicht ausschließlich (17%) oder ausschließlich Deutsch war (83%). Die Quote fehlender Werte lag vor der Imputation bei 9%. Die Analysen mit den Hintergrundmerkmalen HISEI und Muttersprache führten qualitativ zu denselben Ergebnissen. Die Nachmittagsangebote zeigten keinen Effekt auf die Entwicklung sprachlicher Bildungsungleichheiten im Verlauf der neunten Jahrgangsstufe.

3.2.4 Modellierung von Zuwachswerten

In den Hauptanalysen wurde die Leistung zum ersten Messzeitpunkt als Kontrollvariable in die Modelle aufgenommen. Ein anderer Ansatz der längsschnittlichen Datenmodellierung ist die Analyse von Zuwachswerten. Zuwachswerte sind definiert als die Differenz zwischen der Leistung zum zweiten und ersten Messzeitpunkt. Lord (1967) zeigt, dass beide Analysestrategien zu unterschiedlichen Ergebnissen führen können, wenn sich die Leistungen zum ersten Messzeitpunkt zwischen den untersuchten Gruppen unterscheiden (hier: Schulen mit und ohne Nachmittagsangeboten). In Anlehnung an Wright (2006) replizierten wir unsere Analysen daher mit Zuwachswerten (vgl. Tabelle 1). Im Ergebnis bestätigten sich unsere Hauptergebnisse: Die betrachteten Angebotsvariablen hatten weder auf die Differenzwerte in Deutsch und Englisch, noch auf die Ungleichheit statistisch signifikante Effekte.

4. Diskussion

Deutsche Schulen waren im vergangenen Jahrhundert primär halbtägig organisiert, während ganztägige Modelle auf kleinere Reformbewegungen beschränkt waren. Erst seit Beginn des aktuellen Jahrhunderts werden immer mehr Ganztagschulen eingerichtet und Halbtagschulen in ganztägige Formen umgewandelt. Hierfür ausschlaggebend war das etwa vier Milliarden schwere Investitionsprogramm Zukunft Bildung und Betreuung (IZBB, 2003 bis 2009). Diese Reform und die mit ihr verknüpften Investitionen fielen in eine Zeit der vielfach postulierten evidenzbasierten Bildungspolitik, bei der die Bildungssteuerung Forschungsergebnisse über effektive Interventionen und Reformen nutzen soll (vgl. Bellmann, 2016). Vor diesem Hintergrund haben wir uns der Frage gewidmet, ob die Analyse entsprechend zurückliegender Daten empirische Evidenz für diese Begründungslinie des

massiven Ausbaus ganztägiger Bildung zu Beginn des 21. Jahrhunderts bietet. Der besondere Wert der vorliegenden Arbeit liegt darin, dass wir anhand von Daten aus den Jahren 2003/2004 die Situation vor bzw. zu Beginn des massiven Ganztagsausbaus in Deutschland betrachten können. Für diesen Zeitpunkt liegt keine empirische Evidenz dafür vor, dass schulische Bildungsangebote am Nachmittag lernförderliche Effekte hatten: Wir finden nach Kontrolle der Vorleistung keine Effekte von Nachmittagsangeboten in den Bereichen Hausaufgabenbetreuung, fachlicher Arbeitsgemeinschaften oder Ganztagsunterricht auf das Leistungsniveau oder das Ausmaß an herkunftsbezogener Bildungsungleichheit in Deutsch und Englisch in der neunten Jahrgangsstufe.

Es ist naheliegend zu fragen, ob unsere Forschungsergebnisse auch für heutige ganztägige Bildungsangebote gelten oder ob mittlerweile Effekte auf den Leistungsstand und das Ausmaß sozialer Bildungsungleichheiten bestehen. Offensichtlich können wir diese weiterführende Frage nicht mit den von uns genutzten Daten beantworten. Allerdings fanden sich auch in empirischen Studien, in denen jüngere Daten verwendet wurden, keine positiven Leistungseffekte von Ganz- versus Halbtagsschulen (Strietholt et al., 2015) oder der Ganztagsbeteiligung (Bellin & Tamke, 2010; Fischer et al., 2016; Lossen et al., 2016). Des Weiteren sei an dieser Stelle darauf hingewiesen, dass Ganztagschulen aus pädagogischer Sicht qualitativ häufig nicht so ausgestaltet sind, dass Lernmehrwerte der Schüler/-innen erwartet werden können. Dies mag auch darin begründet liegen, dass das Ganztagschulausbauprogramm IZBB hauptsächlich Infrastruktur finanzierte, diese Förderung jedoch kaum an pädagogische Ausgestaltungskriterien knüpfte. Holtappels warnte bereits 2006 (S. 26) vor einer „Angebots-Qualitäts-Falle“, also einer quantitativen Steigerung niedrigqualitativer Gestaltungsformen von Ganztags. Insbesondere in Bezug auf die Reduktion von Bildungsungleichheit stellt sich zudem die Frage, ob nicht spezielle Förderprogramme für Bildungsbenachteiligte erfolgversprechendere Interventionen darstellen als ein allgemeiner Ausbau, bei dem Mitnahmeeffekte für bildungsprivilegierte Schüler/-innen zu erwarten sind.

Zusammenfassend lässt sich festhalten, dass sich für Deutschland weder zu Beginn, noch im späteren Verlauf des Ausbaus nachmittäglicher Betreuung und Förderung empirische Hinweise finden, dass ganztägige Modelle einen Lernmehrwert gegenüber halbtägigen bieten. Vor diesem Hintergrund steht der Ausbau von Ganztags im Widerspruch zu Forderungen evidenzgesteuerter oder -gestützter Bildungspolitik. Böttcher kritisiert die Ganztagschulausbaureform mit Blick auf die postulierten Bildungsziele sogar als kaum mehr als „eine ‚programmlose‘ Investition“ (2015, S. 50). Um Missverständnissen vorzubeugen sei

an dieser Stelle noch einmal explizit betont, dass wir nicht bestreiten wollen, dass ganztägige Bildung und Betreuung effektiv lernförderlich wirken *kann*. Wir verstehen die vorliegende Studie in Abgrenzung zu Modellprogrammen als Politikevaluation.

Während es aus pädagogischer Sicht unbefriedigend sein mag, dass ganztägige Beschulung scheinbar wenig Mehrwert für Leistungsentwicklungen bietet, so möchten wir an dieser Stelle unterstreichen, dass diese Perspektive die Effektivität von Ganztagschulen nur fragmentiert abbildet. Der Ganztagschulausbau hat nicht nur bildungs-, sondern auch arbeitsmarktbezogene Ziele. Durch die längere schulische Betreuung soll ebenfalls auf die steigende Quote von Ein-Eltern-Familien reagiert und die Erwerbstätigkeit von Eltern erleichtert werden (vgl. Holtappels, 2006; Radisch & Klieme, 2003). Mit Blick auf diese Ziele legen erste Analysen des sozioökonomischen Panels nahe, dass erweiterte Betreuungsmöglichkeiten im Vor- und Grundschulalter die Arbeitsmarktbeteiligung von Müttern steigern (Bauernschuster & Schlotter, 2015; Gambaro, Marcus & Peter, 2016), auch wenn die positiven Effekte nicht in allen Studien repliziert wurden (Dehos & Paul, 2017). Arbeitsmarkt- und bildungsbezogene Ziele von Ganztags stehen nicht unbedingt in Konkurrenz. Scheinbar ist für die erstgenannte Zieldimension jedoch die Betreuung von Kindern und Jugendlichen ausreichend, wohingegen das Erreichen bildungsbezogener Ziele insbesondere von der Ausgestaltung des Nachmittags abhängt.

Betrachtet man Befunde zu Ausgestaltungsformen von Ganztagschulen in Verbindung mit Überlegungen zu Wirksamkeitsbedingungen, dann lässt sich das Ausbleiben von Effekten mit einer zu geringen Qualität des Ganztags in Deutschland begründen. Im Kontrast dazu zeigen US-amerikanische Befunde, dass hochqualitative „extended education“ durchaus leistungsförderlich und ungleichheitskompensatorisch wirken kann. Augenscheinlich ist, dass einige der amerikanischen Programme deutlich spezifischer auf Lernen ausgerichtet sind und teilweise auch explizit benachteiligte Schüler/-innen als Zielgruppe haben. Das gilt nicht für das Breitenangebot in Deutschland. In diesem Zusammenhang sind Implementationsstudien zu begrüßen, in denen der Frage nachgegangen wird, *wie* Nachmittagsangebote in Deutschland lernförderlich gestaltet und umgesetzt werden können. Beispielsweise zu nennen sind die Projekte StEG-Lesen und StEG-Tandem, die zwischen 2016 und 2019 durchgeführt werden (DIPF, 2017).

4.1 Reichweite der Studie

Die vorliegende Studie ergänzt den Forschungsstand um Befunde aus dem Anfangsstadium des Ganztagsausbaus, für das bislang keine Wirksamkeitsstudien vorlagen. Darüber hinaus bildet sie für den Vergleich von Schulen mit und ohne Nachmittagsangebot die erste Studie die auf längsschnittliche Daten zurückgreift, wobei Neuntklässler/-innen an Regelschulen zu Beginn und Ende des Schuljahrs 2003/2004 untersucht wurden. Andere Längsschnittstudien untersuchten ausschließlich Teilnahmeeffekte innerhalb von Ganztagschulen und beschränkten sich auf Stichproben in der Grundschulzeit und frühen Sekundarstufe. Neben Effekten auf das Leistungsniveau betrachten wir die Entwicklung herkunftsbezogener Bildungsungleichheit, einen Aspekt der im deutschsprachigen Raum bislang kaum untersucht wurde.

Gleichzeitig sei darauf hingewiesen, dass die vorliegende Studie Beobachtungsdaten nutzt. Vergleiche zwischen Schulen mit und ohne Nachmittagsangebote basieren demnach auf natürlicher Variation innerhalb des deutschen Schulwesens. So konnten wir aufgrund der Tatsache, dass die Angebote durchschnittlich eher selten stattfanden nicht robust untersuchen, welche Effekte bei einem täglichen Angebot resultieren. Aus anderen Studien geht außerdem hervor, dass in dem Zeitraum auch innerhalb von Schulen mit einem Angebot weitere Heterogenität bestand (vgl. z.B. Höhmann et al., 2004). Beispielsweise wäre denkbar, dass zwar kein allgemeiner Leistungsvorteil an anbietenden Schulen zu finden war, aber an solchen, die ihr Nachmittagsprogramm ausschließlich mit qualifizierten Lehrkräften gestalteten, oder an Schulen, die eine sehr hohe Teilnahmequote erreichten. Allerdings handelt es sich hierbei zunächst nur um Hypothesen, die es empirisch zu prüfen gilt (vgl. z.B. Radisch, 2009).

4.2 Fazit

Die Befunde der vorliegenden Studie deuten darauf hin, dass sich Schulen in 2003/2004 zwar je nach Schulform und -komposition in ihrer Lernförderlichkeit in Deutsch und Englisch unterschieden, nicht jedoch danach, ob sie nachmittägliche Angebote bereitstellten. Dieser Befund mag vor dem Hintergrund der bisherigen empirischen Forschung sowie des weiteren Diskurses nicht überraschen, bleibt aber aus pädagogischer Sicht unbefriedigend (vgl. Hertel et al., 2008). Wenn man dieser Einschätzung folgt und evidenzbasierte Bildungspolitik betrieben werden soll, so bleibt eine alte Forderung aktuell:

„Die Bildungskommission des Deutschen Bildungsrates empfiehlt, Ganztagschulen als Versuchsschulen in einer für eine wissenschaftliche Kontrolle ausreichenden Zahl einzurichten“ (Deutscher Bildungsrat, 1968, S. 12).

5. Literaturverzeichnis

- Afterschool Alliance (2014). *Taking a deeper dive into afterschool: Positive outcomes and promising practices*, Washington, D.C.
- Bauernschuster, S. & Schlotter, M. (2015). Public child care and mothers' labor supply—Evidence from two quasi-experiments. *Journal of Public Economics*, 123, 1–16.
- Baumert, J. & Schümer, G. (2001). Lebens- und Lernbedingungen von Jugendlichen: Schulformen als selektionsbedingte Lernmilieus. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider et al. (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 427–454). Opladen: Leske + Budrich.
- Beck, B., Bundt, S., & Gomolka, J. (2008). Ziele und Anlage der DESI-Studie. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 11–25). Weinheim/Basel: Beltz.
- Beck, B. & Klieme, E. (Hrsg.) (2007). *Sprachliche Kompetenzen: Konzepte und Messung. DESI-Studie (Deutsch-Englisch-Schülerleistungen-International)*. Weinheim [u.a.]: Beltz.
- Bellin, N. & Tamke, F. (2010). Bessere Leistungen durch Teilnahme am offenen Ganztagsbetrieb? *Empirische Pädagogik*, 24(2), 93–112.
- Bellmann, J. (2016). Datengetrieben und/oder evidenzbasiert? *Zeitschrift für Erziehungswissenschaft*, 19(S1), 147–161.
- Böttcher, W. (2015). Ganze Tage in der Schule: Politik und Wissenschaft zwischen Anspruch und Wirklichkeit. In T. Hascher, T. S. Idel, S. Reh & W. Thole (Hrsg.), *Bildung über den ganzen Tag. Forschungs- und Theorieperspektiven der Erziehungswissenschaft* (S. 39–53). Opladen [u.a.]: Barbara Budrich Publishers.
- BMBF (2003a). *Investitionsprogramm „Zukunft Bildung und Betreuung“: Ganztagschulen. Zeit für mehr*, Bundesministerium für Bildung und Forschung.
http://www.bmbf.de/pub/ganztagschulen-zeit_fuer_mehr.pdf.
- BMBF (2003b). *Verwaltungsvereinbarung Investitionsprogramm „Zukunft Bildung und Betreuung“ 2003-2007*.

https://www.ganztagsschulen.org/_media/20030512_verwaltungsvereinbarung_zukunft_bildung_und_betreuung.pdf.

- Dahrendorf, R. (1966). *Bildung ist Bürgerrecht: Plädoyer für eine aktive Bildungspolitik*. Hamburg: Nannen.
- Dehos, F. & Paul, M. (March 2017). *The effects of after-school programs on maternal employment*. *Ruhr economic papers*: Bd. No. 686. Bochum: RUB.
- Deutscher Bildungsrat (1968). *Einrichtung von Schulversuchen mit Ganztagschulen: Empfehlungen der Bildungskommission*. Bonn.
- Deutscher Bildungsrat (1969). *Einrichtung von Schulversuchen mit Gesamtschulen: Empfehlungen der Bildungskommission*. Stuttgart: Klett.
- Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American journal of community psychology*, 45(3-4), 294–309.
- Eichler, W. (2008). Sprachbewusstheit Deutsch. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 112–119). Weinheim/Basel: Beltz.
- Feldman, A. F. & Matjasko, J. L. (2005). The Role of School-Based Extracurricular Activities in Adolescent Development: A Comprehensive Review and Future Directions. *Review of Educational Research*, 75(2), 159–210.
- Fend, H. (1982). *Gesamtschule im Vergleich: Bilanz der Ergebnisse des Gesamtschulbesuchs*. Weinheim: Beltz.
- Fischer, N., Holtappels, H. G., Klieme, E., Rauschenbach, T., Stecher, L., & Züchner, I. (Hrsg.) (2011). *Ganztagsschule: Entwicklung, Qualität, Wirkungen: Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG)*. Weinheim/Basel: Beltz Juventa.
- Fischer, N. & Klieme, E. (2013). Quality and effectiveness of German all-day schools. Results of the study on the development of all-day schools in Germany. In J. Ecarius, E. Klieme, L. Stecher & J. Woods (Hrsg.), *Extended education – an international perspective. Proceedings of the international conference on extracurricular and out-of-school time educational research* (S. 27–52). Opladen [u.a.]: Barbara Budrich Publishers.

- Fischer, N., Sauerwein, M. N., Theis, D., & Wolgast, A. (2016). Vom Lesenlernen in der Ganztagschule: Leisten Ganztagsangebote einen Beitrag zur Leseförderung am Beginn der Sekundarstufe I? *Zeitschrift für Pädagogik*, 62(6), 780–796.
- Gambaro, L., Marcus, J., & Peter, F. (2016). *School entry, afternoon care and mother's labour supply*: DIW Discussion Papers, No. 1622.
- Harsch, C. & Schröder, K. (2008). Textrekonstruktion Englisch. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 149–156). Weinheim/Basel: Beltz.
- Hartig, J., Jude, N., & Wagner, W. (2008). Methodische Grundlagen der Messung und Erklärung sprachlicher Kompetenzen. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 34–54). Weinheim/Basel: Beltz.
- Hertel, S., Klieme, E., Radisch, F., & Steinert, B. (2008). Nachmittagsangebote im Sekundarbereich und ihre Nutzung durch die Schülerinnen und Schüler. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme et al. (Hrsg.), *PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (S. 297–318). Münster [u.a.]: Waxmann.
- Hömann, K., Holtappels, H. G., & Schnetzer, T. (2004). Ganztagschule: Konzeptionen, Forschungsbefunde, aktuelle Entwicklungen. In H. G. Holtappels, K. Klemm, H. Pfeiffer, H.-G. Rolff & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung. Daten, Beispiele und Perspektiven* (S. 253–290). Weinheim: Juventa.
- Holtappels, H. G. (2006). Stichwort: Ganztagschule. *Zeitschrift für Erziehungswissenschaft*, 9(1), 5–29.
- Holtappels, H. G., Jarsinski, S., & Rollett, W. (2011). Teilnahme als Qualitätsmerkmal für Ganztagschulen: Entwicklung von Schülerteilnahmequoten auf Schulebene. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG)* (S. 97–119). Weinheim/Basel: Beltz Juventa.
- Holtappels, H. G., Radisch, F., Rollett, W., & Kowoll, M. E. (2010). Bildungsangebot und Schülerkompetenzen in Ganztagsgrundschulen. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried & E.-M. Lankes (Hrsg.), *IGLU 2006 – die Grundschule auf dem Prüfstand* (S. 165–198). Münster: Waxmann.

- Holtappels, H. G. & Rollett, W. (2008). Organisationskultur, Entwicklung und Ganztagschulausbau. In H. G. Holtappels, E. Klieme, T. Rauschenbach & L. Stecher (Hrsg.), *Ganztagschule in Deutschland. Ergebnisse der Ausgangserhebung der „Studie zur Entwicklung von Ganztagschulen“ (StEG)* (2. korrigierte Auflage) (S. 209–226). Weinheim/München: Juventa.
- Ipfling, H.-J. & Lorenz, U. (1979). *Schulversuche mit Ganztagschulen*. Schulversuche und Bildungsforschung: Bd. 28. Mainz: v. Hase & Koehler.
- Jerrim, J. & Micklewright, J. (2014). Socio-economic Gradients in Children's Cognitive Skills: Are Cross-Country Comparisons Robust to Who Reports Family Background? *European Sociological Review*, 30(6), 766–781.
- Klieme, E. (2012). *Deutsch-Englisch-Schülerleistungen-International (DESI)*. Version: 1. IQB – Institut zur Qualitätsentwicklung im Bildungswesen. Datensatz. http://doi.org/10.5159/IQB_DESI_v1.
- KMK (2002). *PISA 2000 - Zentrale Handlungsfelder: Zusammenfassende Darstellung der laufenden und geplanten Maßnahmen in den Ländern (Stand 07.10.2002)* (Beschluss der 299. Kultusministerkonferenz vom 17./18.10.2002).
- KMK (2006). *Bericht über die allgemein bildenden Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: 2002 bis 2004*, Bonn.
- KMK (2016). *Allgemeinbildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: Statistik 2010 bis 2014*, Berlin.
- Lord, F. M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin*, 72, 304–305.
- Lossen, K., Tillmann, K., Holtappels, H. G., Rollett, W., & Hannemann, J. (2016). Entwicklung der naturwissenschaftlichen Kompetenzen und des sachunterrichtsbezogenen Selbstkonzepts bei Schüler/-innen in Ganztagsgrundschulen: Ergebnisse der Längsschnittstudie StEG-P zu Effekten der Schülerteilnahme und der Angebotsqualität. *Zeitschrift für Pädagogik*, 62(6), 760–779.
- Picht, G. (1964). *Die deutsche Bildungskatastrophe: Analyse und Dokumentation*. Olten [u.a.]: Walter-Verlag.
- Radisch, F. (2009). *Qualität und Wirkung ganztägiger Schulorganisation: Theoretische und empirische Befunde*. Studien zur ganztägigen Bildung. Weinheim [u.a.]: Juventa.

- Radisch, F. & Klieme, E. (2003). *Wirkung ganztägiger Schulorganisation: Bilanzierung der Forschungslage. Literaturbericht im Rahmen von „Bildung Plus“*. Frankfurt, M.: DIPF.
- Radisch, F., Klieme, E., & Bos, W. (2006). Gestaltungsmerkmale und Effekte ganztägiger Angebote im Grundschulbereich. *Zeitschrift für Erziehungswissenschaft*, 9(1), 30–50.
- Riley, R. W. & Peterson, T. K. (2013). Conclusion: Leveraging the power of quality afterschool and summer learning: An important strategy for student success across America. In T. K. Peterson (Hrsg.), *Expanding minds and opportunities. Leveraging the power of afterschool and summer learning for student success*. Washington, D.C.: Collaborative Communications Group.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.
- Schüpbach, M. (2015). Effects of extracurricular activities and their quality on primary school-age students' achievement in mathematics in Switzerland. *School Effectiveness and School Improvement*, 26(2), 279–295.
- DIPF (2017). *StEG (2016-2019): Die Teilprojekte der Studie*. <http://www.projektsteg.de/content/steg-die-teilprojekte-der-studie>.
- Steinert, B. (2005). *Deutsch Englisch Schülerleistungen International. Dokumentation der Erhebungsinstrumente für Schulleiterinnen und Schulleiter, Mitglieder des Fachkollegiums Deutsch und Mitglieder des Fachkollegiums Englisch. (Material zur Bildungsforschung 14/1)*. Unveröffentlichtes Dokument, Frankfurt, M.: DIPF.
- Strietholt, R., Manitius, V., Berkemeyer, N., & Bos, W. (2015). Bildung und Bildungsungleichheit an Halb- und Ganztagschulen. *Zeitschrift für Erziehungswissenschaft*, 18(4), 737–761.
- von Davier, M., Gonzalez, E. J., & Mislevy, R. J. (2009). What are plausible values and why are they useful? *IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments* (Vol. 2, pp. 9-36). Hamburg/Princeton NJ: IEA-ETS Research Institute.
- Willems, A. S., Wendt, H., & Radisch, F. (2015). Domänenspezifische Kompetenzen und Chancengerechtigkeit im Vergleich von Ganz- und Halbtagsgrundschultypen: Zur Rolle individueller Herkunftsmerkmale und der Komposition der Schülerschaft. In H. Wendt, T. C. Stubbe, K. Schwippert & W. Bos (Hrsg.), *10 Jahre international vergleichende Schulleistungsforschung in der Grundschule. Vertiefende Analysen zu IGLU und TIMSS 2001 bis 2011* (S. 219–238). Münster/New York: Waxmann.

- Wright, D. B. (2006). Comparing groups in a before-after design: When t test and ANCOVA produce different results. *British Journal of Educational Psychology*, 76, 663–675.
- Züchner, I. & Fischer, N. (2014). Kompensatorische Wirkungen von Ganztagschulen – Ist die Ganztagschule ein Instrument zur Entkopplung des Zusammenhangs von sozialer Herkunft und Bildungserfolg? *Zeitschrift für Erziehungswissenschaft*, 17, 349–367.

2.3 Beitrag III: Student achievement and educational inequality in half- and all-day schools: Evidence from Germany

Isa Steinmann & Rolf Strietholt

Die im Folgenden abgedruckte Version des Manuskripts ist das Preprint des Beitrags, der sich im International Journal for Research on Extended Education in Druck befindet.

Abstract

Several countries have expanded extended education in recent years. In Germany, the most substantial educational reform is the ongoing transformation of the traditional half-day school system into an all-day school system. Among politicians, expectations are high that all-day schools will promote student achievement and reduce social achievement inequalities. To test these assumptions, we used representative data from the National Educational Panel Study (NEPS) to estimate two-level latent growth models for achievement in grades 5, 7, and 9. The analyses revealed initial achievement differences but no differences in achievement growth or changes in inequality throughout secondary school. This suggests that selection mechanisms are at work but that half- and all-day schools are not differentially effective. We discuss these findings in light of the international debate on the quality of extended education.

Student achievement and educational inequality in half- and all-day schools: Evidence from
Germany

1. Introduction and Research Question

Learning takes place in various contexts, which can be depicted on a continuum from informal to formal settings. Informal learning is not organized and takes place unintentionally and continuously in everyday life. By contrast, formal learning takes place in organized, highly structured contexts that are designated for learning (e.g., regular school lessons). Between these two poles lies non-formal learning (Werquin, 2010). Extended education—like private tutoring or extra-curricular activities at schools—is a group of non-formal contexts that is intended to promote learning and is pedagogically structured but less formalized than regular classes (see Stecher & Maschke, 2013). Especially school-based, non-formal extended education is increasingly politically relevant in many countries because it is (a) expected to improve student learning outcomes and (b) more open to external influence than, for instance, private tutoring (see Kuger, 2016; Plantenga & Remery, 2013; Vest, Mahoney, & Simpkins, 2013).

In Germany, the largest education policy reform of the past decades concerns school-based extended education. In the 20th century, the school day in most schools consisted of morning classes but no afternoon program.⁷ Between 2003 and 2009, a federal investment program of more than four billion euros (BMBF, 2003) prompted a large increase in the proportion of all-day schools from 16% in 2002 to 68% in 2016 (KMK, 2008, 2018). The program supported the founding of new all-day schools and the infrastructural development of existing ones. One aim behind the massive expansion of all-day schools was that the extended supervision of children would facilitate maternal employment (e.g., Fischer, Theis, & Züchner, 2014; Plantenga & Remery, 2013). However, politicians also expected the schools to increase opportunities to support learning and reduce educational gaps between social groups (e.g., BMBF, 2003; Fischer et al., 2014). The definition of all-day schools adopted by

⁷ In the 1970s, some initiatives started to establish “Gesamtschulen” in Germany. “Gesamtschulen” were a new school type that intended to overcome the traditional ability grouping in different school types and combine ability tracks within schools. Furthermore, they were organized as all-day schools. However, the share of those schools was small and they were limited to few federal states.

the investment program was: (a) All-day schools provide lunch and at least a seven-hour program on at least three days per week. (b) This program is supervised by (and conducted in close cooperation with) the school administration, which is also accountable for the program. (c) It is conceptually connected to the regular classes (KMK, 2008). Expectations were high that all-day schooling improves learning outcomes because all-day schools provide additional time and personnel resources that can be used to enhance schooling, for example, by providing more individual support and diverse learning activities (BMBF, 2003). Learning benefits were especially expected for socially disadvantaged students who have less stimulating environments in the afternoon at home and for children whose families do not speak German at home. Therefore, all-day schooling was expected to contribute to a reduction of social achievement gaps (e.g., BMBF, 2003; Steiner, 2009; Züchner & Fischer, 2014).

Based on the aforementioned all-day school definition, it is reasonable to expect that all-day schooling would extend formal learning opportunities. In fact, however, at all-day schools the regular morning classes are typically supplemented by non-formal components in the afternoon, not by additional formal instruction. Although the definition requires that the all-day program and regular classes are conceptually related, research indicates that this does not hold at about half of all-day schools (StEG, 2015). Generally, the definition gives schools considerable room for interpretation: It does not define qualitative characteristics such as the pedagogical content or the qualifications of the supervising staff. The schools decide whether participation in afternoon activities is nonmandatory, partially mandatory (e.g., for students in certain grades), or fully mandatory for all students. As a result, the organization of all-day schools varies within and between states, primary and secondary schools, and school types/academic tracks (KMK, 2015; StEG, 2015). In addition, all-day programs are often not supervised by teachers but by differently and often less qualified staff. Programs designed to promote student learning are much less common than those promoting leisure and sports activities (StEG, 2013, 2016). Moreover, the student enrollment rates are often low, even at schools with a broad all-day program, which sets limitations to the potential benefit of all-day over half-day schools. However, students with working parents, low socioeconomic status, or an immigrant background are especially likely to make use of all-day schooling, even though many schools charge fees for lunch and/or attending programs (StEG, 2016; Steiner, 2011). The observation that all-day programs seem to be able to reach disadvantaged groups reinforces the idea that all all-day schools may decrease social inequalities in achievement (cf. also Steiner, 2009; Steinmann, 2018; Züchner & Fischer, 2014).

The heterogeneous characteristics and implications of all-day schooling are especially relevant in light of models on the effectiveness of extended education, which highlight the importance of *how* extended time resources are used. Such models suggest that both program characteristics and student participation mediate effects on student outcomes. Among other things, the assumed quality prerequisites for effective extended education include the types of activities and their level of structure (e.g., curriculum, alignment with learning) as well as personnel resources (e.g., small staff-to-child ratios, high staff qualifications) and students' frequent and intensive participation (cf. Fischer & Klieme, 2013; Miller & Truong, 2009). The models suggest that extended education programs that are closer to the formal pole of the informal-formal learning continuum are more successful in promoting student learning. Therefore, one empirical question is whether non-formal learning provided in addition to regular schooling at all-day schools actually improves the learning outcomes of (groups of) students. In the present study, we investigated whether all-day schools are more successful in promoting student achievement and reducing educational inequality between social groups than traditional half-day schools.

2. Review of Literature

Despite the substantial recent investment in all-day schooling in Germany, few studies have evaluated its effects on achievement and inequality. Studies with robust research designs are particularly rare. One group of longitudinal studies compared students who participated in non-formal all-day programs with students who did not attend all-day programs. None of these studies found effects on student achievement after controlling for prior achievement and further background characteristics (Bellin & Tamke, 2010; Fischer, Sauerwein, Theis, & Wolgast, 2016; Linberg, Struck, & Bäumer, 2018; Lossen, Tillmann, Holtappels, Rollett, & Hannemann, 2016; Steinmann, Strietholt, & Caro, 2018). Bellin and Tamke (2010) further investigated if students with a migration background profit more from the all-day participation than native peers but found no support for this assumption. However, there are general issues with studies at the student level, resulting from possible spillover effects within schools that are related to remedial education measures (e.g., nonparticipating students may receive more attention during regular classes).

Studies that investigated all-day schooling as a school-level measure circumvent some issues related to selection mechanisms and spillover effects that operate at the individual level within schools. Only three studies used proxies or test measures of prior achievement as

controls to compare performance at half- and all-day schools or at schools with and without afternoon programs offering homework supervision and remedial courses. Again, none of these studies revealed student achievement effects (Linberg et al., 2018; Steinmann & Strietholt, in print; Strietholt, Manitius, Berkemeyer, & Bos, 2015). Two of the three studies investigated effects on educational inequality and found null results (Steinmann & Strietholt, in print; Strietholt et al., 2015). They examined for example inequalities in achievement scores between students with a high and a low social status or between students with and without German as first language. However, the longest investigated time span was two years. Furthermore, all these studies failed to take into account the ongoing changes from half- to all-day schools because they only determined the schools' organization form at one time point. This lack of precision likely led to biased effect estimates.

In order to contextualize the findings for Germany, we briefly summarize findings on the circumstances under which non-formal extended education programs are found to foster learning. In the US, extended education programs (e.g., afterschool programs, summer schools) have been studied extensively in experimental and quasi-experimental studies. Meta-analyses and literature reviews of this research suggest that some programs had no effects while others showed positive effects on student achievement. In contrast, positive effects were observed for programs with the following characteristics: They were designed to promote specific competences, they were closely linked to the regular curriculum, or they employed evidence-based educational approaches (Apsler, 2009; Cooper, Charlton, Valentine, Muhlenbruck, & Borman, 2000; Durlak, Weissberg, & Pachan, 2010; Lauer et al., 2006). Programs targeting at-risk students showed particularly positive effects (Durlak et al., 2010; Patall, Cooper, & Allen, 2010). Effective programs also employed highly qualified staff (Feldman & Matjasko, 2005; Lauer et al., 2006). However, programs that did not meet these quality characteristics showed smaller and often no effects on student outcomes (Apsler, 2009; Durlak et al., 2010; Roth, Malone, & Brooks-Gunn, 2010). In summary, programs that were located at the formal rather than the informal end of the learning continuum—i.e., that were more comparable to regular schooling—showed the most promising results. Additionally, some studies indicated that disadvantaged groups like students with a low socioeconomic status profit more from extended education, which implies that all-day schools reduce social achievement gaps (Lauer et al., 2006; Patall et al., 2010).

3. Hypotheses

The present study aimed to compare achievement levels and social inequalities in achievement in half- and all-day schools in Germany. Specifically, we investigated three research questions: Did the school's organization form have an impact on (a) reading and mathematics achievement, (b) achievement inequality regarding social status, and (c) achievement gaps between students who did and did not learn German as a first language? The political sphere expects all-day schooling to boost achievement levels and reduce educational inequality in achievement, while the scientific debate and previous findings cast doubt on these optimistic expectations. We used representative longitudinal large-scale data with test information for grades 5, 7, and 9 to investigate the three research questions.

4. Methods

4.1 Data

This paper used data from the National Educational Panel Study (NEPS; Blossfeld, Roßbach, & von Maurice, 2011)⁸. NEPS collected longitudinal data on achievement, educational processes, and educational organizations, as well as on returns to education for different age cohorts. In the present study, we focused on a sample of secondary school students who were first tested and surveyed in grade 5 (Frahm et al., 2011; Strietholt et al., 2013). We limited our analyses to the $N=3444$ fifth graders who remained in the survey, i.e. they stayed at the same 164 schools until grade 9. Further, we excluded 20 schools whose principals' did not specify their schools' organization form at any measurement point (for more information, see section 4.2.5). In total, we investigated a sample of $N=3024$ students at 144 regular schools in 15 federal states⁹ who were followed from grade 5 (school year 2010/11) until grade 9 (school

⁸ NEPS drew a representative random sample of fifth graders in the school year 2010/11 (Starting Cohort 3, doi:10.5157/NEPS:SC3:7.0.1). From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LifBi) at the University of Bamberg in cooperation with a nationwide network.

⁹ Berlin was excluded, because no school entirely covered grades 5–9. This is because the transition from primary to secondary schools in Berlin typically takes place between grades 6 and 7.

year 2014/15). They were on average 10.9 years old in grade 5 and 49% were female.

4.2 Variables

4.2.1 Student Achievement

Mathematics and reading achievement scores are available for grades 5, 7, and 9. The achievement scales are comparable over time because the paper-pencil-based tests were linked by the anchor-item design they employ (Fischer, Rohm, Gnamb, & Carstensen, 2016). We used longitudinally linked weighted maximum likelihood estimates.¹⁰ Table 1 shows roughly linear achievement increases between the measurement points. The overall increases between grades 5–9 corresponded to a bit more than one standard deviation. While there was some variance across measurement points in the reliability of the tests, the EAP/PV reliability was constantly high, ranging between .76 and .81 (Duchhardt & Gerdes, 2012; Krannich et al., 2017; Pohl, Haberkorn, Hardt, & Wiegand, 2012; Scharl, Fischer, Gnamb, & Rohm, 2017; Schnittjer & Gerken, 2017).¹¹

¹⁰ We rescaled all achievement scores by multiplying the original by 100 to improve the readability of the estimates in the main analyses.

¹¹ Reliability information on the mathematics test in grade 9 has not yet been published.

Table 1 Descriptive statistics for student and school characteristics

	All schools	Divided by school organization form		
		Half-day school	All-day school	Changed status
<i>Student characteristics</i>	<i>N=3024</i>	<i>N=731</i>	<i>N=1326</i>	<i>N=967</i>
Student achievement				
Mathematics grade 5 (<i>M(SD)</i>)	-1.67 (113.40)	-35.12 (120.51)	14.87 (110.92)	6.50 (104.94)
Mathematics grade 7 (<i>M(SD)</i>)	74.83 (122.68)	37.35 (129.96)	92.81 (118.08)	84.53 (115.48)
Mathematics grade 9 (<i>M(SD)</i>)	151.27 (116.96)	114.26 (118.47)	166.25 (117.66)	163.67 (108.84)
Reading grade 5 (<i>M(SD)</i>)	0.19 (122.64)	-31.94 (126.29)	12.63 (121.10)	11.54 (117.01)
Reading grade 7 (<i>M(SD)</i>)	70.44 (132.19)	39.40 (132.08)	77.84 (136.33)	86.07 (123.92)
Reading grade 9 (<i>M(SD)</i>)	125.25 (107.92)	98.64 (107.20)	136.65 (110.62)	133.53 (102.17)
Student background				
Parental education grade 5 (<i>M(SD)</i>)	2.68 (1.13)	2.47 (1.11)	2.77 (1.10)	2.69 (1.16)
Language of origin German grade 5	88.0%	78.9%	91.3%	90.4%
<i>School characteristics</i>	<i>N=144</i>	<i>N=43</i>	<i>N=56</i>	<i>N=45</i>
School type				
Hauptschule/Volksschule	27.8%	43.4%	17.7%	21.2%
School with several courses of education	13.8%	12.6%	12.0%	16.3%
Realschule	17.3%	16.6%	13.6%	20.9%
Gesamtschule	8.5%	7.6%	11.7%	6.9%
Gymnasium	32.6%	19.8%	45.0%	34.6%
School composition				
Social composition grade 5 (<i>M(SD)</i>)	2.60 (0.54)	2.35 (0.49)	2.77 (0.54)	2.71 (0.49)
Language composition grade 5 (<i>M(SD)</i>)	0.85 (0.21)	0.76 (0.25)	0.92 (0.12)	0.87 (0.19)

Note. The descriptive analyses were based on imputed data of $N=3024$ students at 144 schools; “w_t_cal” was used as a sampling weight. Unstandardized social and language composition variables were used.

4.2.2 Parental Education and Language of Origin

We investigated two facets of student background, which were both assessed in computer-assisted telephone interviews with parents in grade 5 (see Table 1). We operationalized social background as parental education level in line with the International Standard Classification

of Education. Information on parents' highest general educational qualification were categorized from 0 ("level 0/1A: Inadequately completed general education") to 5 ("level 6: Doctoral degree and postdoctoral lecture qualification"). We treated this variable as continuous. We operationalized immigrant background by the student's language of origin. Answers to the question regarding the language students learned in the first three years in the family were dichotomized to 0 ("other than German") and 1 ("German").

4.2.3 School Organization: Half- and All-Day Schools

The main explanatory variable was school organization form. In questionnaires in grades 5, 7, and 9, principals were asked whether their school was a half- or all-day school. In the questionnaires, all-day schools were further categorized into nonmandatory, partially mandatory, and fully mandatory all-day schools. The sample sizes were, however, too small to estimate the effects of schools that remained fully mandatory over time, for example. For this reason, we subsume them as all-day schools. Due to the massive investment in all-day schooling, several schools converted from half- to all-day schools in the period of investigation, while some also changed from all- to half-day schools (see also KMK, 2018; Steiner, 2011). For a clearer interpretation of our main explanatory variable, we categorized such schools as a separate group in our analyses. The sample covered 43 half-day schools, 56 all-day schools, and 45 schools with a mixed status over time. The organization form was analyzed as two dummy variables, with half-day schools as the reference category. Table 1 depicts the characteristics of these groups and shows that half-day school students constituted a less privileged group than those attending all-day schools.

4.2.4 Covariates

Germany has a stratified secondary school system with different ability tracks. The school types corresponding to those tracks were part of the explicit strata in the NEPS sampling design and were used as covariates in the form of four dummy variables ("school with several courses of education", "Realschule", "Gesamtschule", and "Gymnasium", with "Hauptschule/Volksschule" as reference). To control for differences in student composition, we aggregated the student background information on parental education (social composition) and the language of origin (language composition) at school level. Both variables were treated as continuous variables in the analyses (see Table 1).

4.2.5 Missing Value Imputation

The dataset used in this analysis contained missing data (see Appendices A1 and A2). Twenty schools whose principals did not respond to the organization form question at any of the three measurement points were excluded from the analyses because we regarded the variable base as too poor for sound imputation. Apart from this, we imputed missing values five times by using two-level predictive mean matching, which is a simple extension of ordinary multiple imputation technique for non-clustered data (van Buuren & Groothuis-Oudshoorn, 2011). In order to replicate the data structure, we included the sampling weight as an imputation predictor as well as a rich set of further variables (see Appendix A2): all aforementioned variables, complementary variables, their repetitions in other waves, and counterparts in other instruments. The Appendix A2 also depicts which variables lied on school and student levels in the two-level imputation. We reran all analyses for the five imputed datasets and combined the estimates using Rubin's rules (1987).

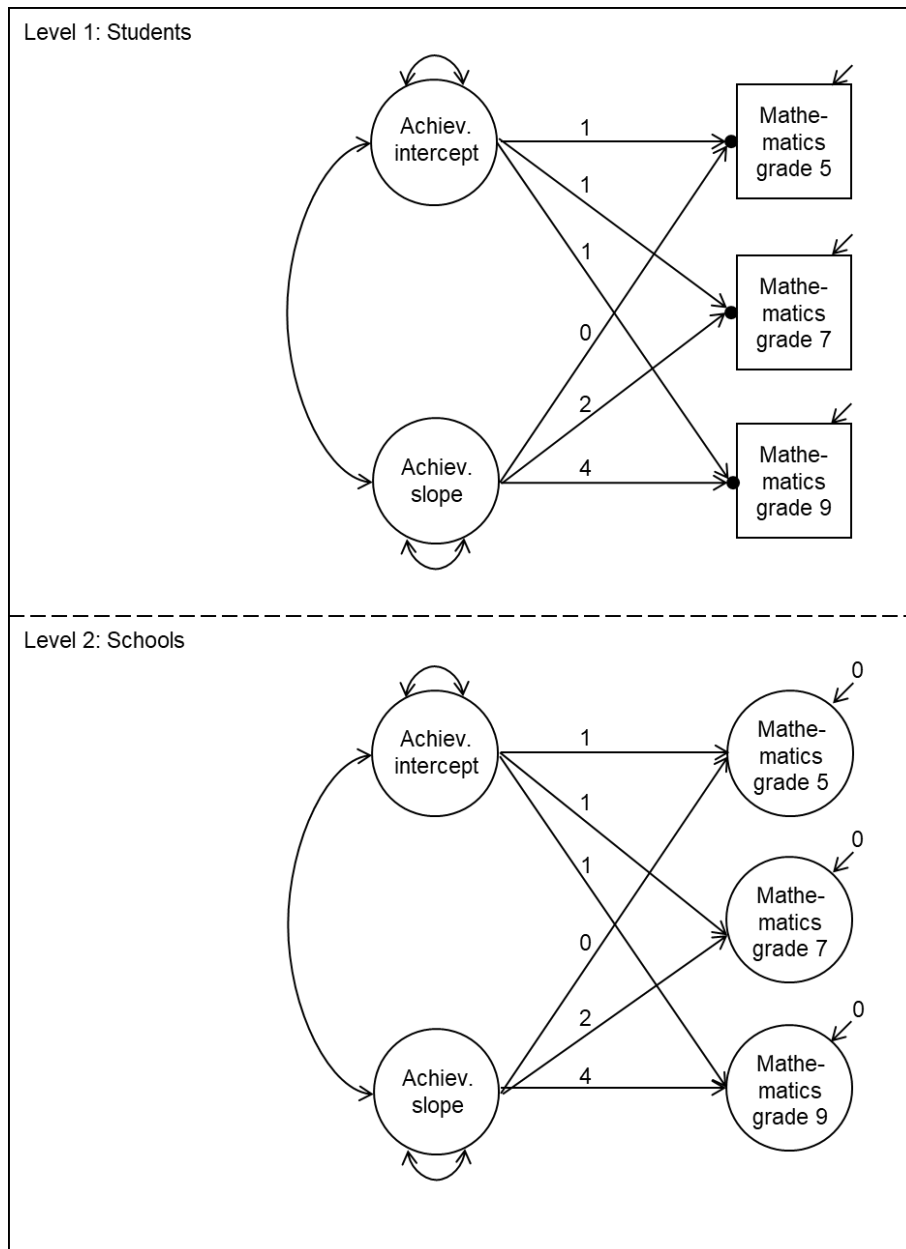
4.3 Analyses

To explain our general analytical approach, we will describe how we modeled achievement growth and changes in inequality in achievement before outlining how we tested for differences between half- and all-day schools. All analyses were replicated for mathematics and reading in separate models.

4.3.1 Modeling Effects on Student Achievement

Two-level linear latent growth modeling was used to investigate achievement growth. At the student level, achievement scores in grades 5, 7, and 9 were used to model an achievement intercept for achievement in grade 5 and an achievement slope for the annual growth in achievement (see Figure 1). As in the student-level models, for the school-level models the schools' achievement scores in grades 5, 7, and 9 were used to model an achievement intercept and slope. The school level achievement intercept reflects that schools could vary in their initial achievement in grade 5 and the achievement slope indicates that the schools could differ in their growth rates.

Figure 1 Measurement model for latent growth in mathematics achievement



Note. The two-level latent growth model shown in this figure was estimated; the dots in the individual-level graph represent random intercepts; the random intercepts are shown in circles in the school-level graph because they are continuous latent variables that vary across schools; school-level residuals are constrained to zero to avoid negative variance. The school-level achievement intercept and slope serve as dependent variables in further analyses.

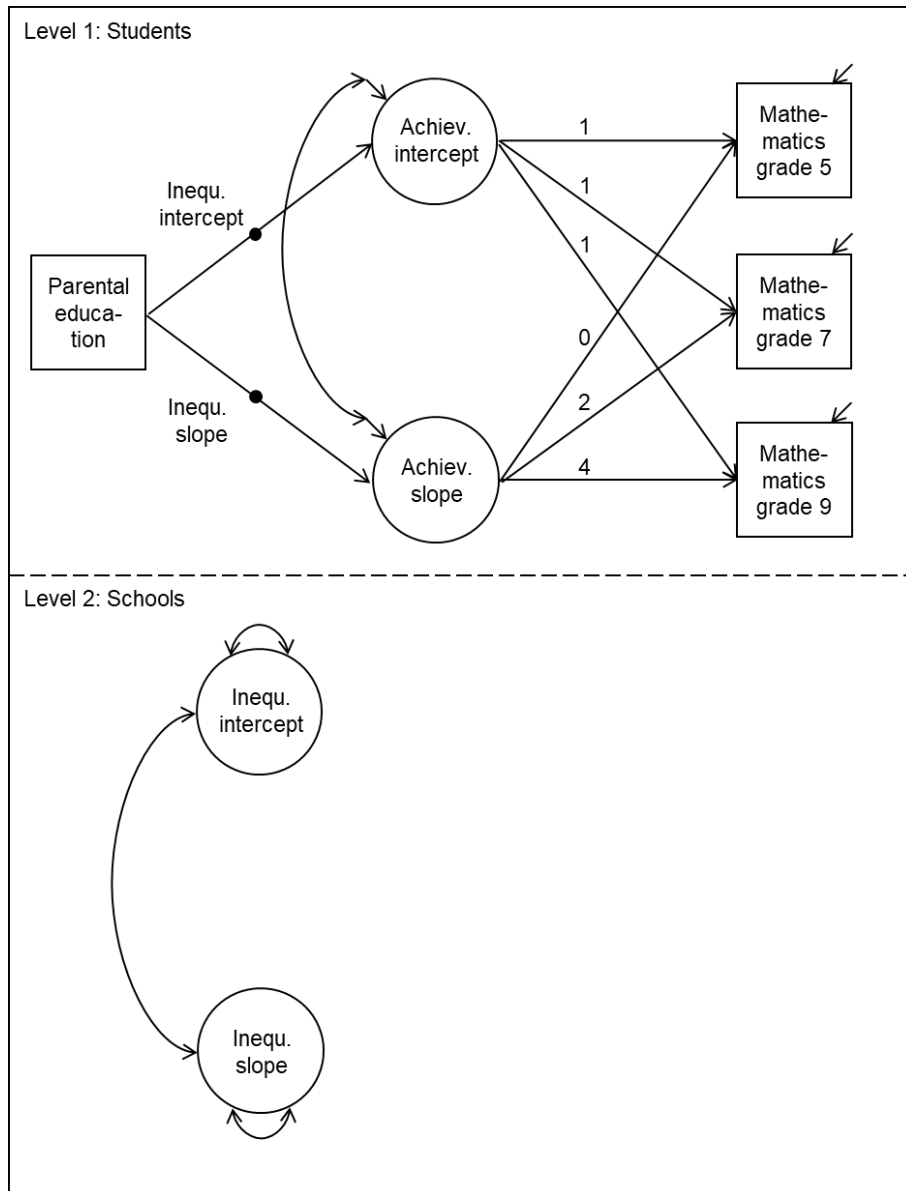
To answer the first research question on the effects of organization form on student achievement, we regressed both school achievement intercept and slope on the dummies for the organization form. The key parameter of interests is the achievement slope, because we were primarily interested in effects on learning progress. The main advantage of using

longitudinal modeling is that it reduced the risk of confounding variables biasing the effect estimation of interest. Even with longitudinal data, confounding variables may bias the analyses. To further minimize the risk of confounding variables, we also controlled for school type and composition.

4.3.2 Modeling Effects on Inequality Regarding Parental Education

To model social inequalities within schools, we extended the previous student level model by regressing the achievement intercept and slope on parental education (see Figure 2). We call the within-school association between the achievement intercept and parental education the inequality intercept, and the within-school association between the achievement slope and parental education the inequality slope. The inequality intercept and slope were modeled as random parameters.

Figure 2 Measurement model for latent growth in social inequality in mathematics achievement



Note. The two-level latent growth model shown in this figure was estimated; the dots in the individual-level graph represent random parameters that vary across schools; the inequality intercept represents the achievement gap between students from educated and uneducated parents in grade 5 and it varies across schools; the inequality slope is the gap in achievement growth between students from educated and uneducated parents and it varies across schools. The school-level achievement intercept and slope serve as dependent variables in further analyses.

To answer the second research question on the impact of half- versus all-day schools on educational inequality, we regressed the school level inequality intercept and inequality slope on the school organization dummy variables. The first parameter reflects differences between half- and all-day schools in the initial degree of social inequality in achievement in grade 5.

The second parameter is the estimate for differences between school organization forms in terms of changes in social inequality in achievement. We were mainly interested in the second parameter. Again, we also included school controls (see previous section).

4.3.3 Modeling Effects on Inequality Regarding Language of Origin

In order to answer the third research question, we replaced parental education with language of origin and replicated the analyses described in the previous section. The resulting model investigated initial achievement gaps and gaps in achievement growth between students with and without German as a first language.

5. Results

5.1 Effects on Student Achievement

Table 2 summarizes the results of the latent growth analyses on student achievement for models with and without control variables for mathematics and reading. The column 1 shows that the mathematics achievement level in grade 5 was significantly higher in all-day schools than in half-day schools. We also observed higher performance levels in schools that changed organization form, but the difference was not significant and should not be over-interpreted because of the difficulty in interpreting the status of these mixed schools. However, the observed differences at the beginning of secondary school may be due to selection effects. The more important estimate of the effect of the organization form on achievement is the parameter of the regression of the achievement slope on the organization form (column 2). The key result is that there is no evidence of differing achievement growth rates in half- and all-day schools.

Table 2 Predictors of growth in achievement at half- and all-day schools

	Mathematics				Reading			
	Model 1a		Model 1b		Model 1c		Model 1d	
	Achiev. intercept (1)	Achiev. slope (2)	Achiev. intercept (3)	Achiev. slope (4)	Achiev. intercept (5)	Achiev. slope (6)	Achiev. intercept (7)	Achiev. slope (8)
All-day school ¹	52.32* (19.58)	-0.16 (1.85)	12.81 (9.11)	-0.28 (1.88)	46.30* (20.15)	-1.76 (1.65)	6.86 (11.24)	-0.34 (1.80)
Changed status ¹	47.83 (27.86)	0.82 (1.91)	13.47 (11.34)	1.16 (2.05)	48.99 (28.80)	-2.00 (1.85)	17.22 (12.08)	-0.57 (1.97)
School with several courses of education ²	-	-	44.06* (16.26)	-2.85 (3.16)	-	-	28.55 (22.69)	1.96 (3.72)
Realschule ²	-	-	68.89* (13.19)	-5.59* (2.54)	-	-	58.16* (19.04)	-2.46 (2.95)
Gesamtschule ²	-	-	37.17 (23.93)	-0.78 (3.86)	-	-	30.13 (26.66)	3.28 (3.97)
Gymnasium ²	-	-	135.64* (25.09)	-3.15 (3.10)	-	-	109.10* (29.45)	-3.57 (4.06)
Social composition	-	-	19.83 (10.51)	0.66 (1.19)	-	-	26.73* (9.42)	-0.77 (1.40)
Language composition	-	-	7.89* (3.85)	-1.43* (0.70)	-	-	3.64 (3.59)	-0.79 (0.66)

Note. The two-level analyses were based on imputed data of $N=3024$ students at 144 schools; “w_t_cal” was used as a sampling weight; maximum likelihood estimation with robust standard errors; unstandardized parameters with standard errors in parentheses; * statistically significant at $p<0.05$. ¹ The reference category was “half-day school”. ² The reference category was “Hauptschule/Volksschule”. Standardized social and language composition variables were used in order to improve the interpretability of the results.

The observed large difference in the achievement intercept vanished after controlling for key school covariates (column 3). This finding confirms the presumption that differences in initial achievement levels were mainly due to differences in school intake. Importantly, however, the effects of organization form on the achievement slope did not change after controlling for key covariates (column 4).

The findings for reading replicated those for mathematics. Initially observed differences in performances levels in grade 5 were no longer significant when further school characteristics were controlled for. Furthermore, the achievement slopes were unrelated to the organization form.

5.2 Effects on Inequality Regarding Parental Education

Table 3 summarizes the results of the analyses on social inequality in student achievement for models with and without control variables for mathematics and reading. The column 1 shows that the degree of social inequality in grade 5 was significantly higher at all-day schools than

at half-day schools. However, this difference at the beginning of secondary school may be due to selection effects. The more important estimate for effects on inequality is the parameter of the regression of the inequality slope on the organization form (column 2). The key result is that there is no evidence of differing inequality growth rates in half- and all-day schools.

Table 3 Predictors of growth in achievement inequality related to parental education at half- and all-day schools

	Mathematics				Reading			
	Model 2a		Model 2b		Model 2c		Model 2d	
	Inequ. intercept (1)	Inequ. slope (2)	Inequ. intercept (3)	Inequ. slope (4)	Inequ. intercept (5)	Inequ. slope (6)	Inequ. intercept (7)	Inequ. slope (8)
All-day school ¹	19.43* (7.61)	-1.25 (0.90)	4.39 (3.00)	-0.08 (0.82)	17.25* (7.91)	-1.58 (0.82)	2.36 (3.90)	-0.17 (0.75)
Changed status ¹	18.86 (10.30)	-1.01 (0.86)	5.06 (4.26)	0.14 (0.73)	18.41 (11.04)	-1.55 (1.01)	5.49 (4.71)	-0.41 (0.76)
School with several courses of education ²	-	-	20.92* (5.99)	-1.73 (1.28)	-	-	15.06* (5.91)	0.17 (1.15)
Realschule ²	-	-	29.79* (5.27)	-2.74* (1.20)	-	-	26.58* (4.88)	-1.39 (0.96)
Gesamtschule ²	-	-	18.23* (7.71)	-0.66 (1.66)	-	-	18.23* (7.72)	0.45 (1.29)
Gymnasium ²	-	-	48.97* (6.07)	-1.82 (1.14)	-	-	41.45* (6.54)	-1.65 (1.13)
Social composition	-	-	7.30* (2.01)	-1.49* (0.38)	-	-	9.08* (2.23)	1.48* (0.43)
Language composition	-	-	3.43* (1.54)	-0.15 (0.34)	-	-	3.45* (1.64)	-0.30 (0.29)

Note. The two-level analyses were based on imputed data of $N=3024$ students at 144 schools; “w_t_cal” was used as a sampling weight; maximum likelihood estimation with robust standard errors; unstandardized parameters with standard errors in parentheses; * statistically significant at $p<0.05$. ¹ The reference category was “half-day school”. ² The reference category was “Hauptschule/Volksschule”. Standardized social and language composition variables were used in order to improve the interpretability of the results.

The observed difference in the inequality intercept was no longer significant when controlling for school covariates (column 3), which supports the assumption that initial inequality differences between schools were mainly due to differences in school intake. Importantly, the null effects of organization form on the inequality slope remained stable when controlling for school covariates (column 4).

The reading analyses produced equivalent findings. Differences between inequality levels in grade 5 for half- and all-day schools were no longer significant when school controls

were included in the analyses. The inequality slopes did not significantly differ between half- and all-day schools.

5.3 Effects on Inequality Regarding Language of Origin

Table 4 summarizes the findings on language-related inequalities in mathematics and reading from models with and without controls. The analyses on the effects of school organization form on inequalities related to students' language of origin were identical to the analyses on social inequalities except that parental education was replaced with language of origin. The analyses basically replicated the results for parental education (see previous section). The findings did not support the assumption that all-day schools would differ from half-day schools with respect to the relationship between achievement growth and language of origin (see Table 4). There were no significant differences in language-related inequality levels in grade 5 between half- and all-day schools, when controlling for school type and composition.

Table 4 Predictors of growth in achievement inequality related to language of origin at half- and all-day schools

	Mathematics				Reading			
	Model 3a		Model 3b		Model 3c		Model 3d	
	Inequ. intercept (1)	Inequ. slope (2)	Inequ. intercept (3)	Inequ. slope (4)	Inequ. intercept (5)	Inequ. slope (6)	Inequ. intercept (7)	Inequ. slope (8)
All-day school ¹	51.47* (19.10)	-0.33 (2.02)	11.70 (9.33)	-0.43 (1.99)	42.90* (20.70)	-1.39 (1.84)	4.22 (11.49)	-0.10 (1.92)
Changed status ¹	47.61 (27.41)	0.54 (2.03)	11.86 (11.55)	1.12 (2.07)	44.90 (28.70)	-1.32 (1.90)	11.57 (12.60)	-0.16 (2.00)
School with several courses of education ²	-	-	44.61* (15.32)	-3.21 (3.16)	-	-	30.00 (15.48)	2.08 (3.06)
Realschule ²	-	-	68.08* (12.04)	-5.09 (2.63)	-	-	58.99* (13.10)	-1.96 (2.49)
Gesamtschule ²	-	-	32.87 (19.88)	1.90 (3.62)	-	-	37.03 (21.75)	4.58 (3.44)
Gymnasium ²	-	-	132.39* (17.09)	-2.34 (2.92)	-	-	108.63* (18.89)	-1.97 (3.04)
Social composition	-	-	21.16* (5.75)	0.14 (1.07)	-	-	26.36* (5.88)	-1.34 (0.97)
Language composition	-	-	5.60 (3.90)	-0.66 (0.78)	-	-	4.26 (3.50)	-0.43 (0.74)

Note. The two-level analyses were based on imputed data of $N=3024$ students at 144 schools; “w_t_cal” was used as a sampling weight; maximum likelihood estimation with robust standard errors; unstandardized parameters with standard errors in parentheses; * statistically significant at $p<0.05$. ¹The reference category was

“half-day school”.² The reference category was “Hauptschule/Volksschule”. Standardized social and language composition variables were used in order to improve the interpretability of the results.

5.4 Robustness Analyses

5.4.1 Student Participation Profiles

Previous research showed that student participation in all-day programs varies in many regards. For example, not all all-day school students attend programs—especially not academic programs—and if they do, they may not spend substantial amounts of time in them (e.g., StEG, 2016). It could be assumed that all-day schools with more preferable participation characteristics had achievement-increasing or inequality-decreasing effects. Therefore, using ranking and median-splitting, we divided the 56 all-day schools into two groups based upon eight participation variables (see Appendix A3). In additional analyses, we reran all models for mathematics and reading with the new dummy variables for low- and high-profile all-day schools and mixed schools, using half-day schools as reference. The results showed that high-profile all-day schools did not show higher achievement growth or lower inequality gradients than half-day schools.

5.4.2 Alternative Social Inequality Indicators

In the main analyses, we chose parental education and language of origin (both assessed in parent interviews) as student background indicators. In order to increase the generalizability of our findings, we reran models 2b, 2d, 3b, and 3d for mathematics and reading with alternative indicators. These indicators were books at home and countries of birth, and both were assessed in the student questionnaires in grade 5. The results were qualitatively identical to the main findings.

5.4.3 Federal States

Although previous research documents that the federal states conduct all-day schools differently (e.g., KMK, 2015), we did not control for this in the main analyses because the samples from some states included only few schools. In additional analyses, we extended models 1b, 1d, 2b, 2d, 3b, and 3d for mathematics and reading by adding dummy variables for federal states. The results were qualitatively the same as the main findings.

5.4.4 Schools with Missing Information on Organization Form

We excluded 20 schools whose principals did not provide information on the organization form in the main analyses because we only had limited information to impute the missing data (see section 4.2.5). At the same time, we acknowledge that the exclusion of schools limits the generalizability of our findings. To address this issue, we replicated all analyses for the full sample of $N=3444$ students at 164 schools with imputed missing data on the organization form based on the limited available data. Again, the results were qualitatively the same as the main findings.

6. Discussion

The present study investigated whether the school organization form—i.e., half- and all-day schools—affected the development of student achievement and educational inequalities. The longitudinal comparison of schools that continued to offer half-day schooling versus those that offered all-day schooling between 2010 and 2015 in two-level latent growth curve models showed no evidence supporting the assumption that organization form influenced mathematics or reading achievement growth and changes in educational inequality in the course of secondary schooling. Neither the achievement slopes nor the inequality slopes (related to social or language background) differed significantly between half- and all-day schools.

We found, however, that half- and all-day schools operated under different conditions. Simple models indicated higher initial levels of achievement and educational inequalities at all-day schools than at half-day schools. These differences were no longer significant when controlling for further school characteristics. Therefore, the differences in the achievement and inequality intercepts seemed to result from selection effects. For example, in the tracked German school system, students are allocated to different school types after primary school. This results in both performance-related and social segregation between school types (e.g., Maaz, Trautwein, Lüdtke, & Baumert, 2008). In the investigated sample, the most selective school type "Gymnasium" was overrepresented among all-day schools.

The fact that our study did not find significant differences between half- and all-day schools in terms of achievement and inequality development is in line with previous studies (Linberg et al., 2018; Steinmann & Strietholt, in print; Strietholt et al., 2015). One possible explanation for the absence of an effect is the quality of all-day schools in Germany. The literature suggests that international extended education programs should, for example, be

aligned with learning goals and supervised by teachers to be effective, i.e. they should be rather formal than non-formal learning opportunities (e.g., Durlak et al., 2010; Miller & Truong, 2009). Previous research on the quality of all-day schools suggests that this is not the case in Germany (e.g., StEG, 2013, 2015, 2016). Indeed, the programs offered by all-day schools in the present sample were on average attended for less than three hours per week. The average shares of students who attended the academic all-day programs offering homework support, remedial education, or subject-specific programs were 15% or lower.

6.1 Scope

The study extends the existing research in three important regards. Most importantly, this is the first study that compared the developments of half- and all-day schools over a period of four years. The first of the three previous robust studies used cross-sectional data (Strietholt et al., 2015), the second had a one-year longitudinal section (Steinmann & Strietholt, in print), and the third had a two-year longitudinal section (Linberg et al., 2018). A second important benefit is that the present study compared schools that remained half- or all-day schools over four years and therefore took into account the changes in organizational structure undertaken by many schools in the investigated time span. Indeed, 31% of the schools reported a mixed organization structure over time. In the previous studies, organization form was assessed only once (Linberg et al., 2018; Steinmann & Strietholt, in print; Strietholt et al., 2015). Therefore, the all-day schools might have changed their organization structure immediately before or after this single point of assessment. Such schools might not have made enough progress in the school development process of becoming all-day schools in order to be fully effective (cf. StEG, 2013, 2015). This explanation for the null effects of all-day schools is less likely in the present study. Third, the study minimizes the risk of the regression-towards-the-mean phenomenon because it uses three measurement points for student achievement.

Apart from these advantages, the scope of the present study has certain limitations. Importantly, we only investigated effects on student achievement development and social disparities in this development. Needless to say, these are not the only important outcomes of extended education. For example, all-day schooling has been found to have beneficial effects on students' psychosocial development (StEG, 2016). Additionally, our analytical sample consisted of students who did not change schools between grades 5–9, which limits the generalizability of our results. Another major issue is that we simply compared half- and all-day schools, although previous research has shown that the quality of actual extended

education at all-day schools is rather low. The data we used only allowed a further investigation of mediators to a limited extent. In additional analyses, we found that all-day schools with favorable student participation characteristics did not have significantly different achievement or inequality developments than half-day schools either. It was, however, not possible to investigate the effects of all-day schools with mandatory all-day participation for all students.

6.2 Future Research

We derive some conclusions for future research from the discussed scope of the present and previous studies. Given that U.S. studies showed that high-quality extended education can increase student achievement, especially among disadvantaged students (e.g., Durlak et al., 2010; Lauer et al., 2006), more research is needed on the circumstances under which all-day schools in Germany can be effective. For example, one current project aims to develop a reading support program that can also be effectively delivered by staff without teacher training (DIPF, 2018). Generally, one reason why international findings on effective extended education programs cannot be directly transferred to all-day schools is that all-day schooling is a rather nonspecific intervention with diverse implications, as the schools are highly autonomous in how they organize their all-day program. Causal relationships between all-day schools' setups and resources, all-day program features, participation characteristics, and student outcomes should be examined in greater detail in future studies (cf. Fischer & Klieme, 2013; Steinmann, 2018; Vest et al., 2013).

From an international perspective, German all-day schooling is an interesting example of an up scaling of non-formal extended education. However, in line with previous research, our study suggests that all-day schools rather provide childcare than extended formal learning opportunities to support student achievement. The study therefore contributes to the international state of research by illustrating that the extension of time at school is not sufficient to extend formal learning opportunities. In fact, the lack of standardization and highly qualified staff for example seem to limit the effectiveness of German all-day schools. This conforms to findings on comparable afterschool programs and summer schools in the US (cf. Apsler, 2009; Durlak et al., 2010; Roth et al., 2010). From an overall perspective, these findings highlight that international comparisons need to pay particular attention to qualitative characteristics of extended education programs.

6.3 Conclusion

Our findings do not support the assumption that all-day schools make a difference for student achievement or educational inequality development in Germany. Given the discussed literature, we draw two tentative conclusions: First, all-day schools do not yet seem to be of sufficient quality to increase student achievement or to decrease social inequalities, at least not on a large scale (see StEG, 2015, 2016; Steinmann, 2018). Second, in order to attain the educational goals of higher achievement gains and lower social inequalities, policy makers may be advised to consider investing in the quality of all-day schools (cf. Fischer et al., 2014; Lossen et al., 2016; StEG, 2016).

7. References

- Apsler, R. (2009). After-school programs for adolescents: A review of evaluation research. *Adolescence, 44*, 1–19.
- Bellin, N., & Tamke, F. (2010). Bessere Leistungen durch Teilnahme am offenen Ganztagsbetrieb? *Empirische Pädagogik, 24*(2), 93–112.
- Blossfeld, H.-P., Roßbach, H.-G., & von Maurice, J. (Eds.) (2011). *Education as a lifelong process – The German National Educational Panel Study (NEPS). Zeitschrift für Erziehungswissenschaft: Sonderheft 14*.
- BMBF (2003). *Investitionsprogramm “Zukunft Bildung und Betreuung”: Ganztagschulen. Zeit für mehr*. Bonn: BMBF.
- Cooper, H., Charlton, K., Valentine, J. C., Muhlenbruck, L., & Borman, G. D. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development, 65*(1), 1–127.
- DIPF (2018). *StEG-Lesen: Förderung der Lesekompetenz an Grundschulen durch innovative Arbeitsgemeinschaft im Ganztage*. Retrieved from <https://www.projektsteg.de/content/steg-lesen>
- Duchhardt, D., & Gerdes, A. (2012). *NEPS technical report for mathematics – scaling results of starting cohort 3 in fifth grade (NEPS Working Paper No. 19)*. Bamberg: Otto-Friedrich-Universität, NEPS.
- Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American journal of community psychology, 45*(3–4), 294–309.

- Feldman, A. F., & Matjasko, J. L. (2005). The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. *Review of Educational Research*, 75(2), 159–210.
- Fischer, L., Rohm, T., Gnams, T., & Carstensen, C. H. (2016). *Linking the data of the competence tests (NEPS Survey Paper No. 1)*. Bamberg: LIfBi, NEPS.
- Fischer, N., & Klieme, E. (2013). Quality and effectiveness of German all-day schools. Results of the Study on the Development of All-Day Schools in Germany. In J. Ecarius, E. Klieme, L. Stecher, & J. Woods (Eds.), *Extended education – an international perspective: Proceedings of the international conference on extracurricular and out-of-school time educational research* (pp. 27–52). Opladen [i.a.]: Barbara Budrich Publishers.
- Fischer, N., Sauerwein, M. N., Theis, D., & Wolgast, A. (2016). Vom Lesenlernen in der Ganztagschule: Leisten Ganztagsangebote einen Beitrag zur Leseförderung am Beginn der Sekundarstufe I? *Zeitschrift für Pädagogik*, 62(6), 780–796.
- Fischer, N., Theis, D., & Züchner, I. (2014). Narrowing the gap? The role of all-day schools in reducing educational inequality in Germany. *International Journal for Research on Extended Education*, 2(1), 79–96.
- Frahm, S., Goy, M., Kowalski, K., Sixt, M., Strietholt, R., Blatt, I., Bos, W., & Kanders, M. (2011). Transition and development from lower secondary to upper secondary school. *Zeitschrift für Erziehungswissenschaft*, 14(special issue, supplement 2), 217–232.
- KMK (2008). *Allgemein bildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: Statistik 2002 bis 2006*. Bonn: KMK.
- KMK (2015). *Ganztagschulen in Deutschland: Bericht der Kultusministerkonferenz vom 03.12.2015*. Berlin, Bonn: KMK.
- KMK (2018). *Allgemein bildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: Statistik 2012 bis 2016*. Berlin: KMK.
- Krannich, M., Jost, O., Rohm, T., Koller, I., Pohl, S., Haberkorn, K., . . . Gnams, T. (2017). *NEPS technical report for reading: Scaling results of starting cohort 3 for grade 7 (Update NEPS Survey Paper No. 14)*. Bamberg: LIfBi, NEPS.
- Kuger, S. (2016). Curriculum and learning time in international school achievement studies. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), *Assessing Contexts of Learning* (pp. 395–422). Cham: Springer.

- Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research, 76*(2), 275–313.
- Linberg, T., Struck, O., & Bäumer, T. (2018). Vorzug Ganztagschule? Zusammenhänge mit der Kompetenzentwicklung im Bereich Lesen und Mathematik. *Zeitschrift für Erziehungswissenschaft, 14*(2), 51.
- Lossen, K., Tillmann, K., Holtappels, H. G., Rollett, W., & Hannemann, J. (2016). Entwicklung der naturwissenschaftlichen Kompetenzen und des sachunterrichtsbezogenen Selbstkonzepts bei Schüler/-innen in Ganztagsgrundschulen: Ergebnisse der Längsschnittstudie StEG-P zu Effekten der Schülerteilnahme und der Angebotsqualität. *Zeitschrift für Pädagogik, 62*(6), 760–779.
- Maaz, K., Trautwein, U., Lüdtke, O. & Baumert, J. (2008). Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives, 2*(2), 99–106.
- Miller, B. M., & Truong, K. A. (2009). The role of afterschool and summer in achievement. The untapped power of afterschool and summer to advance student achievement. *Zeitschrift für Pädagogik, 55*(54th supplement), 124–142.
- Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the school day or school year: A systematic review of research (1985–2009). *Review of Educational Research, 80*(3), 401–436.
- Plantenga, J., & Remery, C. (2013). *Childcare services for school age children: A comparative review of 33 countries*. Luxembourg: Publications Office of the European Union.
- Pohl, S., Haberkorn, K., Hardt, K., & Wiegand, E. (2012). *NEPS technical report for reading – scaling results of starting cohort 3 in fifth grade (NEPS Working Paper No. 15)*. Bamberg: Otto-Friedrich-Universität, NEPS.
- Roth, J. L., Malone, L. M., & Brooks-Gunn, J. (2010). Does the amount of participation in afterschool programs relate to developmental outcomes? A review of the literature. *American journal of community psychology, 45*(3–4), 310–324.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.

- Scharl, A., Fischer, L., Gnambs, T., & Rohm, T. (2017). *NEPS technical report for reading: Scaling results of starting cohort 3 for grade 9 (NEPS Survey Paper No. 20)*. Bamberg: LfBi, NEPS.
- Schnittjer, I., & Gerken, A.-L. (2017). *NEPS technical report for mathematics: Scaling results of starting cohort 3 in grade 7 (NEPS Survey Paper No. 16)*. Bamberg: LfBi, NEPS.
- Stecher, L., & Maschke, S. (2013). Research on extended education in Germany – A general model with all-day schooling and private tutoring as two examples. *International Journal for Research on Extended Education*, 1(1), 31–52.
- StEG (2013). *Ganztagsschule 2012/2013: Deskriptive Befunde einer bundesweiten Befragung*. Frankfurt a. M. [i.a.]: DIPF [i.a.].
- StEG (2015). *Ganztagsschule 2014/2015: Deskriptive Befunde einer bundesweiten Befragung*. Frankfurt a. M. [i.a.]: DIPF [i.a.].
- StEG (2016). *Ganztagsschule: Bildungsqualität und Wirkungen außerunterrichtlicher Angebote: Ergebnisse der Studie zur Entwicklung von Ganztagsschulen 2012–2015*. Frankfurt a. M. [i.a.]: DIPF [i.a.].
- Steiner, C. (2009). Mehr Chancengleichheit durch die Ganztagsschule? *Zeitschrift für Pädagogik*, 55(54th supplement), 81–105.
- Steiner, C. (2011). Teilnahme am Ganztagsbetrieb. Zeitliche Entwicklung und mögliche Selektionseffekte. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher, & I. Züchner (Eds.), *Ganztagsschule: Entwicklung, Qualität, Wirkungen: Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagsschulen (StEG)* (pp. 57–75). Weinheim [i.a.]: Beltz Juventa.
- Steinmann, I. (2018). Ganztagsschulen als neue Organisationsformen – Entwicklungen und Evaluation. *Enzyklopädie Erziehungswissenschaft Online*. doi: 10.3262/EEO08180405.
- Steinmann, I., & Strietholt, R. (in print). Effekte nachmittäglicher Bildungsangebote auf sprachliche Kompetenzentwicklungen. *Zeitschrift für Pädagogik*.
- Steinmann, I., Strietholt, R., & Caro, D. (2018). Participation in extracurricular activities and student achievement: Evidence from German all-day schools. *School Effectiveness and School Improvement*. doi 10.1080/09243453.2018.1540435.
- Strietholt, R., Manitius, V., Berkemeyer, N., & Bos, W. (2015). Bildung und Bildungsungleichheit an Halb- und Ganztagsschulen. *Zeitschrift für Erziehungswissenschaft*, 18(4), 737–761.

- Strietholt, R., Naujokat, K., Mai, T., Kretschmer, S., Jarsinski, S., Goy, M., Frahm, S., Kanders, M., Bos, W., & Blatt, I. (2013). The National Educational Panel Study (NEPS) in Germany. An overview of design, research options, and access, with a focus on lower secondary school. *European Educational Research Journal*, 12(4), 568–579.
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–67.
- Vest, A. E., Mahoney, J. L., & Simpkins, S. D. (2013). Patterns of out-of-school time use around the world: Do they help to explain international differences in mathematics and science achievement. *International Journal for Research on Extended Education*, 1(1), 71–85.
- Werquin, P. (2010). *Recognising non-formal and informal learning: Outcomes, policies and practices*. Paris: OECD.
- Züchner, I., Fischer, N. (2014). Kompensatorische Wirkungen von Ganztagschulen – Ist die Ganztagschule ein Instrument zur Entkopplung des Zusammenhangs von sozialer Herkunft und Bildungserfolg? *Zeitschrift für Erziehungswissenschaft*, 17, 349–367.

Appendices

Appendix A1 Descriptive statistics of student and school characteristics prior to imputation

	All schools	Divided by school organization form			
		Half-day school	All-day school	Changed status	Missing ¹
<i>Student characteristics</i>	<i>N=3024</i>	<i>N=297</i>	<i>N=744</i>	<i>N=433</i>	<i>N=1550</i>
Student achievement					
Mathematics grade 5 (<i>M(SD)</i>)	-1.35 (114.89)	2.86 (119.56)	23.34 (107.78)	11.74 (100.68)	-14.51 (118.33)
Mathematics grade 7 (<i>M(SD)</i>)	75.32 (122.51)	75.23 (132.76)	106.10 (112.20)	89.47 (119.53)	59.90 (122.44)
Mathematics grade 9 (<i>M(SD)</i>)	152.13 (117.12)	148.21 (118.74)	176.72 (113.08)	176.41 (101.60)	136.05 (120.15)
Reading grade 5 (<i>M(SD)</i>)	1.82 (122.81)	1.71 (129.18)	20.61 (114.26)	13.19 (110.75)	-8.02 (126.73)
Reading grade 7 (<i>M(SD)</i>)	70.96 (131.81)	62.93 (131.71)	82.02 (134.78)	85.41 (118.14)	63.89 (134.18)
Reading grade 9 (<i>M(SD)</i>)	127.09 (107.76)	124.14 (106.72)	138.55 (106.52)	133.02 (95.82)	121.65 (111.62)
Student background					
Parental education grade 5 (<i>M(SD)</i>)	2.81 (1.10)	2.67 (1.02)	3.02 (1.06)	2.87 (1.11)	2.73 (1.11)
Language of origin German grade 5	92.1%	90.2%	93.9%	95.4%	90.5%
<i>School characteristics</i>	<i>N=144</i>	<i>N=18</i>	<i>N=29</i>	<i>N=19</i>	<i>N=78</i>
School type					
Hauptschule/Volksschule	27.8%	22.3%	12.7%	3.7%	37.5%
School with several courses of education	13.8%	4.9%	6.6%	9.0%	17.8%
Realschule	17.3%	47.5%	14.8%	39.2%	8.0%
Gesamtschule	8.5%	0.0%	17.0%	15.4%	6.4%
Gymnasium	32.6%	25.3%	49.0%	32.6%	30.2%
School composition					
Social composition grade 5 (<i>M(SD)</i>)	2.59 (0.61)	2.66 (0.45)	2.92 (0.51)	2.89 (0.34)	2.43 (0.65)
Language composition grade 5 (<i>M(SD)</i>)	0.79 (0.34)	0.90 (0.18)	0.94 (0.07)	0.94 (0.12)	0.71 (0.40)

Note. The descriptive analyses were based on non-imputed data of $N=3024$ students at 144 schools; “w_t_cal” was used as a sampling weight. ¹ Students at schools where school organization form information is missing for at least one out of three measurement points. Unstandardized social and language composition variables were used.

Appendix A2 Variables in the imputation model on the student and school level: Sources, measurement points, and percentages of missing values

	Source	2010 /11	2011 /12	2012 /13	2013 /14	2014 /15
Replication weight	cohort profile	0.0%				
School level variables						
School type	cohort profile	0.0%				
Federal state	cohort profile	0.0%				
Social composition	aggregated student data	0.6%				
Language composition	aggregated student data	0.6%				
Cross-sectional organization form variable	principal quest.	35.9%		24.4%		39.3%
School offer of						
Homework supervision	principal quest.		15.5%		36.3%	
Enrichment groups: students with high grades	principal quest.		18.5%		36.6%	
Remedial instruction: students with low grades	principal quest.		16.6%		35.9%	
Remedial instruction: non-native speakers	principal quest.		17.4%		37.7%	
Language of origin instruction	principal quest.		17.4%		37.9%	
Subject-specific offering in mathematics	principal quest.		19.5%		39.5%	
Subject-specific offers in science	principal quest.		17.8%		38.1%	
Subject-specific offers in German	principal quest.		19.8%		41.1%	
Subject-specific offers in foreign languages	principal quest.		18.2%		40.2%	
Sports offers	principal quest.		17.6%		38.8%	
Music/art offers	principal quest.		16.0%		37.7%	
Religion offers	principal quest.		24.0%		43.0%	
Trades and home economics offers	principal quest.		16.7%		38.6%	
Technology/new media offers	principal quest.		16.4%		36.0%	
Community activities/student government	principal quest.		15.5%		37.8%	
Forms of social learning	principal quest.		15.9%		38.1%	
Forms of intercultural learning	principal quest.		15.3%		43.4%	
Required free-time activities	principal quest.		16.0%		38.7%	
Voluntary free-time activities	principal quest.		15.7%		38.9%	
All-day provision for grade 8	principal quest.				35.5%	
Number of staff in all-day program						
Without definite university degree	principal quest.				56.2%	
With university degree	principal quest.				61.0%	
No. days all-day progr. in 8 th and/or 12 th grade	principal quest.				38.9%	
Student level variables						
Student achievement						
Mathematics	student test	8.9%		4.8%		7.1%
Reading	student test	9.0%		4.9%		15.1%
Orthography	student test	8.7%		4.8%		7.4%
Perceptual speed	student test	8.9%				11.8%
Cognitive reasoning	student test	9.1%				11.5%
Declarative metacognition	student test		4.4%			11.2%
ICT literacy	student test		5.3%			7.3%
Scientific competence	student test		5.3%			7.3%
Reading speed	student test	8.7%				7.1%
Time between tests	student test			13.5%		11.2%
Student participation in						
Homework support	student quest.		10.8%		13.4%	
Remedial education	student quest.		11.4%		13.7%	
Subject-specific programs	student quest.		12.2%		14.7%	
Subject-unrelated projects	student quest.		10.9%		14.4%	

Leisure facilities	student quest.	12.1%	14.9%	
Participation frequency	student quest.	16.8%	24.4%	
Student rating of all-day program				
Enjoyment	student quest.	35.3%	55.4%	
Useful in class	student quest.	36.4%	56.1%	
Wish for more offers	student quest.	36.4%	56.2%	
Learning new things	student quest.	37.1%	56.5%	
Preferred free time	student quest.	37.3%	56.7%	
Made new friends	student quest.	37.3%	56.7%	
Improving grades	student quest.	37.4%	56.7%	
Not alone in the afternoons	student quest.	37.5%	56.8%	
Student background				
Books at home	student quest.	14.2%	7.9%	8.0%
Parental education	parent interv.	37.9%	43.3%	55.0%
Partner parental education	parent interv.	50.3%	49.6%	60.2%
Family native-born	student quest.	8.2%	8.2%	
Language of origin	parent interv.	37.1%	37.1%	
Gender	cohort profile	4.8%	1.3%	
Age in years	cohort profile	12.1%	7.3%	

Appendix A3 Descriptive statistics for the subsamples of all-day schools with high and low student participation profiles for robustness analyses

	All-day schools	Divided by student participation characteristics	
		Low profile all-day school	High profile all-day school
<i>School characteristics</i>	<i>N=56</i>	<i>N=28</i>	<i>N=28</i>
Average student participation frequency ¹			
Average frequency grade 6 (<i>M(SD)</i>)	2.47 (1.48)	1.79 (0.78)	2.97 (1.66)
Average frequency grade 8 (<i>M(SD)</i>)	1.61 (1.38)	1.19 (0.72)	1.92 (1.64)
Percentage of students participating in programs ²			
Homework support grade 6 (<i>M(SD)</i>)	0.15 (0.15)	0.07 (0.07)	0.21 (0.16)
Homework support grade 8 (<i>M(SD)</i>)	0.09 (0.08)	0.04 (0.04)	0.12 (0.09)
Remedial education grade 6 (<i>M(SD)</i>)	0.13 (0.11)	0.08 (0.08)	0.17 (0.12)
Remedial education grade 8 (<i>M(SD)</i>)	0.10 (0.08)	0.05 (0.05)	0.14 (0.09)
Subject-specific learning offers grade 6 (<i>M(SD)</i>)	0.06 (0.07)	0.04 (0.04)	0.08 (0.08)
Subject-specific learning offers grade 8 (<i>M(SD)</i>)	0.09 (0.08)	0.05 (0.05)	0.11 (0.09)

Note. The descriptive analyses were based on imputed data of $N=1326$ students at 56 all-day schools; “w_t_cal” was used as a sampling weight. All variables were taken from student questionnaires. ¹ Aggregated variable on hours per week students made use of all-day offers overall; open answering format. ² Aggregated variables on the participation in specific all-day programs; binary answering format with 0 (“no participation”) and 1 (“participation”).

2.4 Beitrag IV: Participation in extracurricular activities and student achievement: Evidence from German all-day schools

Isa Steinmann, Rolf Strietholt & Daniel Caro

Die im Folgenden abgedruckte Version des Manuskripts ist das Preprint des in School Effectiveness and School Improvement veröffentlichten Beitrags.

Abstract

The present study investigates the effects of participation in academic extracurricular activities on student achievement in mathematics and reading in the course of secondary schooling. We use two samples of nonmandatory all-day schools that are part of the National Educational Panel Study (NEPS): N=1131 fifth graders at 43 schools were followed until grade 7 and N=1545 seventh graders at 64 schools until grade 9. We compare students participating in all-day schools' extracurricular activities homework support, remedial education, and subject-specific programs to nonparticipating students. After controlling for prior achievement and further student background variables as well as estimating school fixed effects, we find no effects of participation in extracurricular activities on student achievement for both samples and both outcomes. The study contributes to the evidence base on the effects of extracurricular activities. We discuss the findings in light of the international discourse on effective extracurricular activities.

Participation in extracurricular activities and student achievement: Evidence from German all-day schools

Introduction

Much learning takes place outside of regular school lessons. At the age of 15, for example, students in OECD countries spend an average of 18 hours per week learning the language of instruction, mathematics, and science and spend 62% of this time in regular school lessons, 26% studying individually, and 12% in so-called out-of-school-time lessons (OECD, 2011). The out-of-school-time lessons encompass activities taking place at school (e.g., all-day schooling), home (e.g., private tutoring), or elsewhere (e.g., summer schools). School-based extracurricular activities are of particular interest for policy making and educational research as they are more amenable to regulation by reforms than private study or tutoring (cf. Kuger, 2016).

Two strands of arguments for extending extracurricular activities in schools dominate the public and scientific discourses. From a social policy perspective, the additional time children spend at school should improve the employability of both parents. From an educational perspective, the additional time at school may be used to provide extended learning opportunities, individual support, or to enrich and diversify the school day in order to encourage the academic and psychosocial development of children (cf. Fischer, Theis, & Züchner, 2014; Plantenga & Remery, 2013; Stecher & Maschke, 2013). We focus on the educational perspective in the present study.

While the desire to support children's development is a universal one, the definition, organization, and use of additional time varies across and within countries. For instance, educational systems differ in terms of their programs (e.g., remedial education, sports clubs), the amount of time allocated to them, staff qualifications, staff-to-child ratios, and enrolment rates (OECD, 2015; Plantenga & Remery, 2013). This international divergence in extended education characteristics needs to be discussed in the context of theoretical effectiveness models. As different models emphasize, extracurricular programs need to meet quality criteria to affect student outcomes like school achievement (cf. Fischer & Klieme, 2013; Miller & Truong, 2009; Vandell, Reisner, & Pierce, 2007). For example, the programs should be designed to foster skills, linked to regular school lessons, and supervised by qualified staff. Furthermore, even if a school provides high-quality programs, development effects can only be expected for those students who actually participate over a certain timespan. If attendance

is voluntary, programs need to be attractive to students so that they decide to attend. We can therefore assume that variations between countries in the use of additional time impact the effectiveness of extended education. One can thus expect to find internationally differing effects on outcomes like student achievement: While attending extended learning programs might promote learning in some educational systems, this might not be the case in others.

Extended Education and Student Achievement: International Variation in Programs and their Effectiveness

The term extended education refers to programs that differ in their scope and aims. For this reason it is important to acknowledge such differences when reviewing the previous research on different extended education programs. We limited our literature review to research on student test scores and did not include studies that considered, for example, school grades or motivational outcomes. In the US, so-called after-school programs are rather common. Such programs are mostly organized independently of the schools, operate after regular school finishes, and provide adult supervision for participating students. They differ in terms of their opening hours, facilities (e.g., dedicated facilities vs. community centers), the types of activities (e.g., childcare vs. academic enrichment), and the qualifications of the supervising staff (e.g., trained vs. layperson) (e.g., Scott-Little, Hamann, & Jurs, 2002). Meta-analyses and literature reviews show positive mean effects of after-school programs on student achievement in the US. The primary studies discussed in these reviews are both small-scale interventions and evaluations of interstate initiatives of various program types. Durlak, Weissberg, and Pachan (2010), for example, identified a positive overall effect of after-school programs on student outcomes. Specifically, programs with certain quality criteria had a positive impact on student learning, while there were neutral effects in the absence of such criteria: After-school programs needed to include sequenced learning steps and active student work on dedicated material, be focused in terms of allocated time and attention, and have explicit learning objectives in order to affect student outcomes. The relevance of after-school programs' quality also becomes apparent when reviewing further meta-analyses and literature reviews. Generally, higher effect sizes were associated with the specificity of the programs' goals, like the aim to promote learning (Apsler, 2009; Cooper, Charlton, Valentine, Muhlenbruck, & Borman, 2000; Lauer et al., 2006), a close link between the programs' contents and regular school curricula (Fashola, 1998; Scott-Little et al., 2002), staff qualifications, and a low teacher-student ratio (Cooper et al., 2000; Fashola, 1998; Feldman &

Matjasko, 2005; Scott-Little et al., 2002). Primary school programs were somewhat more effective than programs for older students (Cooper et al., 2000; Scott-Little et al., 2002). However, the positive findings could not be replicated in low-quality programs with opposite characteristics (Apsler, 2009; Durlak et al., 2010; Roth, Malone, & Brooks-Gunn, 2010). Furthermore, researchers identified favorable characteristics of student participation. Students who attended more frequently and consistently (Apsler, 2009; Feldman & Matjasko, 2005), as well as at-risk students, benefited more than others (Patall, Cooper, & Allen, 2010; Scott-Little et al., 2002). Studies on further extracurricular activities like summer schools (which take place during the summer holidays) basically replicate the findings on after-school programs (Lauer et al., 2006).

Compared to the rich body of research on extended education in the US, the number of studies on this topic published in other countries is much smaller. In the following, we provide evaluation examples ranging from small-scale local interventions to large-scale education extensions in entire school systems. We will stress the characteristics of the different investigated programs to illustrate cross-national differences. In Ireland, a small-scale randomized trial found that attending an intervention that aimed to promote reading had literacy-increasing effects over the first school year. This 36-week-long after-school training program was tailored to students who have trouble learning to read in disadvantaged schools and comprised three 90-minute sessions per week (Biggart, Kerr, O'Hare, & Connolly, 2013). By contrast, another small-scale randomized field experiment with elementary school students in the Netherlands found no significant effects of attending a different extended day program on mathematics and language achievements. On average, this pilot program comprised two additional hours of Dutch instruction, two hours of mathematics, and a one-hour excursion per week. The authors suggest that the intervention period of 11 weeks might have been too short to impact academic achievement (Meyer & van Klaveren, 2013). In contrast to the after-school programs in the US, which are long term and widely available, the two programs in Ireland and the Netherlands were small pilot interventions that were implemented outside of regular classes but in the school context.

In Italy, a difference-in-difference study found positive effects of a larger-scale funding project on mathematics (but not on language) achievement over one year in lower secondary schools. The funding enabled more than 200 low-achieving schools to provide additional instruction in language and/or mathematics outside of regular school hours for some classes. All students were required to attend these additional subject-specific activities,

which were conducted by specifically trained teachers. Further analyses suggest that the effects were more pronounced for the least advantaged students (Battistin & Meroni, 2016). This project was—similar to the Irish and Dutch studies—implemented in the school context but comprised additional regular classes instead of specific interventions.

In Chile, small mathematics- and language-promoting effects in tenth grade were found over two years in a difference-in-difference study on a nation-wide extended education reform. The school days were switched from two shifts—some students had classes in the morning, others in the afternoon—to a full school day for all public schools. In contrast to the previously mentioned European studies, this intervention entailed reforming the whole school system and had very heterogeneous implications for how the additional time was used. Unlike after-schooling in the USA, it was closely connected to the schools. Some of the Chilean schools used the additional time for extended regular classes or extracurricular activities that were more or less aligned with learning. The difference-in-difference study compared students who studied under all-day conditions with students who studied under half-day conditions, i.e. whose schools were still on the (government-determined) waiting list to implement the reform. Further analyses showed that high-achieving students at rural schools especially benefitted from the extended day (Bellei, 2009).

In the German-speaking part of Switzerland, the primary school system comprises both all-day schools and half-day schools. While half-day schools finish before lunch, all-day schools provide lunch and additional supervised time in the afternoon, which is typically used for homework supervision and recreational activities. Participation in these activities is usually voluntary. Students who intensively attended the extracurricular activities at their all-day schools were found to have small mathematics and reading achievement advantages compared to their nonparticipating peers over the first two school years. However, these findings need to be interpreted with caution because the treatment groups of the studies were very small (less than 40 students) and included students from more privileged backgrounds (Schüpbach, 2014, 2015).

Similarly, there are all-day schools in addition to traditional half-day schools in Germany. As a result of a massive state funding program that promoted the extension of all-day schools in the beginning of this century, the nation-wide share of all-day schools increased from 16% in 2002 to 65% in 2015 (KMK, 2008, 2017). All-day schools are schools that provide lunch and an all-day program of at least seven hours on at least three days per week. At about half of the all-day schools, participation is nonmandatory: at the other half, it

is obligatory either for certain groups of students or for all students. These are called nonmandatory, partially, and fully mandatory all-day schools respectively (KMK, 2008). Because of the autonomy that German (and Swiss) schools have in designing and organizing their all-day activities, the term all-day schooling refers to programs that are characterized by a low level of standardization in comparison to the previously mentioned programs in other countries. For example, the all-day programs are not necessarily intended to foster learning but also incorporate leisure activities (StEG-Consortium, 2015, 2016).

To summarize, the German state of research is less promising than the international one. In the following, we review well-designed studies aiming to establish causal inference of all-day schooling effects on student achievement. So far, no experiment has been conducted in this field. A study comparing half-day and all-day schools found no significant differences in science, reading, and mathematics achievement. The authors used cross-sectional representative large-scale data from primary and secondary school samples from Germany and applied a matching procedure to balance systematic differences between half- and all-day schools. Retrospective information on early literacy skills and teacher ratings were used as proxies for prior achievement to mimic a longitudinal design. The analyses further showed that the socioeconomic and immigrant-background-related gaps in student achievement did not differ between half-day and all-day schools (Strietholt, Manitius, Berkemeyer, & Bos, 2015).

Only three longitudinal studies with test measures of prior achievement evaluated the student achievement effects of attending extracurricular activities at all-day schools instead of the aggregated effects on the school level, like Strietholt and colleagues' study. Bellin and Tamke (2010) found that first-grade students who in general attended the all-day program at their nonmandatory all-day schools attained significantly higher reading scores in second grade after controlling for prior achievement in first grade. However, the effect between the two measurement points was small and no longer significantly different from zero when also controlling for cognitive ability in the first grade. Although these findings are informative, it is important to acknowledge that the authors did not distinguish between domain-specific learning and other activities in the afternoon. In addition, the study was confined to one federal state.

In a second and larger longitudinal study with two measurement points, researchers found that neither participating in reading nor in media programs (e.g., school newspaper or drama club) affected students' reading development between the beginning and middle of the

fifth grade (Fischer, Sauerwein, Theis, & Wolgast, 2016). In further analyses, participating students were differentiated according to the voluntary or involuntary nature of their participation: Those who participated voluntarily in reading programs showed achievement advantages over involuntary participants. For media programs this effect was not significantly different from zero.

In a third longitudinal study with two measurement points, participation in science-related extracurricular activities in all-day schools at the end of the third grade was found to have no significant effect on science achievement at the end of the fourth grade in 10 federal states. Researchers did find positive effects on science achievement when considering participation in programs in the middle or end of the fourth grade. Although these positive effects were significant, the effect sizes were extremely small. The authors therefore concluded that the study could not substantiate a general participation effect. They also tested whether broad participation (i.e., the overall number of attended science programs), continuous participation (i.e., the number of school terms), or the perceived quality of the extracurricular activities impacted achievement but found no evidence to support these hypotheses (Lossen, Tillmann, Holtappels, Rollett, & Hannemann, 2016).

Before summarizing the research on achievement effects, it seems worth remembering that a considerable degree of freedom constitutes German all-day schooling. Overall, however, previous studies show no achievement effect of all-day versus half-day schools or of participation in extracurricular activities at all-day schools in Germany. This also holds for studies that focus explicitly on the effects of domain-specific all-day learning activities on achievement in the respective domain (Fischer et al., 2016; Lossen et al., 2016). Additionally, the longitudinal studies strengthen the assumption that attending after-school activities is subject to selection mechanisms that can bias effect estimations. In the first study, participating students had cognitive and reading advantages over nonparticipating students, and for most of them German was their first language. Controlling for prior achievement in the analyses did not remove all confounding through unobserved covariates (Bellin & Tamke, 2010). Reading program participants were found to be low-performing readers on average, and media programs were attended by students with higher reading learning goals and higher reading self-concepts on average (Fischer et al., 2016). Attendees of science-related programs tended to be male, have an immigrant background, and a low socio-economic status (Lossen et al., 2016). Additionally, schools' academic programs differed between school types and

were attended by less than 20% of students on average (Fischer et al., 2016; Lossen et al., 2016).

However, the state of research on the effects of all-day schooling on student achievement is scarce and not comprehensive. The studies only look at students from the first to the fifth grades, the comparison intervals are only between half and one year, and they only cover reading and science-related programs and outcome measures.

The Current Study

The overview of the effectiveness research reveals considerable differences in how extended education programs have been defined, organized, and studied internationally. These differences reflect a natural limitation of research on extended education across countries. In light of this limitation, we aimed to test whether participation in academic extracurricular activities at German secondary all-day schools affected student achievement in reading and mathematics. Specifically, we focused on a representative sample of nonmandatory all-day schools offering homework support, remedial education, and subject-specific programs where some students did and did not participate in these activities. We focused on this subset of schools in order to exploit the within-schools variation in the explanatory variables and thus circumvent bias from school-level confounding factors.

We separately investigated the following three research questions. (1) Did students who participated in homework support have achievement advantages over nonparticipating classmates? (2) Did students who participated in remedial education have achievement advantages over nonparticipants? (3) Did students who attended subject-specific programs have achievement advantages over those who did not attend such programs? We studied developments in the core learning areas of reading and mathematics achievement. Following the discussed literature and political expectations, we hypothesized we would find mathematics and reading achievement advantages among students who participated in (1) homework support, (2) remedial education, and (3) subject-specific programs in comparison to nonparticipants.

Method and Analyses

Data

This paper uses representative data for Germany on secondary schools from the National Educational Panel Study (NEPS; Blossfeld, Roßbach, & Maurice, 2011).¹² As we aimed to investigate the effects of participating in extracurricular activities, we strove to compare students participating in these activities to nonparticipating classmates (see analysis plan section below) and thus focused on nonmandatory all-day schools. We excluded data from special-needs schools because differences between the assessment materials precluded comparisons with the regular school sample. Due to these restrictions, our analyses are limited to regular secondary schools with nonmandatory all-day schooling, which applies to about one out of three secondary schools (KMK, 2017). Additionally, we aimed to investigate achievement developments and therefore only included students who stayed at the same schools between measurement points.

We considered two samples: In the first sample of $N=1131$ students at $N=43$ nonmandatory all-day schools, the students were tested in grades 5 (2010/11) and 7 (2012/13). The average age in grade 5 was $M=10.88$ years ($SD=0.53$) and 48.3% of the students were girls. The average number of students per school was $M=26.30$ ($SD=12.07$). The schools were from 10 federal states and 23 were Gymnasien (the school type with the strictest ability selection in the tracked German school system). The second sample comprised data from $N=1545$ students at $N=64$ schools who were tested in grades 7 (2012/13) and 9 (2014/15). The average age was $M=12.91$ years in grade 7 ($SD=0.53$) and 51.9% of the students were girls. The average number of students per school was $M=24.14$ ($SD=9.70$). The schools in this second sample were located in 13 federal states and 32 were Gymnasien.

Although there was an overlap across the two samples, we analyzed the data separately for three reasons. First, due to panel attrition, an additional student sample was added to the initial NEPS sample in 2012/13 for which no prior information was available.

¹² Starting Cohort Grade 5, doi:10.5157/NEPS:SC3:7.0.1. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LifBi) at the University of Bamberg in cooperation with a nationwide network.

Second, only a few students attended specific extracurricular activities consistently between grades 5 and 9. Third, due to the ongoing reforms many schools changed from half-day to all-day or from nonmandatory to mandatory schooling and vice versa during the investigated timespan. In order to clearly define our treatment, such schools were excluded from the analyses. Treating the two samples separately, however, circumvented some issues related to the treatment status and minimized the number of excluded cases.

Outcomes

The outcome variables were mathematics and reading achievement scores. The paper-pencil-tests in mathematics and reading were administered in grades 5, 7, and 9. The scores were comparable and linked across assessment cycles to make it possible to study change in the respective domains (NEPS, 2011, 2016). We used weighted maximum likelihood estimates (WLE) of student abilities based on item response theory models (Pohl & Carstensen, 2012). In sample 5–7, the mean WLE posttest scores in grade 7 were $M=0.99$ in mathematics ($SD=1.18$, 5.7% missing) and $M=0.87$ in reading ($SD=1.31$, 5.8% missing). In sample 7–9, the mean achievement scores in grade 9 were $M=1.71$ in mathematics ($SD=1.24$, 8.2% missing) and $M=1.29$ in reading ($SD=1.22$, 16.0% missing).

The mathematics test assessed student achievement in abstract thinking and mathematical problem solving in the four content areas quantity (e.g., percentage calculation), space and shape (e.g., geometric figures), change and relationship (e.g., algebraic curves), as well as data and chance (e.g., probability calculation). The EAP/PV reliability of the test was 0.802 and was determined on the basis of all fifth graders taking the NEPS test (Duchhardt & Gerdes, 2012). The reading test assessed how students derived information from continuous texts, drew conclusions, reflected, and assessed contents. The reading material covered informational, commenting, literary, instructional, and advertising texts. The EAP/PV reliability of the test was 0.811 and was also determined based on the performance of all tested fifth graders (Pohl, Haberkorn, Hardt, & Wiegand, 2012).

Treatments

The explanatory variables were different forms of participation in three extracurricular activities: (a) “homework support/supervision,” (b) “remedial groups/remedial education,” and (c) “subject-specific learning offers (e.g., extra or enhancement courses in math or German)”. The binary variables were 0 if a student did not attend the respective activity and 1

if he or she did attend. The variables were administered in grades 6 and 8. In sample 5–7, 17.2% of students took part in homework support (9.4% missing), 15.5% in remedial education (9.8% missing), and 7.4% in subject-specific programs (10.8% missing). The participation rates were lower in sample 7–9: 8.3% in homework support (10.8% missing), 8.0% in remedial education (11.0% missing), and 6.3% in subject-specific programs (11.4% missing). In sample 5–7, only 5.0% attended two and 1.2% all three activities. In sample 7–9, 4.4% attended two and 1.6% three activities. In nonmandatory all-day schools, participation in extracurricular activities must be declared in advance and is binding for at least half a school year (KMK, 2008), which therefore reflects the minimum treatment dosage.

Covariates

Participation in extracurricular activities at nonmandatory all-day schools is not randomized. The literature shows that the characteristics of extracurricular programs vary between federal states and individual schools, for instance, in terms of participation rates (Fischer & Klieme, 2013; Prein, Rauschenbach, & Züchner, 2009; StEG-Consortium, 2015; Steiner, 2011). Previous research also documents systematic differences at the individual level. Students attending extracurricular activities differ from their nonparticipating classmates in terms of cognitive and family background measures. The degree and direction of these differences depend on the program type (Bellin & Tamke, 2010; Fischer et al., 2016; Fischer & Klieme, 2013; Lossen et al., 2016; Prein et al., 2009; Schüpbach, 2014, 2015; Steiner, 2011). Based on the literature, we determined a series of covariates reflecting selection mechanisms at student level. These include prior cognitive and noncognitive measures (tests, grades, school satisfaction) and family background (social and employment status). We also added measures of alternative parental learning support that might interrelate with enrollment in extracurricular activities (homework help by parents, private tutors) (see appendix A1).

In order to illustrate selection mechanisms, table 1 compares the covariate statistics of both participating and nonparticipating students in sample 5–7 with regard to all three treatments. The findings suggest that these treatment and control groups had heterogeneous characteristics. For example, students who participated in homework support or remedial education had significantly lower pretest scores in mathematics and reading, poorer grades, and more often received homework help from a tutor than nonparticipants. With regard to subject-specific programs, participating and nonparticipating students differed in fewer

respects. In sample 7–9, we found similar differences in student background variables that reflect mechanisms of selection into treatment (see appendix A2).

Table 1. Treatment and control groups’ comparison statistics of covariates in sample 5–7.

	Homework support			Remedial education			Subject-specific programs		
	<i>M(SD)</i> or % in categ. treated	<i>M(SD)</i> or % in categ. controls	<i>p</i> value group differ- ence	<i>M(SD)</i> or % in categ. treated	<i>M(SD)</i> or % in categ. controls	<i>p</i> value group differ- ence	<i>M(SD)</i> or % in categ. treated	<i>M(SD)</i> or % in categ. controls	<i>p</i> value group differ- ence
Mathematics pretest	-0.33 (1.14)	0.27 (1.08)	0.00	-0.17 (1.01)	0.23 (1.12)	0.00	-0.02 (1.04)	0.19 (1.12)	0.12
Reading pretest	-0.22 (1.32)	0.23 (1.16)	0.00	-0.30 (1.11)	0.24 (1.20)	0.00	-0.08 (1.14)	0.17 (1.20)	0.08
Latest school grade mathematics	3.61 (0.90)	3.91 (0.83)	0.00	3.67 (0.78)	3.89 (0.86)	0.00	3.75 (0.82)	3.87 (0.85)	0.36
Latest school grade German	3.48 (0.86)	3.89 (0.78)	0.00	3.55 (0.87)	3.87 (0.79)	0.00	3.64 (0.92)	3.83 (0.80)	0.09
School satisfaction	7.51 (2.57)	7.86 (2.37)	0.13	7.21 (2.59)	7.90 (2.36)	0.00	7.72 (2.47)	7.80 (2.40)	0.78
Books at home	2.75 (1.37)	3.12 (1.30)	0.00	2.91 (1.26)	3.08 (1.33)	0.19	2.93 (1.30)	3.06 (1.32)	0.49
Parent employment yes	80.3%	71.1%	0.15	67.9%	73.6%	0.18	53.5%	74.2%	0.00
Homework help parents	2.85 (1.03)	2.65 (0.98)	0.03	2.57 (0.94)	2.70 (1.00)	0.19	2.67 (1.03)	2.68 (0.99)	0.94
Homework help tutor	0.52 (1.15)	0.17 (0.63)	0.00	0.49 (1.11)	0.18 (0.66)	0.00	0.48 (1.13)	0.21 (0.72)	0.04

Note. Based on weighted imputed data. $N=1131$ students at 43 schools.

Missing Values

We imputed missing data that occur on student level with regard to all outcomes, treatments, and covariates five times by using predictive mean matching. The imputation model used all aforementioned treatments, outcomes, and covariates along with further variables from student, parent, and principal surveys, which are all listed in appendix A3. The same variables were included in the imputation models of both samples. We included respective sampling weights to replicate the data structure. Dummy variables for school identification were included to account for the nested data structure (cf. Lüdtke, Robitzsch, & Grund, 2017). All further analyses were replicated for the five imputed datasets, and the estimates were combined using Rubin’s (1987) rules.

Analysis Plan

A standard issue in observational studies is that the treatment may be correlated with other predictors of the outcome. In the present study, we used longitudinal data from a representative sample of German nonmandatory all-day schools to estimate the effects of attending extracurricular activities on student achievement. The empirical issues can be depicted in a simple linear model in which the achievement Y of student i in school s is a function of his or her participation in extracurricular activities (E) and other confounding variables. It is useful to distinguish between confounding variables at the individual (I) and school (S) levels:

$$Y_{is} = b_0 + b_1 E_{is} + b_2 I_{is} + b_3 S_s + e_{is} \quad (1)$$

We were mainly interested in estimating b_1 , which is the effect of attending an extracurricular activity on achievement, while holding other predictors of student performance constant. To identify this parameter, the error term e_{is} has to be orthogonal to other explanatory variables and especially to participation in extracurricular activities. We applied two strategies to meet this assumption at school and individual levels.

At the school level, unobserved school composition or school quality measures may be correlated with the provision of extracurricular courses. If, for example, schools operating in deprived areas provide more extracurricular learning opportunities, children from disadvantaged families might participate more frequently in such activities. Hence, the estimation of the effect of extracurricular activities would be biased because e_{is} would be correlated with the explanatory variable. In the same vein, other school measures like educational tracks may be correlated with the supply of extracurricular activities. To avoid bias from any selection mechanisms at a school level, we estimated a fixed effects model to which we added school dummies (μ_s). School fixed effects can be interpreted as different intercepts for all schools, which is why any variation in outcomes owed to school characteristics is absorbed in these school fixed effects. By implication, we estimated b_1 based on the variation in participation in extracurricular activities within schools (i.e., we compared participating and nonparticipating students from the same school). The fact that this approach avoids school feature bias is reflected in equation (2), which no longer contains any school-level covariates:

$$Y_{is} = b_0 + b_1 E_{is} + b_2 I_{is} + \mu_s + e_{is} \quad (2)$$

Selection mechanisms within schools remain an issue in the identification of the effects of extracurricular activities. For example, students who are low performing or from disadvantaged families were found to enroll in homework support more often (see table 1 and appendix A2). We attempted to address this issue by including extensive controls for prior achievement and family background measures. For this purpose, we used rich test and survey information from achievement tests, student questionnaires, and parent computer-assisted telephone interviews that were conducted prior to the treatment (see appendix A1). Furthermore, some students attended more than one extracurricular activity at a time, which could have led to an overestimation of the effects of the respective activities. To address this issue, we included all three treatment variables simultaneously in the analysis models.

One tradeoff that arises when using extensive control variables is the increasing risk of an overadjustment bias, where some covariates are correlated with the error term. This leads to an underestimation of the treatment effects (see Castellano, Rabe-Hesketh, & Skrondal, 2014). To address this issue, we introduced the covariates stepwise.¹³

We replicated the analyses for two outcomes (mathematics, reading) and two samples (sample 5–7, sample 7–9), resulting in four sets of models. The NEPS target population are students in schools offering lower secondary education in Germany. NEPS applied a stratified two-stage probability-proportional-to-size sampling strategy to sample schools as primary sampling units at the first stage and students at the second stage. We used sampling weights to account for different sampling probabilities as well as initial and wave-specific nonresponse in order to generalize our findings to the underlying student population (for technical details see Steinhauer & Zinn, 2016).

Results

As a baseline for our main models with school fixed effects and extensive control variables, we regressed the outcomes on participation in the three extracurricular activities without any

¹³ A further issue of regression adjustments relates to the linearity assumption. To address this issue, we applied nonparametric propensity score matching to adjust for confounding variables (Stuart, 2010). In the matching model, we included all above mentioned student covariates and used 1:1 nearest-neighbor matching with replacement and calipers. Furthermore, we conducted exact matching on the school identification variable. The matching analyses lead to qualitatively the same results as the regression adjustments.

further controls. These analyses estimated the cross-sectional association between the outcomes and participation in extracurricular activities. In the sample from grade 5 to 7, students who had attended homework support or remedial education in grade 6 had lower mathematics and reading achievement scores in grade 7. Participation in subject-specific programs was not significantly associated with performance levels (see table 2, columns 1 and 5). In the sample 7–9, only participating in homework support was significantly related to lower mathematics achievement in grade 9 (see table 3, columns 1 and 5). However, cross-sectional estimates were affected by selection mechanisms at the school and individual levels and should not be interpreted as causal effects.

Table 2. Treatment effects on mathematics and reading in sample 5–7.

	Mathematics					Reading			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
Homework support	-0.40*	0.01	0.03	0.05	-0.39*	-0.07	-0.04	-0.11	
	(0.11)	(0.08)	(0.08)	(0.08)	(0.11)	(0.09)	(0.09)	(0.11)	
Remedial education	-0.26*	0.00	-0.01	-0.03	-0.38*	-0.09	-0.08	-0.10	
	(0.11)	(0.07)	(0.07)	(0.08)	(0.13)	(0.11)	(0.10)	(0.11)	
Subject-specific programs	-0.02	0.06	0.06	0.01	-0.35	-0.26	-0.25	-0.29	
	(0.15)	(0.11)	(0.11)	(0.11)	(0.18)	(0.14)	(0.14)	(0.15)	
Mathematics pretest	-	0.65*	0.56*	0.55*	-	0.28*	0.25*	0.23*	
		(0.03)	(0.03)	(0.03)		(0.04)	(0.04)	(0.05)	
Reading pretest	-	0.18*	0.17*	0.16*	-	0.49*	0.45*	0.41*	
		(0.03)	(0.03)	(0.03)		(0.04)	(0.04)	(0.04)	
Latest school grade mathematics	-	-	0.28*	0.24*	-	-	0.03	-0.03	
			(0.04)	(0.04)			(0.05)	(0.05)	
Latest school grade German	-	-	-0.03	-0.06	-	-	0.21*	0.17*	
			(0.05)	(0.05)			(0.05)	(0.05)	
School satisfaction	-	-	-0.01	-0.01	-	-	0.00	0.01	
			(0.01)	(0.01)			(0.01)	(0.01)	
Books at home	-	-	0.05*	0.03	-	-	0.00	0.00	
			(0.02)	(0.02)			(0.03)	(0.03)	
Parent employment	-	-	0.00	-0.06	-	-	0.06	0.04	
			(0.06)	(0.07)			(0.08)	(0.09)	
Homework help parents	-	-	-0.02	-0.03	-	-	-0.02	-0.03	
			(0.03)	(0.03)			(0.04)	(0.04)	
Homework help tutor	-	-	0.05	0.04	-	-	0.06	0.05	
			(0.04)	(0.04)			(0.05)	(0.05)	
School fixed effects	no	no	no	yes	no	no	no	yes	

Note. Based on weighted imputed data; unstandardized *b* values; standard errors in parentheses; * $p < 0.05$; $N = 1131$ students at 43 schools.

In the main analyses, we regressed the outcomes on the treatment variables while controlling stepwise for prior achievement, further student covariates, and school fixed effects. After controlling for prior achievement, the effects of the treatments homework support, remedial education, and subject-specific programs on mathematics and reading achievement were not

significantly different from zero (tables 2 and 3, columns 2 and 6). The results were qualitatively the same after controlling for further student covariates (tables 2 and 3, columns 3 and 7) and estimating school fixed effects (tables 2 and 3, columns 4 and 8). These findings further support the assumption that negative associations between outcomes and treatments in models without covariate control and school fixed effect estimation were confounded by selection mechanisms. The main analyses provided no evidence for an effect of participation in extracurricular activities on student achievement.

Table 3. Treatment effects on mathematics and reading in sample 7–9.

	Mathematics				Reading			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Homework support	-0.39*	-0.10	-0.11	-0.09	-0.06	0.17	0.16	0.07
	(0.16)	(0.12)	(0.12)	(0.12)	(0.17)	(0.15)	(0.15)	(0.15)
Remedial education	-0.18	0.01	0.09	0.11	-0.12	0.00	0.03	-0.05
	(0.16)	(0.14)	(0.13)	(0.10)	(0.15)	(0.12)	(0.12)	(0.10)
Subject-specific programs	-0.03	0.03	0.04	0.02	-0.12	-0.03	-0.02	0.14
	(0.19)	(0.15)	(0.15)	(0.10)	(0.20)	(0.16)	(0.16)	(0.10)
Mathematics pretest	-	0.58*	0.51*	0.46*	-	0.20*	0.19*	0.12*
		(0.07)	(0.06)	(0.03)		(0.05)	(0.05)	(0.04)
Reading pretest	-	0.14*	0.14*	0.08*	-	0.39*	0.35*	0.30*
		(0.03)	(0.03)	(0.02)		(0.03)	(0.03)	(0.03)
Latest school grade mathematics	-	-	0.15*	0.21*	-	-	-0.03	0.00
			(0.06)	(0.03)			(0.04)	(0.03)
Latest school grade German	-	-	0.00	0.01	-	-	0.15*	0.15*
			(0.06)	(0.05)			(0.06)	(0.05)
School satisfaction	-	-	-0.01	-0.01	-	-	-0.02	-0.02
			(0.02)	(0.01)			(0.01)	(0.01)
Books at home	-	-	0.03	0.03	-	-	0.06	0.04
			(0.03)	(0.02)			(0.03)	(0.02)
Parent employment	-	-	-0.16	-0.08	-	-	0.03	0.00
			(0.21)	(0.07)			(0.11)	(0.10)
Homework help parents	-	-	-0.02	-0.01	-	-	-0.03	-0.02
			(0.06)	(0.03)			(0.03)	(0.03)
Homework help tutor	-	-	-0.07	-0.03	-	-	-0.01	0.00
			(0.07)	(0.02)			(0.03)	(0.03)
School fixed effects	no	no	no	yes	no	no	no	yes

Note. Based on weighted imputed data; unstandardized *b* values; standard errors in parentheses; * $p < 0.05$;

$N = 1545$ students at 64 schools

Discussion

The present study investigates whether attending academic extracurricular activities results in learning benefits for students at nonmandatory all-day schools in Germany. The results suggest that participating in homework support, remedial education, and subject-specific programs has no effects on mathematics and reading achievements in the course of secondary

schooling. These results are in line with previous research that showed no effects of attending all-day schools' extracurricular activities on achievement in Germany (Bellin & Tamke, 2010; Fischer et al., 2016; Lossen et al., 2016).

Another strand of findings concerns participation characteristics. We found relatively low rates of attendance in the investigated academic program types. Further, we found selection mechanisms associated with student attendance. Students who participated in homework support and remedial education especially tended to come from a disadvantaged academic and social background. These findings are in line with theoretical and empirical deliberations in the literature, too (Fischer et al., 2016; Fischer & Klieme, 2013; Lossen et al., 2016; Steiner, 2011). We also found that students attending the programs more frequently made use of private tutoring. Considering the remedial nature of the selection mechanisms, this might not be surprising. However, in light of international research in this field, one might conclude that observed treatments may relate to unobserved additional treatments.

With regard to policymaking, the major result of our study is that the investigated forms of all-day schooling are not found to foster student learning and therefore seem insufficiently designed and organized for this particular purpose. International examples show that programs with high-quality characteristics can serve as valuable supports for student learning (e.g., Apsler, 2009; Biggart et al., 2013; Durlak et al., 2010; Lauer et al., 2006). This implies that promising international findings could be replicated when investigating high-quality extracurricular activities. However, they remain rare in Germany. For example, the schools' extracurricular programs are not necessarily connected to the morning's classes or supervised by teachers or other highly qualified staff (e.g., Fischer & Klieme, 2013; StEG-Consortium, 2015, 2016; Steiner, 2011).

The present study has substantive and methodological merits and limitations, as the dataset it used was not specifically designed for investigating the present research questions. A first set of issues relates to the measures of the explanatory variables. We used fairly simple treatment variables, namely participation in extracurricular activities. An advantage of NEPS is that it surveys various common activities, including programs with an explicit focus on learning, as investigated in the analyses. At the same time, future research may investigate less common but possibly effective academic activities that were not surveyed in NEPS. In this regard, it may also be worth investigating domain-specific courses because such specific interventions are conceptually more closely related outcomes in the respective domains. However, previous studies on such domain-specific programs found no significant

participation effects (Fischer et al., 2016; Lossen et al., 2016). Additionally, even nonacademic treatments might affect student learning indirectly (e.g., participation in sports may increase cognitive capacities). Another limitation relates to how the participation in all-day activities was measured in NEPS. The binary treatments of participation vs. nonparticipation were only assessed once (in grade 6 respectively 8) between the achievement tests. Therefore, their minimal dosage was a half year of participation, and we were not able to investigate more continuous attendance.

Apart from the simple treatment variables, it must be acknowledged that we were unable to further investigate possible mediating roles of program quality such as the qualification of supervising staff. One possible explanation for our neutral results is that the quality of extracurricular activities in German all-day schools is too low on average. This explanation is largely consistent with meta-analyses showing that extracurricular activities have no effect on student achievement unless they meet high quality standards (Apsler, 2009; Durlak et al., 2010; Roth et al., 2010). The heterogeneous quality of German all-day schools and of the activities can generally be expected to lead to heterogeneous effects on student outcomes. This relates to general limitations of research with observational data. Given that we find a neutral overall effect, however, this would imply that negative effects in some schools counteract positive effects in others. Following this argument, the recent expansion of all-day schools in Germany apparently did not lead to the development of high-quality extracurricular programs in the afternoon, at least not overall (cf. also StEG-Consortium, 2016). In the present study, we were however only able to investigate the overall effect of unstandardized treatments that are subsumed under the terms homework support, remedial education, and subject-specific programs. Even students from the same schools might have attended different courses for homework support, for instance. Future research with experimental designs could provide insights into the effects of well-defined extracurricular interventions on closely connected student outcomes.

Furthermore, the scope of our study is limited to specific educational outcomes. Although we considered key measures of student achievement in different grades and the core domains of mathematics and reading, there are other important outcomes such as emotions, attitudes, social and problematic behavior (e.g., drug use), or school dropout (cf. Fischer et al., 2014; Fischer & Klieme, 2013; StEG-Consortium, 2016; Vandell et al., 2007). For example, previous studies found that all-day schooling positively affected student outcomes like learning goal orientation, academic self-concepts, motivation, social behavior, and school

grades. Typically, such effects were mediated by the programs' quality characteristics and by continuous and intensive student participation (cf. Fischer et al., 2016; Kuhn & Fischer, 2011; Lossen et al., 2016; Sauerwein, Theis, & Fischer, 2016; StEG, 2016). The inconsistent finding on effects of all-day schooling on school grades and test scores in particular deserves closer attention. If one assumes that teacher judgements are not only influenced by students' performance but also by their behavior (e.g., completion of homework, social behavior in class) or motivation, the inconsistency may suggest that all-day schooling affects behavioral and motivational constructs rather than performance.

For a holistic and realistic evaluation of all-day schooling, it seems worthwhile acknowledging noneducational outcomes. The increase in parental labor market participation is another aim of all-day schools. Some initial studies suggest that the recent extension of all-day schools in Germany has had a positive impact on mothers' labor market involvement (Bauernschuster & Schlotter, 2015; Gambaro, Marcus, & Peter, 2016). This finding reflects that extending all-day schools has led to an increase in the time students spend at school. From an educational perspective, however, a simple extension of time seems to be insufficient to affect students' learning outcomes measured in standardized tests.

Other issues are linked to the identification of effects with nonexperimental data. As a first step our analytical approach relies on school fixed effects that effectively control for any confounding variables at a school level and by implication also at higher levels. Within schools, we then considered student-level confounders because students (or their parents) select into participation. We sought to minimize the risk of confounded effect estimates by using rich pretest and background data as controls. The most important remaining risk is time-invariant confounding, which is correlated with the treatment and outcome variables but not with the observed covariates.¹⁴ Related to the previous point, the control of pretest and key background measures in linear models may not be effective if the assumption of predictor linearity does not hold. We partly addressed this issue by replicating the analyses using nonparametric propensity score matching in unreported analyses. These analyses qualitatively replicated the findings of our main analyses. However, two limitations must be stressed. First, although we applied different matching algorithms, we were not able to remove all differences between the treatment and control groups. A reasonable explanation for this is that

¹⁴ However, controlling for observed covariates also (at least partly) controls for unobserved covariates if they are correlated with the observed ones (Stuart, 2010).

the exact matching on the school makes it difficult to find good matches because the number of classmates is limited within schools. A second limitation could be even more complex selection mechanisms, which would have to be modeled as interaction terms between covariates, for example. Again, sample size sets a natural limitation for investigations with the present data.

A last methodological note relates to the regression-towards-the-mean phenomenon. We observed that participation in extracurricular activities is partly due to a selection based on prior school grades. Such differences may not only result from actual performance differences but also in part from chance. If so, regression towards the mean could mask the positive effects of extracurricular activities on achievement, resulting in a neutral overall effect. Again, as we included not only grades but also various test scores as controls, we think that regression toward the mean did not considerably bias our effect estimates.

However, our results should be interpreted with caution, as it is difficult—even with longitudinal data—to identify the impact of attending extracurricular activities on student achievement. Obviously it would be easier to identify this effect in carefully implemented small-scale interventions with randomized group allocation. However, the external validity of results is a common problem for experimental approaches, which is why we believe that the present study is nonetheless an important complement to the state of research. We think that our results should contribute to policy discussions on the quality of extracurricular activities when they are implemented on a large scale.

References

- Apsler, R. (2009). After-school programs for adolescents: A review of evaluation research. *Adolescence, 44*, 1–19.
- Battistin, E., & Meroni, E. C. (2016). Should we increase instruction time in low achieving schools? Evidence from Southern Italy. *Economics of Education Review, 55*, 39–56. <https://doi.org/10.1016/j.econedurev.2016.08.003>
- Bauernschuster, S. & Schlotter, M. (2015). Public child care and mothers' labor supply— Evidence from two quasi-experiments. *Journal of Public Economics, 123*, 1–16.
- Bellei, C. (2009). Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile. *Economics of Education Review, 28*(5), 629–640. <https://doi.org/10.1016/j.econedurev.2009.01.008>
- Bellin, N., & Tamke, F. (2010). Bessere Leistungen durch Teilnahme am offenen Ganztagsbetrieb? *Empirische Pädagogik, 24*(2), 93–112.
- Biggart, A., Kerr, K., O'Hare, L., & Connolly, P. (2013). A randomised control trial evaluation of a literacy after-school programme for struggling beginning readers. *International Journal of Educational Research, 62*, 129–140. <https://doi.org/10.1016/j.ijer.2013.07.005>
- Blossfeld, H.-P., Roßbach, H.-G., & von Maurice, J. (Eds.) (2011). Education as a lifelong process – The German National Educational Panel Study (NEPS). *Zeitschrift für Erziehungswissenschaft: Sonderheft 14*.
- Castellano, K. E., Rabe-Hesketh, S., & Skrandal, A. (2014). Composition, context, and endogeneity in school and teacher comparisons. *Journal of Educational and Behavioral Statistics, 39*(5), 333–367.
- Cooper, H., Charlton, K., Valentine, J. C., Muhlenbruck, L., & Borman, G. D. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development, 65*(1), 1–127.
- Duchhardt, D., & Gerdes, A. (2012). *NEPS technical report for mathematics - scaling results of starting cohort 3 in fifth grade (NEPS Working Paper No. 19)*. Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.
- Durlak, J. A., Weissberg, R. P., & Pachan, M. (2010). A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents. *American journal of community psychology, 45*(3-4), 294–309. <https://doi.org/10.1007/s10464-010-9300-6>

- Fashola, O. S. (1998). *Review of extended-day and after-school programs and their effectiveness. Report: No. 24*. Baltimore, MD: Center for Research on the Education of Students Placed at Risk.
- Feldman, A. F., & Matjasko, J. L. (2005). The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. *Review of Educational Research*, 75(2), 159–210.
- Fischer, N., & Klieme, E. (2013). Quality and effectiveness of German all-day schools. Results of the study on the development of all-day schools in Germany. In J. Ecarius, E. Klieme, L. Stecher, & J. Woods (Eds.), *Extended education – an international perspective: Proceedings of the international conference on extracurricular and out-of-school time educational research* (pp. 27–52). Opladen, Berlin, Toronto: Barbara Budrich Publishers.
- Fischer, N., Sauerwein, M. N., Theis, D., & Wolgast, A. (2016). Vom Lesenlernen in der Ganztagschule: Leisten Ganztagsangebote einen Beitrag zur Leseförderung am Beginn der Sekundarstufe I? *Zeitschrift für Pädagogik*, 62(6), 780–796.
- Fischer, N., Theis, D., & Züchner, I. (2014). Narrowing the gap? The role of all-day schools in reducing educational inequality in Germany. *International Journal for Research on Extended Education, IJREE*, 2(1), 79–96.
- Gambaro, L., Marcus, J., & Peter, F. (2016). *School entry, afternoon care and mother's labour supply*: DIW Discussion Papers, No. 1622.
- KMK – Standing Conference of the Ministers of Education and Cultural Affairs of the Laender (2008). *Allgemein bildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: Statistik 2002 bis 2006*, Bonn.
- KMK – Standing Conference of the Ministers of Education and Cultural Affairs of the Laender (2017). *Allgemeinbildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland*. Retrieved from <https://www.kmk.org/dokumentation-statistik/statistik/schulstatistik/allgemeinbildende-schulen-in-ganztagsform.html>
- Kuger, S. (2016). Curriculum and learning time in international school achievement studies. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), *Assessing contexts of learning* (pp. 395–422). Cham: Springer International Publishing.
- Kuhn, H. P., & Fischer, N. (2011). Entwicklung der Schulnoten in der Ganztagschule. Einflüsse der Ganztagsbeteiligung und der Angebotsqualität. In E. Klieme, T. Rauschenbach, L. Stecher, & I. Züchner (Eds.), *Ganztagschule: Entwicklungen*,

- Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG) (207–226). Weinheim, Basel: Beltz Juventa.*
- Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research, 76*(2), 275–313.
<https://doi.org/10.3102/00346543076002275>
- Lossen, K., Tillmann, K., Holtappels, H. G., Rollett, W., & Hannemann, J. (2016). Entwicklung der naturwissenschaftlichen Kompetenzen und des sachunterrichtsbezogenen Selbstkonzepts bei Schüler/-innen in Ganztagsgrundschulen: Ergebnisse der Längsschnittstudie StEG-P zu Effekten der Schülerteilnahme und der Angebotsqualität. *Zeitschrift für Pädagogik, 62*(6), 760–779.
- Lüdtke, O., Robitzsch, A., & Grund, S. (2017). Multiple imputation of missing data in multilevel designs: A comparison of different strategies. *Psychological methods, 22*(1), 141–165. <https://doi.org/10.1037/met0000096>
- Meyer, E., & van Klaveren, C. (2013). The effectiveness of extended day programs: Evidence from a randomized field experiment in the Netherlands. *Economics of Education Review, 36*, 1–11. <https://doi.org/10.1016/j.econedurev.2013.04.002>
- Miller, B. M., & Truong, K. A. (2009). The role of afterschool and summer in achievement. The untapped power of afterschool and summer to advance student achievement. *Zeitschrift für Pädagogik, 55*(54. Beiheft), 124–142.
- NEPS – National Educational Panel Study (2011). Starting cohort 3 main study 2010/11 (A28) students, 5th grade, regular schools: Information on the competence tests. Retrieved from https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/1-0-0/C_A28_en.pdf
- NEPS – National Educational Panel Study (2016). Starting cohort 3: grade 5 (SC3) wave 3: Information on the competence test. Retrieved from https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/3-1-0/C_A30_en.pdf
- OECD (2011). *PISA quality time for students: Learning in and out of school* (1. Aufl.). PISA: OECD Publishing. Retrieved from <http://dx.doi.org/10.1787/9789264087057-en>
- OECD (2015). PF4.3: Out-of-school-hours care services. Retrieved from <http://www.oecd.org/els/soc/PF4-3-Out-of-school-hours-care.pdf>

- Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the school day or school year: A systematic review of research (1985-2009). *Review of Educational Research, 80*(3), 401–436. <https://doi.org/10.3102/0034654310377086>
- Plantenga, J., & Remery, C. (2013). *Childcare services for school age children: A comparative review of 33 countries*. Luxembourg: Publications Office of the European Union.
- Pohl, S., & Carstensen, C. H. (2012). *NEPS technical report – scaling the data of the competence tests (NEPS Working Paper No. 14)*. Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.
- Pohl, S., Haberkorn, K., Hardt, K., & Wiegand, E. (2012). *NEPS technical report for reading - scaling results of starting cohort 3 in fifth grade (NEPS Working Paper No. 15)*. Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.
- Prein, G., Rauschenbach, T., & Züchner, I. (2009). Analysen zur Selektivität von offenen Ganztagschulen. In F. Prüß, S. Kortas, & M. Schöpa (Eds.), *Die Ganztagschule: von der Theorie zur Praxis: Anforderungen und Perspektiven für Erziehungswissenschaften und Schulentwicklung* (81-97). Weinheim: Juventa.
- Roth, J. L., Malone, L. M., & Brooks-Gunn, J. (2010). Does the amount of participation in afterschool programs relate to developmental outcomes? A review of the literature. *American journal of community psychology, 45*(3-4), 310–324. <https://doi.org/10.1007/s10464-010-9303-3>
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.
- Sauerwein, M., Theis, D., & Fischer, N. (2016). How youths' profiles of extracurricular and leisure activity affect their social development and academic achievement. *International Journal for Research on Extended Education, IJREE, 4*(1), 103–124.
- Schüpbach, M. (2014). Extended education and social inequality in Switzerland: Compensatory effects? *Journal for Educational Research Online, 6*(3), 95–114.
- Schüpbach, M. (2015). Effects of extracurricular activities and their quality on primary school-age students' achievement in mathematics in Switzerland. *School Effectiveness and School Improvement, 26*(2), 279–295. <https://doi.org/10.1080/09243453.2014.929153>

- Scott-Little, C., Hamann, M. S., & Jurs, S. G. (2002). Evaluations of after-school programs: A meta-evaluation of methodologies and narrative synthesis of findings. *American Journal of Evaluation*, 23(4), 387–419.
- Stecher, L., & Maschke, S. (2013). Research on extended education in Germany - A general model with all-day schooling and private tutoring as two examples. *International Journal for Research on Extended Education, IJREE*, 1(1), 31–52.
- StEG-Consortium (2015). Ganztagsschule 2014/2015: Deskriptive Befunde einer bundesweiten Befragung. Retrieved from http://www.projektsteg.de/sites/default/files/StEG_Bundesbericht%202015_final_0.pdf
- StEG-Consortium (2016). *Ganztagsschule: Bildungsqualität und Wirkungen außerunterrichtlicher Angebote: Ergebnisse der Studie zur Entwicklung von Ganztagsschulen 2012-2015*. Retrieved from http://www.projektsteg.de/sites/default/files/StEG_Brosch_FINAL.pdf
- Steiner, C. (2011). Teilnahme am Ganztagsbetrieb: Zeitliche Entwicklung und mögliche Selektionseffekte. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher, & I. Züchner (Eds.), *Ganztagsschule: Entwicklung, Qualität, Wirkungen: Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagsschulen (StEG)* (pp. 57–75). Weinheim und Basel: Beltz Juventa.
- Steinhauer, H. W. & Zinn, S. (2016). *NEPS technical report for weighting: Weighting the sample of starting cohort 3 of the National Educational Panel Study (Waves 1 to 5)*. Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Strietholt, R., Manitius, V., Berkemeyer, N., & Bos, W. (2015). Bildung und Bildungsungleichheit an Halb- und Ganztagsschulen. *Zeitschrift für Erziehungswissenschaft*, 18(4), 737–761. <https://doi.org/10.1007/s11618-015-0634-6>
- Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science : a review journal of the Institute of Mathematical Statistics*, 25(1), 1–21. <https://doi.org/10.1214/09-STS313>
- Vandell, D., Reisner, E., & Pierce, K. (2007). *Outcomes linked to highquality afterschool programs: Longitudinal findings from the study of promising afterschool programs*. Washington, D.C.: Policy Studies Associates Inc.

Appendices

Appendix A1. Descriptive covariate statistics for sample 5–7 and sample 7–9.

	Sample 5–7			Sample 7–9		
	Source	Missing	<i>M(SD)</i> or % in categories	Source	Missing	<i>M(SD)</i> or % in categories
Mathematics pretest	student test grade 5	5.4%	0.17 (1.11)	student test grade 7	3.6%	0.93 (1.25)
Reading pretest	student test grade 5	5.3%	0.15 (1.20)	student test grade 7	3.5%	0.89 (1.38)
Latest school grade mathematics ¹	student quest. grade 5	9.5%	3.86 (0.85)	student quest. grade 7	11.4%	3.26 (0.97)
Latest school grade German ¹	student quest. grade 5	9.9%	3.82 (0.81)	student quest. grade 7	10.5%	3.35 (0.80)
School satisfaction ²	student quest. grade 5	7.2%	7.80 (2.41)	student quest. grade 7	4.0%	6.95 (2.27)
Books at home ³	student quest. grade 5	10.1%	3.05 (1.32)	student quest. grade 7	5.9%	3.11 (1.44)
Parent employment ⁴	parent inter. grade 5	32.2%	72.7% yes	parent inter. grade 7	41.8%	78.7% yes
Homework help parents ⁵	student quest. grade 5	8.7%	2.68 (1.00)	student quest. grade 7	8.1%	2.33 (1.02)
Homework help tutor ⁵	student quest. grade 5	14.6%	0.23 (0.76)	student quest. grade 7	8.9%	0.52 (1.15)

Note. Based on weighted imputed data; ¹ from 0=“failing” to 5=“very good”; ² from 0=“completely dissatisfied” to 10=“completely satisfied”; ³ from 0=“0-10 books” to 5=“>500 books”; ⁴ yes=“parent employed as a side-job, in part-time, or in full-time”; ⁵ from 0=“never” to 4=“always”.

Appendix A2. Treatment and control groups' comparison statistics of covariates in sample 7–9.

	Homework support			Remedial education			Subject-specific programs		
	<i>M(SD)</i> or % in categ. treated	<i>M(SD)</i> or % in categ. controls	<i>p</i> value group differ- ence	<i>M(SD)</i> or % in categ. treated	<i>M(SD)</i> or % in categ. controls	<i>p</i> value group differ- ence	<i>M(SD)</i> or % in categ. treated	<i>M(SD)</i> or % in categ. controls	<i>p</i> value group differ- ence
Mathematics pretest	0.43 (1.19)	0.97 (1.25)	0.00	0.39 (1.15)	0.97 (1.25)	0.00	0.82 (1.33)	0.93 (1.25)	0.40
Reading pretest	0.33 (1.45)	0.95 (1.36)	0.00	0.48 (1.39)	0.93 (1.37)	0.01	0.68 (1.38)	0.91 (1.38)	0.22
Latest school grade mathematics	3.19 (1.03)	3.27 (0.96)	0.46	3.22 (0.93)	3.26 (0.97)	0.68	3.11 (1.08)	3.27 (0.96)	0.16
Latest school grade German	3.18 (0.88)	3.37 (0.79)	0.04	3.17 (0.82)	3.37 (0.79)	0.02	3.06 (0.83)	3.37 (0.79)	0.00
School satisfaction	6.23 (2.59)	7.01 (2.23)	0.01	6.48 (2.36)	6.99 (2.26)	0.02	6.32 (2.52)	6.99 (2.24)	0.01
Books at home	2.62 (1.42)	3.15 (1.43)	0.00	2.76 (1.50)	3.14 (1.43)	0.06	3.15 (1.49)	3.10 (1.43)	0.80
Parent employment yes	77.8% yes	78.8% yes	0.83	80.7% yes	78.5% yes	0.78	71.9% yes	79.1% yes	0.31
Homework help parents	2.27 (1.06)	2.33 (1.01)	0.62	2.35 (0.96)	2.33 (1.02)	0.85	2.32 (1.08)	2.33 (1.01)	0.91
Homework help tutor	1.25 (1.55)	0.45 (1.08)	0.00	0.97 (1.52)	0.48 (1.10)	0.00	0.83 (1.39)	0.50 (1.13)	0.09

Note. Based on weighted imputed data. *N*=1545 students at 64 schools.

Appendix A3. Variables in the imputation model.

	Source	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9
Student achievement variables						
Mathematics ¹	student test	X		X		X
Reading ¹	student test	X		X		X
Orthography ¹	student test	X		X		X
Perceptual speed ²	student test	X				X
Cognitive reasoning ²	student test	X				X
Declarative metacognition ³	student test		X			X
ICT literacy ¹	student test		X			X
Scientific competence ¹	student test		X			X
Reading speed ²	student test	X				X
All-day related variables						
Student use of						
Homework support ⁴	student quest.		X		X	
Remedial education ⁴	student quest.		X		X	
Subject-specific programs ⁴	student quest.		X		X	
Subject-unrelated projects ⁴	student quest.		X		X	
Leisure facilities ⁴	student quest.		X		X	
Participation frequency ⁵	student quest.		X		X	
School offer of						
Homework supervision ⁶	principal quest.		X		X	
Remedial instruction ⁶	principal quest.		X		X	

Beiträge der Arbeit | Beitrag IV

Subject-specific offering in mathematics ⁶	principal quest.	X	X	
Subject-specific offering in German ⁶	principal quest.	X	X	
Sports offerings ⁶	principal quest.	X	X	
Music offerings ⁶	principal quest.	X	X	
Free-time activities ⁶	principal quest.	X	X	
Student background variables				
Latest school grade mathematics ⁷	student quest.	X	X	X
Latest school grade German ⁷	student quest.	X	X	X
School satisfaction ⁸	student quest.	X	X	X
Books at home ⁹	student quest.	X	X	X
Parent ISCED ¹⁰	parent inter.	X	X	X
Partner ISCED ¹⁰	parent inter.	X	X	X
Parent employment ¹¹	parent inter.	X	X	X
Partner employment ¹¹	parent inter.	X	X	X
Family native-born ¹²	student quest.	X	X	
Language at home ¹³	parent inter.	X	X	
Homework help parents ¹⁴	student quest.	X	X	
Homework help tutor ¹⁴	student quest.	X	X	
Gender ¹⁵	cohort profile	X	X	
Age ¹⁶	cohort profile	X	X	

Note. ¹ Weighted maximum likelihood estimates; ² sum scores; ³ mean scores; ⁴ yes="participation"; ⁵ general participation in hours per week; ⁶ from 0="less than weekly offer" to 3="4-5 times per week"; ⁷ from 0="failing" to 5="very good"; ⁸ from 0="completely dissatisfied" to 10="completely satisfied"; ⁹ from 0="0-10 books" to 5=">500 books"; ¹⁰ from 0="inadequately completed general education" to 10="doctoral degree and postdoctoral lecture qualification"; ¹¹ yes="parent employed as a side-job, in part-time, or in full-time"; ¹² yes="child, parents, and grandparents born in Germany"; ¹³ yes="language at home is other than German"; ¹⁴ from 0="never" to 4="always"; ¹⁵ 0="male" and 1="female"; ¹⁶ in years.

3 Gesamtdiskussion

3.1 Zentrale Befunde

Im Folgenden werden die Vorgehensweisen und Befunde der vier Einzelbeiträge entlang der Forschungsfragen der Dissertation resümiert (s. Abschnitt 1.3). In separaten Abschnitten folgen Reflexionen der Reichweite und Implikationen der Arbeiten sowie ein knappes Fazit.

Die erste Forschungsfrage betraf, inwiefern Ganztagschulen im Vergleich zu Halbtagschulen eine neue Organisationsform mit erweiterten Bildungspotenzialen darstellen und daher über halbtägige Beschulung hinausgehende Effekte erwarten lassen. Hierzu untersuchte der erste Beitrag, wie ganztägige Beschulung vor dem Hintergrund politischer Entwicklungen und Vorgaben ausgestaltet wird und ordnete verschiedene Befunde zur Umsetzung des Ganztagskonzepts in den Kontext pädagogischer Wirkungsannahmen ein. In dem Handbuchartikel wurde herausgearbeitet, dass Ganztagschulen grundsätzlich ein pädagogisches Potenzial zur Leistungssteigerung und zum Bildungsungleichheitsabbau haben, das über Halbtagsmodelle hinausgeht (vgl. z.B. Fischer, Theis & Züchner, 2014; Miller & Truong, 2009; Radisch, 2009; Steiner, 2009). Gleichzeitig werden Ganztagschulen sowohl quantitativ als auch qualitativ sehr unterschiedlich ausgestaltet. Insbesondere offene, additive und nicht von der Schulleitung gesteuerte Ausgestaltungsformen lassen sich lediglich graduell von Halbtagschulen mit Nachmittagsangebot abgrenzen (z.B. vergleichbare Angebote und Teilnahmecharakteristika der Schüler/-innen) (vgl. Holtappels, 2006; StEG-Konsortium, 2015; Wendt, Goy & Walzebug, 2016; Willems, Wendt & Radisch, 2015). Der Beitrag kam zu dem Schluss, dass die KMK-Definition eher strukturelle als pädagogische Kriterien abbildet, während letztere als Voraussetzung für eine leistungssteigernde und ungleichheitsreduzierende Funktion des Ganztags verstanden werden. Daher lassen sich nur bedingt Mehrwerte von Ganztags- im Vergleich zu Halbtagschulen erwarten, insbesondere im Fall von Effekten auf Testleistungen und soziale Ungleichheiten in Testleistungen.

In dem zweiten Beitrag wurde die Forschungsfrage untersucht, ob es zu Beginn des Ganztagschulausbaus einen Effekt des schulischen Anbietens nachmittäglicher Förderungsprogramme auf Leistungsentwicklungen von Schüler/-innen oder auf herkunftsbedingte Bildungsungleichheiten gab. Zunächst ist festzuhalten, dass zu Beginn des IZBB trotz bildungspolitischer Postulierung keine empirischen Belege für die politischen Wirkungsannahmen einer leistungssteigernden und/oder ungleichheitsreduzierenden Funktion

der Ganztagschule vorlagen (vgl. z.B. Höhmann, Holtappels & Schnetzer, 2004; KMK, 2002; Radisch, 2009). Die empirische Studie schloss an dieses Forschungsdesiderat an und analysierte eine Stichprobe von $N=8958$ Neuntklässler/-innen an 184 Regelschulen aus der repräsentativen DESI-Studie aus dem Schuljahr 2003/04 (Deutsch-Englisch-Schülerleistungen-International; Klieme, 2012). In Mehrebenenregressionen wurden sprachliche Fähigkeiten am Ende der neunten Klasse auf das schulische Angebot von Ganztagsprogrammen sowie die Interaktion mit dem sozialen Status regressiert. Hierbei wurden die individuelle Vorleistung, Schulform und sozioökonomische Schulkomposition kontrolliert. Sprachliche Fähigkeiten wurden zu Beginn und Ende des neunten Schuljahres anhand standardisierter Tests in den Bereichen Sprachbewusstheit in Deutsch und Textrekonstruktion in Englisch untersucht. Soziale Bildungsungleichheiten wurden als Zusammenhang dieser Leistungen mit der Anzahl zuhause vorhandener Bücher modelliert, beziehungsweise in weiterführenden Analysen als Zusammenhang mit dem elterlichen Beruf und der Muttersprache. Die betrachteten Ganztagsangebote wurden mindestens wöchentlich angeboten und umfassten Hausaufgabenbetreuung, rhythmisierten Ganztagsunterricht, freiwillige Arbeitsgemeinschaften in Deutsch und Englisch sowie in weiterführenden Analysen Angebote in den Bereichen Sport und Freizeit sowie Musik und Kunst. Des Weiteren wurde untersucht, ob sich Effekte bei einem (nur selten vorkommenden) mehr als wöchentlichen Angebot zeigten. In Bezug auf die Forschungsfrage zeigten sich weder in den Haupt- noch weiterführenden Analysen Effekte des schulischen Anbietens von Ganztagsangeboten auf sprachliche Leistungen oder soziale Bildungsungleichheiten. Darüber hinaus variierten verschiedene Schulformen in ihren Ganztagsangeboten.

Die dritte Forschungsfrage nach Unterschieden zwischen Halb- und Ganztagschulen im späteren Verlauf des Ganztagschulenausbaus in Bezug auf Leistungsentwicklungen von Schüler/-innen und der sozialen Ungleichheit in diesen Leistungsentwicklungen wurde im dritten Beitrag untersucht. Diese Forschungsfrage baute darauf auf, dass durch das IZBB auch eine qualitative Weiterentwicklung ganztägiger Beschulung angestrebt wurde (BMBF, 2003a, 2003b), sodass sich nach dem Abrufzeitraum auch erstmalig erhoffte Effekte zeigen könnten. Daher fokussierte die zweite empirische Studie den Zeitraum zwischen den Schuljahren 2010/11 und 2014/15 und nutzte eine Stichprobe von $N=3024$ Fünftklässler/-innen an 144 Regelschulen der repräsentativen NEPS-Studie (National Educational Panel Study; Blossfeld, Roßbach & von Maurice, 2011), die bis zur neunten Klasse an derselben Schule blieben. In linearen latenten Mehrebenenwachstumsmodellen wurden die Leistungsentwicklungen

beziehungsweise Entwicklungen der sozialen Ungleichheit zwischen drei Messzeitpunkten in Klasse 5, 7 und 9 untersucht. Diese Leistungs- und Ungleichheitsentwicklungen wurden auf die Organisationsform Halb- versus Ganztagschule regressiert, während die Schulform, Schulkomposition und in weiteren Analysen auch die Bundesländerzugehörigkeit kontrolliert wurden. Als Leistungsmaße lagen jeweils standardisierte Tests in den Bereichen Mathematik und Lesen zugrunde. Bildungsungleichheiten wurden als Zusammenhang der Leistung mit der elterlichen Bildung beziehungsweise Muttersprache operationalisiert (in weiteren Analysen mit der Anzahl zuhause vorhandener Bücher und den Geburtsländern der Eltern und Schüler/-innen). Auf Ebene der Schulen wurde kontrastiert, ob diese zwischen den Schuljahren 2010/11 und 2014/15 konstant Halbtags- oder Ganztagschulen blieben oder ob sie ihre Organisationsform wechselten. Außerdem wurde vertiefend geprüft, ob Ganztagschulen mit günstigen Teilnahmemerkmalen der Schüler/-innen (hohe zeitliche Intensität und hohe Teilhabequoten an akademisch ausgerichteten Programmen) Mehrwerte gegenüber Halbtagschulen aufwiesen. Mit Bezug auf die Forschungsfrage zeigten sich weder in den Haupt- noch weiterführenden Analysen Unterschiede in der Leistungs- und Bildungsungleichheitsentwicklung zwischen fünfter und neunter Klasse. Untersucht wurden Schulen, die über diesen Zeitraum durchgehend Halb- oder Ganztagschulen waren. Darüber hinaus waren die Teilnahmequoten an leistungsförderlich ausgerichteten Programmen mit höchstens 15,0% der Schüler/-innen eher gering ausgeprägt. Die Ganztagschulen waren im Vergleich zu den Halbtagschulen häufiger Gymnasien und hatten eher leistungsstarke und privilegierte Schüler/-innen.

In dem vierten Beitrag wurde die Forschungsfrage untersucht, ob es Unterschiede in der Leistungsentwicklung von Schüler/-innen an offenen Ganztagschulen gab, die an lernförderlich ausgerichteten Ganztagsangeboten teilnahmen versus nicht teilnahmen. Diese Forschungsfrage schließt an die Annahme an, dass lernförderliche Effekte ganztägiger Beschulung am ehesten für solche Schüler/-innen zu erwarten sind, die an akademischen Ganztagsangeboten teilnehmen (vgl. z.B. Lossen, Tillmann, Holtappels, Rollett & Hannemann, 2016; Steiner, 2011b; Tillmann, Sauerwein, Hannemann, Decristan, Lossen & Holtappels, 2018). Da die Teilnahmequoten jedoch gerade bei lernförderlich ausgerichteten Programmen und an offenen Ganztagschulen eher gering ausgeprägt sind (z.B. Holtappels, 2008; Holtappels, Jarsinski & Rollett, 2011; StEG-Konsortium, 2015), lässt sich annehmen, dass Ganztagschuleffekte potenziell eher im Vergleich teilnehmender und nichtteilnehmender Schüler/-innen zu finden sind, als auf der aggregierten Schulebene.

Weiterhin wird angenommen, dass für Schüler/-innen offener Ganztagschulen, die nicht am Ganzttag teilnehmen, keine Unterschiede im Vergleich zu einem Halbtagsschulbesuch bestehen (vgl. Bellin & Tamke, 2010; KMK, 2018). Aufgrund dessen analysierte die dritte empirische Studie Teilnahmeeffekte auf Individualebene für den Zeitraum 2010/11 bis 2014/15. Es wurden zwei Stichproben aus der repräsentativen NEPS-Studie betrachtet: Zum einen wurden $N=1131$ Schüler/-innen untersucht, die zwischen fünfter und siebter Klasse an denselben 43 offenen Ganztagschulen blieben, zum anderen $N=1545$ Schüler/-innen, die zwischen siebter und neunter Jahrgangsstufe dieselben 64 offenen Ganztagschulen besuchten. Die Stichproben lagen separaten Regressionsmodellen zugrunde, in denen Leistungsmaße auf die Teilnahme an akademischen Ganztagsangeboten regressiert wurden. Hierbei wurden umfangreiche Schüler/-innenmerkmale wie die Vorleistung und der familiäre Hintergrund kontrolliert, sowie schulfixierte Effekte berechnet. Als Leistungsmaße dienten Testergebnisse aus fünfter und siebter beziehungsweise siebter und neunter Klasse in Mathematik und Lesen. Die untersuchten Ganztagsangebote umfassten Hausaufgabenbetreuung, Förderunterricht und fachbezogene Lernangebote. Es wurden ausschließlich offene Ganztagschulen betrachtet, damit teilnehmende und nichtteilnehmende Schüler/-innen innerhalb derselben Schule verglichen werden konnten. Mit Blick auf die Forschungsfrage war der zentrale Befund, dass keine Leistungsunterschiede zwischen an akademischen Ganztagsangeboten teilnehmenden und nichtteilnehmenden Schüler/-innen derselben offenen Ganztagschulen gefunden wurden, weder zwischen fünfter und siebter, noch zwischen siebter und neunter Jahrgangsstufe. Darüber hinaus zeigten sich Selektionseffekte im Zugang zu den Angeboten, sowie insgesamt niedrige Teilnahmequoten an akademischen Programmen. Außerdem nutzten am Ganzttag Teilnehmende häufiger privaten Nachhilfeunterricht als Nichtteilnehmende.

Das zentrale Forschungsanliegen der Dissertation betraf Mehrwerte der Ganztagschulen gegenüber Halbtagschulen in Hinblick auf die Leistungsentwicklung von Schüler/-innen und den Abbau sozialer Bildungsungleichheiten. Zusammenfassend deuten die Befunde der vier Beiträge darauf hin, dass Ganztagschulen zwar ein theoretisches Potenzial in Hinblick auf entsprechende Effekte haben, in den vorliegenden empirischen Studien jedoch keine Effekte ganztägiger Beschulung auf Leistungs- und Ungleichheitsmaße gefunden werden.

3.2 Reichweite und Limitationen

Die vorliegende Dissertation umfasst drei empirische Studien und einen Handbuchartikel. In Hinblick auf die Reichweite des Gesamtwerks besteht der besondere Mehrwert des Handbuchartikels darin, dass ein allgemeinerer Blick auf das Forschungsfeld vorgenommen wird, indem beispielsweise auch weitere Wirkungsebenen der Ganztagschule betrachtet werden (z.B. Familienleben, motivationale oder verhaltensbezogene Wirkungen auf Schüler/-innen). Die Reichweite und Limitationen der empirischen Primäranalysen werden im Folgenden entlang der verwendeten Stichproben, Methoden, unabhängigen und abhängigen Variablen überblickt sowie in den bisherigen Forschungsstand eingeordnet.

3.2.1 Stichproben

Die drei empirischen Beiträge basierten auf Stichproben der für Deutschland repräsentativen Studien DESI und NEPS. Bei Berücksichtigung der entsprechenden Surveygewichte erlauben die Daten Schlüsse auf alle Regelschulen der Sekundarstufe I (Beck & Klieme, 2007; Blossfeld et al., 2011; Steinhauer & Zinn, 2016). Insgesamt wurden Schüler/-innen zwischen der fünften und neunten Klassenstufe untersucht. Die Befunde sind daher weder direkt auf Förder-, noch Grundschulen zu übertragen. Sowohl die Repräsentativität der zugrundeliegenden Datensätze als auch das Abdecken der Sekundarstufe I stellen besondere Mehrwerte für den Forschungsbereich dar. Die meisten Vorgängerstudien mit robustem Design basierten auf nichtrepräsentativen Daten (Bellin & Tamke, 2010; Fischer, Sauerwein, Theis & Wolgast, 2016; Lossen et al., 2016; Tillmann et al., 2018). Eine Studie nutzte zwar repräsentative Stichproben von Grundschulkindern und Fünfzehnjährigen, verwendete jedoch ein quasi-längsschnittliches Design (Strietholt, Manitius, Berkemeyer & Bos, 2015). Eine weitere analysierte zwar ebenfalls repräsentative Daten der NEPS-Studie, jedoch nur zwischen fünfter und siebter Klasse (Linberg, Struck & Bäumer, 2018).

Ein weiterer Mehrwert der vorliegenden Dissertation liegt in den abgedeckten Untersuchungszeiträumen (2003/04 und 2010/11 bis 2014/15). Wie beispielsweise Holtappels (2006) betont, ist denkbar, dass das Ganztagskonzept vor und nach dem IZBB qualitativ verschieden ausgestaltet wurde und unterschiedliche Wirkungen aufwies. Die empirischen Arbeiten decken sowohl den Zeitraum zu Beginn des IZBB ab, für den bis dato keine robusten Befunde vorlagen (vgl. auch Radisch, 2009), als auch den nach Abschluss des Investitionsprogramms.

3.2.2 Methodisches Vorgehen

In Bezug auf das methodische Vorgehen hatten alle drei Studien zum Ziel, bei der Analyse von Beobachtungsdaten robuste Identifikationsstrategien anzuwenden, sodass kausale Inferenzen abgeleitet werden können. Zunächst ist zu betonen, dass alle drei Studien ein längsschnittliches Untersuchungsdesign mit Modellierungen von zwei oder drei Messzeitpunkten verwendeten. Hierbei wurde die hierarchische Datenstruktur von Schüler/-innen in Schulen auf verschiedene Weise berücksichtigt. Der zweite Beitrag verwendete Mehrebenenregressionen und der dritte lineare latente Mehrebenenwachstumsmodelle. In beiden Beiträgen wurden zusätzlich relevante Schulmerkmale wie die Schulform und Schulkomposition kontrolliert (vgl. Castellano, Rabe-Hesketh & Skrandal, 2014). Der vierte Beitrag, der Effekte auf Individualebene untersuchte, verwendete Regressionsanalysen mit einem umfangreichen Set an Kontrollvariablen auf Individualebene sowie schulfixierte Effekte, durch die Schüler/-innen innerhalb derselben Schulen verglichen und Konfundierungen auf Schul- oder höheren Ebenen umgangen wurden.

Eine weitere Stärke der empirischen Arbeiten besteht in den vielfältigen Replikationen der Analysen (z.B. verschiedene Stichproben, alternative Operationalisierungen) sowie in der Prüfung der Ergebnisse bei Verwendung alternativer Identifikationsstrategien (z.B. Zuwachswertanalyse, vorheriges Propensity Score Matching). Des Weiteren wurden gemäß aktueller Standards multiple Imputationen für fehlende Werte vorgenommen (vgl. Little & Rubin, 2002; Rubin, 1987) und Surveygewichte in den Analysen verwendet, um die Befunde auf die zugrundeliegenden Populationen generalisieren zu können. Die Studien decken Untersuchungszeiträume innerhalb eines Jahres, innerhalb von zwei und innerhalb von vier Jahren ab. Im Gegensatz dazu nutzten viele Vorgängerstudien querschnittliche Datengrundlagen, die keine kausalen Schlüsse zuließen (z.B. Hertel, Klieme, Radisch & Steinert, 2008; Holtappels, Radisch, Rollett & Kowoll, 2010; Radisch, Klieme & Bos, 2006; Willems et al., 2015). Bei den längsschnittlichen Studien lagen zwischen den Messzeitpunkten weniger als ein Jahr bis zu zwei Jahren (Bellin & Tamke, 2010; Fischer et al., 2016; Linberg et al., 2018; Lossen et al., 2016; Tillmann et al., 2018).

Gleichzeitig bleibt festzuhalten, dass sowohl die vorherigen als auch die vorliegenden Studien auf Beobachtungsdaten basieren, was im Vergleich zu experimentellen Designs mit methodischen Einschränkungen verbunden ist. Ein besonderer Mehrwert insbesondere repräsentativer Beobachtungsdaten ist jedoch die ökologische Validität der Befunde, wie sie bei experimentellen Designs häufig als eingeschränkt kritisiert wird.

3.2.3 Unabhängige Variablen

In Bezug auf die unabhängigen Variablen wurden in den vorliegenden Studien verschiedene, sich ergänzende Zugänge und Operationalisierungen gewählt. Generell ist die Identifikation unabhängiger Variablen bei der Untersuchung ganztägiger Beschulung aufgrund der Heterogenität an Ausgestaltungen und Implikationen des Ganztagsbegriffs schwieriger als beispielsweise bei regulärem Unterricht (vgl. auch Fiester, Simpkins & Bouffard, 2005; Züchner & Fischer, 2011). Darüber hinaus wurden die verwendeten Datengrundlagen nicht für die vorliegenden Fragestellungen entwickelt. Alle vorherigen und vorliegenden Studien müssen sich auf eine eher summative Evaluation beschränken und können jeweils nur einzelne Qualitätsmerkmale ganztägiger Beschulung berücksichtigen (vgl. auch Fischer et al., 2016; Lossen et al., 2016; Tillmann et al., 2018).

Zunächst wurden im zweiten Beitrag Effekte des mindestens wöchentlichen schulischen Ganztagsangebots in bestimmten Inhaltsbereichen untersucht, das mit dem Vormittagsunterricht konzeptionell verbunden war und unter Verantwortung der Schulleitung stand. Ein häufigeres als wöchentliches Angebot wurde in weiteren Analysen betrachtet. Die Angebote umfassten rhythmisierten Ganztagsunterricht, Hausaufgabenbetreuung, freiwillige Arbeitsgemeinschaften in Deutsch und Englisch, Angebote in den Bereichen Sport und Freizeit sowie Musik und Kunst. Dies ist die erste Studie, die Ganztageffekte auf diese Weise untersuchte; die Informationen zu den schulischen Angeboten wurden jedoch retrospektiv erhoben.

Komplementierend wurden im dritten Beitrag Ganztags- und Halbtagschulen gemäß der KMK-Definition unterschieden. Die Organisationsforminformation wurde zu Beginn, in der Mitte und zu Ende des Untersuchungszeitraums erhoben, sodass drei Gruppen unterschieden werden konnten: Schulen, die in den betrachteten vier Jahren Ganztags- beziehungsweise Halbtagschulen blieben und solche, die ihre Organisationsform änderten. Fast ein Drittel der betrachteten Schulen wechselte die Organisationsform (vgl. auch Steiner, 2011b). Diese Unterteilung in konstante Ganz- und Halbtagschulen im Gegensatz zu Mischschulen ermöglicht eine trennschärfere Identifikation der untersuchten Effekte, als wenn einmalige Informationen zur Organisationsform herangezogen werden. In diesem Fall können auch solche Schulen in eine der Kategorien fallen, die kurz vor einem Organisationsformwechsel stehen oder ihn gerade erst vollzogen haben. Einmalige Messzeitpunkte der Schulorganisation wurden in beiden (quasi-)längsschnittlichen Vorgängerstudien auf Schulebene herangezogen (Linberg et al., 2018; Strietholt et al., 2015).

Aufgrund der Stichprobengröße war es in der Studie nur bedingt möglich, zusätzlich differenzielle Effekte qualitativer Ausgestaltungsmerkmale zu untersuchen. In vertiefenden Analysen wurde jedoch die Hälfte der Ganztagschulen fokussiert, die potenziell günstigere Teilnahmecharakteristika aufwies, was in keiner Vorgängerstudie untersucht wurde.

In dem vierten Beitrag wurden zwei Stichproben von Schüler/-innen herangezogen, deren Schulen über die zweijährigen Untersuchungszeiträume hinweg offene Ganztagschulen blieben. Daher wurde auch hier der häufigen Organisationsformwechsel Rechnung getragen. Es wurde die Teilnahme versus Nichtteilnahme an den fachlich ausgerichteten Ganztagsangeboten Hausaufgabenbetreuung, Förderunterricht und fachbezogene Lernangebote untersucht. Die Teilnahme an fachbezogenen Lernangeboten wurde bislang von keiner anderen Studie untersucht. Diese inhaltliche Ausrichtung der Programme auf Lernzuwächse steht der Untersuchung der allgemeinen Teilnahme bei Bellin und Tamke (2010) entgegen. Die Angebote wurden jedoch nicht domänenspezifisch abgefragt, wie beispielsweise bei der Untersuchung von Effekten der Teilnahme an naturwissenschaftsbezogenen Programmen auf Sachunterrichtsleistungen bei Fischer und Kolleg/-innen (2016). Außerdem umfasste die Teilnahme lediglich eine Minimaldauer von einem halben Jahr; es konnten keine Effekte einer dauerhafteren Teilnahme oder von der Teilnahme an mehr als einem Programm untersucht werden.

3.2.4 Abhängige Variablen

Die vorliegende Dissertation fokussiert Effekte der Ganztagschule auf Testleistungen und Bildungsungleichheiten in diesen Leistungen. Ein besonderer Mehrwert des ersten Beitrags liegt darin, dass hier auch Befunde anderer Studien zu beispielsweise nichtstandardisierten Schulnoten, motivations- und verhaltensbezogenen Maßen oder der sozialen Chancengleichheit im Zugang zu Lerngelegenheiten thematisiert werden. Stärken der vorliegenden empirischen Arbeiten liegen darin, dass sowohl die abhängigen Variablen der Leistung als auch Bildungsungleichheit jeweils auf längsschnittlich erhobenen und verknüpften standardisierten Tests und unterschiedlichen Operationalisierungen basieren (Beck, Bundt & Gomolka, 2008; Hartig, Jude & Wagner, 2008; Leibniz-Institut für Bildungsverläufe, 2018a, 2018b, 2018c). Die untersuchten Kompetenzbereiche umfassen die Mathematik, Sprachbewusstheit und Lesefähigkeit in Deutsch sowie die Textrekonstruktion in Englisch als Fremdsprache. Sie bilden somit wesentliche Bereiche schulischer Lernfortschritte ab. Frühere Studien untersuchten ebenfalls Lesen und Mathematik als

abhängige Maße, nicht jedoch die Sprachbewusstheit in Deutsch und Fremdsprachkenntnisse in Englisch (Bellin & Tamke, 2010; Fischer et al., 2016; Linberg et al., 2018; Strietholt et al., 2015; Tillmann et al., 2018). In zwei früheren Untersuchungen wurde zusätzlich die naturwissenschaftliche Kompetenz betrachtet (Lossen et al., 2016; Strietholt et al., 2015).

Der zweite und dritte Beitrag untersuchten neben Effekten auf Leistungsentwicklungen auf Schulebene zusätzlich Effekte auf Bildungsungleichheiten in diesen Maßen, die verschieden operationalisiert wurden (vgl. auch Jerrim & Micklewright, 2014). In beiden Studien wurde die Bildungsungleichheit als Ausmaß des Zusammenhangs der Leistungsmaße mit dem sozioökonomischen und migrationsbezogenen Hintergrund der Schüler/-innen modelliert. Der sozioökonomische Hintergrund wurde hierbei anhand der Anzahl zuhause vorhandener Bücher, der elterlichen Bildung oder des elterlichen Berufs operationalisiert. Die migrationsbezogene Herkunft wurde in Form der Muttersprache oder Geburtsländer der Eltern und Schüler/-innen aufgenommen. Nur eine vorherige Studie untersuchte den Einfluss des Ganztags auf das Ausmaß sozialer Bildungsungleichheit an Schulen (operationalisiert anhand des elterlichen Berufs und der Geburtsländer) und nutzte hierbei ein quasi-längsschnittliches Design (Strietholt et al., 2015). Auf der Ebene von Teilnahmeeffekten untersuchten Bellin und Tamke (2010) Schüler/-innen mit Migrationshintergrund als separate Gruppe.

Zusammenfassend lässt sich sagen, dass die vorliegenden empirischen Arbeiten sowohl in Bezug auf die untersuchten Stichproben, verwendeten Methoden, unabhängigen und abhängigen Variablen zwar einigen Einschränkungen unterliegen, den bisherigen Forschungsstand jedoch in bedeutenden Hinsichten ergänzen.

3.3 Implikationen

Dem Forschungsanliegen der vorliegenden Dissertation folgend (s. Abschnitt 1.3) wurde zunächst herausgearbeitet, dass Ganztagschulen gegenüber Halbtagschulen ein gesteigertes lernförderliches und bildungsungleichheitsabbauendes theoretisches Potenzial aufweisen. In empirischen Studien wurden jedoch keine Effekte ganztägiger Beschulung auf Leistungs- und Bildungsungleichheitsmaße gefunden.

In Bezug auf bildungspolitische Implikationen der vorliegenden Arbeit ist zunächst festzuhalten, dass die hier betrachteten politischen Annahmen einer leistungssteigernden und bildungsungleichheitsreduzierenden Funktion von Ganztagschulen zum Zeitpunkt der Verabschiedung des IZBB nicht evidenzbasiert waren (vgl. auch Böttcher, 2015; Radisch, 2009). Wie die vorliegende Dissertation und frühere Studien zeigen, bestehen weder für den Beginn noch weiteren Verlauf des Ganztagschulenausbaus empirische Untermauerungen für diese politischen Annahmen (Bellin & Tamke, 2010; Fischer et al., 2016; Linberg et al., 2018; Lossen et al., 2016; Strietholt et al., 2015; Tillmann et al., 2018). Hierbei muss betont werden, dass das IZBB weitere politische Zielsetzungen verfolgte und demnach auch weitere Wirkungen der Ganztagschule untersucht werden können (s. folgenden Abschnitt 3.4). In Hinblick auf Effekte auf standardisiert gemessene Leistungen oder Bildungsungleichheiten implizieren die Befunde jedoch, dass die Ganztagschulen beziehungsweise Ganztagsangebote durchschnittlich so ausgestaltet zu sein scheinen, dass sich mit den verwendeten State of the Art Evaluationsformaten keine Effekte nachweisen lassen (vgl. Holtappels, 2006; StEG-Konsortium, 2015; Züchner & Fischer, 2011). In diesem Zusammenhang erscheint bedeutsam, dass die KMK-Ganztagsdefinition kaum pädagogisch qualitätssichernde Merkmale, sondern vor allem strukturelle Vorgaben enthielt. Außerdem konnten die fast viereinhalb Milliarden Euro IZBB-Mittel für bauliche und sachliche Ausstattungen, nicht jedoch für beispielsweise Personalaufstockungen verwendet werden (BMBF, 2003a, 2003b). Wie im ersten Beitrag diskutiert, können diese regulatorischen Vorgaben dazu beigetragen haben, dass sich eine sehr heterogene Ganztagschullandschaft entwickelt hat, die, nach Implikationen pädagogischer Wirkmodelle, nicht durchweg leistungssteigernde Effekte erwarten lässt (Böttcher, 2015; Holtappels, 2006; Radisch, 2009; StEG-Konsortium, 2015; Tillmann, 2011; Wendt et al., 2016). Aus pädagogischer Sicht lassen sich über Halbtagschulen hinausgehende Bildungsmehrwerte am ehesten von Ganztagschulen erhoffen, deren Ganztagsprogramm inhaltlich und zeitlich umfangreich, teilnahmegebunden, mit dem Unterricht verknüpft, zeitrythmisiert und von qualifiziertem

Personal gestaltet wird (Miller & Truong, 2009; Radisch, 2009; Stecher, Radisch, Fischer & Klieme, 2007; Steiner, 2009). Solche Schulen sind bislang selten (Fischer, Holtappels, Klieme, Rauschenbach, Stecher & Züchner, 2011; StEG-Konsortium, 2015, 2016), was auch die Ergebnisse der vorliegenden Dissertation erklären könnte.

Auch aktuell bestehen politische Interessen an einem weiteren Ganztagschulausbau mit vergleichbaren politischen Zielsetzungen wie zu Beginn des IZBB (KMK, 2015; Koalitionsparteien, 2018). In Bezug auf mögliche künftige Investitionen in einen weiteren Ganztagsausbau lassen sich aus dem geschilderten Forschungsstand erste Hinweise auf mögliche Schlussfolgerungen ableiten. Die Befunde deuten darauf hin, dass eine im Vergleich zum IZBB engere Verknüpfung der Bewilligung von Fördermitteln an pädagogische Qualitätssicherungskriterien zu einer qualitativen Weiterentwicklung beitragen *könnte* (vgl. auch Böttcher, 2017; Tillmann, 2011). Potenziell können Schulen so stärker darin unterstützt und bestärkt werden, ein qualitativ hochwertiges, über Betreuung hinausgehendes Ganztagsangebot aufzubauen. Tatsächlich berichten Schulleitungen, dass insbesondere die fortlaufende Finanzierung des Ganztagsbetriebs und das Finden und Finanzieren von qualifiziertem Personal eine Herausforderung für die Ganztagschulentwicklung darstellt (Spillebeen, Holtappels & Rollett, 2011; StEG-Konsortium, 2013, 2015). Gleichzeitig ist keineswegs eindeutig, wie solche Qualitätssicherungskriterien aussehen müssten. Qualitätsmerkmale, die in pädagogischen Wirkmodellen (s. Abschnitt 1.1.2) und der US-amerikanischen After-School-Forschung (s. Abschnitt 1.2.2) diskutiert werden, sind bislang nicht systematisch und empirisch für ihre Gültigkeit im Ganztagsschulbereich evaluiert worden. Es ist unklar, inwieweit die theoretischen Annahmen in der Praxis zutreffen und die internationalen Befunde auf das deutsche Schulsystem übertragbar sind, da sich internationale Modelle ganztägiger Bildung und Betreuung defintorisch und organisational stark von Ganztagschulen unterscheiden können (Kuger, Klieme, Jude & Kaplan, 2016; OECD, 2011b; Plantenga & Remery, 2013). Um zukünftige bildungspolitische Entscheidungen zu weiteren Ganztagsschulreformen auf empirische Evidenz stützen zu können, sollten daher zunächst Schulversuche eingerichtet und mit vorzugsweise experimentellen Begleitstudien evaluiert werden (vgl. auch Deutscher Bildungsrat, 1968). Solche Schulversuche könnten erstens aufzeigen, wie theoretisch angenommene und mit internationaler Evidenz gestützte Qualitätsfaktoren (z.B. hochwertige personelle und pädagogische Ausstattung, Verknüpfung der Inhalte mit dem Unterricht) in Ganztagschulen zu implementieren sind und welche Voraussetzungen hierfür gegeben sein müssen. Zweitens kann so geprüft werden, unter

welchen Bedingungen oder Bedingungskonstellationen leistungsförderliche oder bildungsungleichheitsreduzierende Effekte eintreten können, die über halbtägige Beschulung hinausgehen. Solche Schulversuche könnten auf umfangreichen Erkenntnissen aus StEG und anderen Studien aufbauen (BMBF, 2012; Fischer et al., 2011; Holtappels, Klieme, Rauschenbach & Stecher, 2008; StEG-Konsortium, 2015, 2016) oder bestehende Idealvorstellungen ganztägiger Beschulung prüfen (z.B. Radisch, Klemm & Tillmann, 2017). In Bezug auf die Möglichkeiten, wissenschaftlich entwickelte Lerninterventionen in bestehende Ganztagsbetriebe zu implementieren, versprechen laufende StEG-Nachfolgestudien weiterführende Erkenntnisse (DIPF, 2018).

Darüber hinaus scheint eine Betrachtung von Ganztagschulen oder von internationalen Modellen erweiterter Bildung und Betreuung besonders gewinnbringend, wenn die Zeit in Lern- und Freizeitkontexten in einer Gesamtschau untersucht wird. In Frage kommen beispielsweise die Betrachtung von regulärem Unterricht, Selbstlernphasen (z.B. Hausaufgaben), der Teilnahme an extracurricularen Aktivitäten, privater Nachhilfe, aber auch der Ausgestaltung zuhause verbrachter Freizeit. Insbesondere eine international vergleichende Perspektive wäre hierbei interessant, da sowohl die stärker institutionalisierten als auch weniger formalisierten Lernzeiten international variieren. Sowohl die Ausmaße, Inhalte und Kontexte der verbrachten Zeit unterscheiden sich massiv sowohl über den Schultag, die Woche, aber auch das Jahr hinweg (Mahoney, Eccles & Larson, 2005; OECD, 2011b; Plantenga & Remery, 2013; Scheerens, 2014; Werquin, 2010). Dass es sinnvoll erscheint, verschiedene Zeitkontexte nicht nur additiv, sondern auch in ihrer Wechselwirkung zu untersuchen, legt auch ein Befund des vorliegenden vierten Beitrags nahe. Hier zeigte sich, dass zwischen der Nutzung ganztägiger Angebote in der Schule und zusätzlicher privater Nachhilfe ein positiver Zusammenhang bestand. Wenn also Effekte einzelner Lernkontexte herauskristallisiert werden sollen, scheint ein holistischer Blick auf das Verbringen von Lern- und Freizeitphasen gewinnbringend. Insbesondere auch in Hinsicht auf die Kopplung von Lernerfolgen und sozialen Hintergrundmerkmalen ist die Berücksichtigung der außerschulischen Freizeitgestaltung relevant (vgl. z.B. Boudon, 1974; Ditton, 2000; OECD, 2011b; Strietholt et al., in Druck).

3.4 Fazit

In den vergangenen Jahren hat eine enorme Zunahme von Ganztagsplätzen stattgefunden. Eine bedeutsame Wirkungsebene dieses Ausbaus betrifft die verbesserte Vereinbarkeit von Familie und Beruf, wofür in Deutschland wie auch international eine bezahlbare und verlässliche Betreuung unerlässlich ist (Börner, 2011; OECD, 2007, 2011a; Wissenschaftlicher Beirat für Familienfragen, 2006; Züchner, 2011). Darüber hinaus scheinen mit diesen Ganztagsangeboten auch Kinder aus sozial benachteiligten Familien erreicht zu werden, was im Vergleich zu privaten, typischerweise kostenpflichtigen Optionen zu einer chancengleicheren Teilhabe an beispielsweise Sport-, Kunst- und Musikangeboten beiträgt. Auch fachliche Unterstützungsmöglichkeiten wie Hausaufgabenbetreuung oder Förderunterricht sind an der Ganztagschule im Gegensatz zu beispielsweise privater Nachhilfe üblicherweise kostenfrei (vgl. Autorengruppe Bildungsberichterstattung, 2014, 2016; BMBF, 2012; Bundesministerium für Familie, Senioren, Frauen und Jugend, 2017; Stecher & Maschke, 2013; StEG-Konsortium, 2015, 2016; Züchner & Arnoldt, 2011). Darüber hinaus gibt es Hinweise darauf, dass Ganztags zu Entlastungen des Familienlebens beitragen kann, indem beispielsweise die Hausaufgaben erledigung in die Schule verlagert wird (Börner, 2011; StEG-Konsortium, 2016; Züchner, 2011). Auch liegen Befunde dazu vor, dass ganztägige Beschulung unter bestimmten Voraussetzungen positive Effekte auf Schüler/-innen haben kann, zum Beispiel auf die Kontinuität der Schullaufbahn (weniger Klassenwiederholungen und Schulwechsel), oder motivationale oder verhaltensbezogene Eigenschaften von Schüler/-innen (Fischer, Brümmer & Kuhn, 2011; Fischer, Kuhn & Züchner, 2011; Lossen et al., 2016; StEG-Konsortium, 2016; Steiner, 2011a).

Betrachtet man jedoch standardisierte Leistungsmaße und Bildungsungleichheiten in diesen Leistungsmaßen als abhängige Maße, dann zeigt die vorliegende Dissertation, dass sich bislang keine signifikanten Effekte ganztägiger Beschulung nachweisen lassen. Dies gilt erstens für den Beginn des Ganztagschulausbaus, für den Effekte schulischer Ganztagsangebote auf sprachliche Leistungen und Bildungsungleichheiten in Deutsch und Englisch untersucht wurden. Zweitens fanden sich auch im späteren Verlauf des Ganztagschulausbaus keine Unterschiede in der Entwicklung von Leistung und Bildungsungleichheit in Mathematik und Lesen zwischen Halb- und Ganztagschulen. Drittens schließlich wiesen an lernförderlich ausgerichteten Ganztagsangeboten teilnehmende Schüler/-innen offener Ganztagschulen keine mathematischen oder Leseentwicklungsvorteile

gegenüber nichtteilnehmenden auf. Die Befunde der Dissertation sind hierbei konform mit dem weiteren Forschungsstand (s. Abschnitt 1.2.1). Einsichten theoretischer Modelle und internationaler empirischer Befunde deuten darauf hin, dass sich dieses Ausbleiben statistisch signifikanter Effekte mit einer Erhöhung der Qualität ganztägiger Beschulung ändern *könnte*. Dies spricht dafür, bei künftigen Investitionen einen stärkeren Fokus auf eine qualitative statt quantitative Weiterentwicklung ganztägiger Beschulung zu legen (vgl. auch Böttcher, 2015; Ditton, 2000; Scheerens, 2014; StEG-Konsortium, 2016; Tillmann, 2011). Um evidenzbasierte Aussagen darüber treffen zu können, *wie* eine solche qualitative Weiterentwicklung konkret aussehen müsste, bedarf es jedoch weiterer, formativ ausgerichteter Forschung, etwa in Form von Schulversuchen. Die vorliegende Dissertation bietet hierfür Ansatzpunkte und erweitert den gegenwärtigen Stand summativer Evaluationsstudien in bedeutsamer Weise.

3.5 Literaturverzeichnis

- Autorengruppe Bildungsberichterstattung (2014). *Bildung in Deutschland 2014: Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: W. Bertelsmann Verlag.
- Autorengruppe Bildungsberichterstattung (2016). *Bildung in Deutschland 2016: Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: W. Bertelsmann Verlag.
- Beck, B., Bundt, S., & Gomolka, J. (2008). Ziele und Anlage der DESI-Studie. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 11–25). Weinheim, Basel: Beltz.
- Beck, B., & Klieme, E. (Hrsg.) (2007). *Sprachliche Kompetenzen: Konzepte und Messung. DESI-Studie (Deutsch-Englisch-Schülerleistungen-International)*. Weinheim, Basel: Beltz.
- Bellin, N., & Tamke, F. (2010). Bessere Leistungen durch Teilnahme am offenen Ganztagsbetrieb? *Empirische Pädagogik*, 24(2), 93–112.
- Blossfeld, H.-P., Roßbach, H.-G., & von Maurice, J. (2011). Education as a lifelong process – The German National Educational Panel Study (NEPS). *Zeitschrift für Erziehungswissenschaft, Sonderheft 14*.
- BMBF (2003a). *Investitionsprogramm „Zukunft Bildung und Betreuung“: Ganztagschulen. Zeit für mehr*, Bundesministerium für Bildung und Forschung.
http://www.bmbf.de/pub/ganztagschulen-zeit_fuer_mehr.pdf.
- BMBF (2003b). *Verwaltungsvereinbarung Investitionsprogramm „Zukunft Bildung und Betreuung“ 2003-2007*.
https://www.ganztagschulen.org/_media/20030512_verwaltungsvereinbarung_zukunft_bildung_und_betreuung.pdf.
- BMBF (2012). *Ganztägig bilden: Eine Forschungsbilanz*.
http://www.ganztagschulen.org/_media/121206_BMBF_GTS-Forschungsbilanz_bf_df.pdf.
- Börner, N. (2011). Ganztagschule und ihre Auswirkungen auf Familien. *Zeitschrift für Erziehungswissenschaft*, 14, 221–236.
- Böttcher, W. (2015). Ganze Tage in der Schule: Politik und Wissenschaft zwischen Anspruch und Wirklichkeit. In T. Hascher, T. S. Idel, S. Reh & W. Thole (Hrsg.), *Bildung über*

- den ganzen Tag. Forschungs- und Theorieperspektiven der Erziehungswissenschaft* (S. 39–53). Opladen, Berlin, Toronto: Barbara Budrich Publishers.
- Böttcher, W. (2017). Steuerung? Welche Steuerung? In A. Bolder, H. Bremer & R. Epping (Hrsg.), *Bildung für Arbeit unter neuer Steuerung* (S. 73–96). Wiesbaden: Springer Fachmedien Wiesbaden.
- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in Western society*. New York: Wiley.
- Bundesministerium für Familie, Senioren, Frauen und Jugend (2017). *15. Kinder- und Jugendbericht: Bericht über die Lebenssituation junger Menschen und die Leistungen der Kinder- und Jugendhilfe in Deutschland*.
<https://www.bmfsfj.de/blob/115438/d7ed644e1b7fac4f9266191459903c62/15-kinder-und-jugendbericht-bundestagsdrucksache-data.pdf>.
- Castellano, K. E., Rabe-Hesketh, S., & Skrandal, A. (2014). Composition, context, and endogeneity in school and teacher comparisons. *Journal of Educational and Behavioral Statistics*, 39(5), 333–367.
- Deutscher Bildungsrat (1968). *Einrichtung von Schulversuchen mit Ganztagschulen: Empfehlungen der Bildungskommission*. Bonn: Bundesdruckerei.
- DIPF (2018). *StEG: Die Teilprojekte der Studie*. <http://www.projekt-steg.de/content/steg-die-teilprojekte-der-studie>.
- Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. In A. Helmke, W. Hornstein & E. Terhart (Hrsg.), *Qualität und Qualitätssicherung im Bildungsbereich Schule, Sozialpädagogik, Hochschule* (S. 73–92). Weinheim: Beltz.
- Fiester, L. M., Simpkins, S. D., & Bouffard, S. M. (2005). Present and accounted for: Measuring attendance in out-of-school-time programs. *New Direction for Youth Development*, 105, 91-107.
- Fischer, N., Brümmer, F., & Kuhn, H. P. (2011). Entwicklung von Wohlbefinden und motivationalen Orientierungen in der Ganztagschule: Zusammenhänge mit der Prozess- und Beziehungsqualität in den Angeboten. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG)* (S. 227–245). Weinheim, Basel: Beltz Juventa.

- Fischer, N., Holtappels, H. G., Klieme, E., Rauschenbach, T., Stecher, L., & Züchner, I. (Hrsg.) (2011). *Ganztagsschule: Entwicklung, Qualität, Wirkungen: Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagsschulen (StEG)*. Weinheim, Basel: Beltz Juventa.
- Fischer, N., Kuhn, H. P., & Züchner, I. (2011). Entwicklung von Sozialverhalten in der Ganztagsschule: Wirkungen der Ganztagsbeteiligung und der Angebotsqualität. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagsschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagsschulen (StEG)* (S. 246–266). Weinheim, Basel: Beltz Juventa.
- Fischer, N., Sauerwein, M. N., Theis, D., & Wolgast, A. (2016). Vom Lesenlernen in der Ganztagsschule: Leisten Ganztagsangebote einen Beitrag zur Leseförderung am Beginn der Sekundarstufe I? *Zeitschrift für Pädagogik*, 62(6), 780–796.
- Fischer, N., Theis, D., & Züchner, I. (2014). Narrowing the gap? The role of all-day schools in reducing educational inequality in Germany. *International Journal for Research on Extended Education*, 2(1), 79–96.
- Hartig, J., Jude, N., & Wagner, W. (2008). Methodische Grundlagen der Messung und Erklärung sprachlicher Kompetenzen. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 34–54). Weinheim, Basel: Beltz.
- Hertel, S., Klieme, E., Radisch, F., & Steinert, B. (2008). Nachmittagsangebote im Sekundarbereich und ihre Nutzung durch die Schülerinnen und Schüler. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme et al. (Hrsg.), *PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (S. 297–318). Münster, New York, NY, München, Berlin: Waxmann.
- Höhm, K., Holtappels, H. G., & Schnetzer, T. (2004). Ganztagsschule: Konzeptionen, Forschungsbefunde, aktuelle Entwicklungen. In H. G. Holtappels, K. Klemm, H. Pfeiffer, H.-G. Rolff & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung. Daten, Beispiele und Perspektiven* (S. 253–290). Weinheim: Juventa.
- Holtappels, H. G. (2006). Stichwort: Ganztagsschule. *Zeitschrift für Erziehungswissenschaft*, 9(1), 5–29.
- Holtappels, H. G. (2008). Angebotsstruktur, Schülerteilnahme und Ausbaugrad ganztägiger Schulen. In H. G. Holtappels, E. Klieme, T. Rauschenbach & L. Stecher (Hrsg.),

- Ganztagsschule in Deutschland. Ergebnisse der Ausgangserhebung der „Studie zur Entwicklung von Ganztagschulen“ (StEG) (2. korrigierte Auflage) (S. 186–206).*
Weinheim und München: Juventa.
- Holtappels, H. G., Jarsinski, S., & Rollett, W. (2011). Teilnahme als Qualitätsmerkmal für Ganztagschulen: Entwicklung von Schülerbeteiligungquoten auf Schulebene. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagsschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG) (S. 97–119).* Weinheim, Basel: Beltz Juventa.
- Holtappels, H. G., Klieme, E., Rauschenbach, T., & Stecher, L. (Hrsg.) (2008). *Ganztagsschule in Deutschland: Ergebnisse der Ausgangserhebung der „Studie zur Entwicklung von Ganztagschulen“ (StEG) (2. korrigierte Auflage).* Weinheim und München: Juventa.
- Holtappels, H. G., Radisch, F., Rollett, W., & Kowoll, M. E. (2010). Bildungsangebot und Schülerkompetenzen in Ganztagsgrundschulen. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried & E.-M. Lankes (Hrsg.), *IGLU 2006 – die Grundschule auf dem Prüfstand (S. 165–198).* Münster: Waxmann.
- Jerrim, J., & Micklewright, J. (2014). Socio-economic gradients in children’s cognitive skills: Are cross-country comparisons robust to who reports family background? *European Sociological Review*, 30(6), 766–781.
- Klieme, E. (2012). *Deutsch-Englisch-Schülerleistungen-International (DESI)*. Version: 1. IQB – Institut zur Qualitätsentwicklung im Bildungswesen. Datensatz.
http://doi.org/10.5159/IQB_DESI_v1.
- KMK (2002). *PISA 2000 – Zentrale Handlungsfelder: Zusammenfassende Darstellung der laufenden und geplanten Maßnahmen in den Ländern (Stand 07.10.2002).*
https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2002/2002_10_07-Pisa-2000-Zentrale-Handlungsfelder.pdf.
- KMK (2015). *Ganztagsschulen in Deutschland: Bericht der Kultusministerkonferenz vom 03.12.2015.* Berlin, Bonn: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland.
- KMK (2018). *Allgemein bildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland: Statistik 2012 bis 2016.*

https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/GTS_2016_Bericht.pdf.

Koalitionsparteien (2018). *Ein neuer Aufbruch für Europa. Eine neue Dynamik für Deutschland. Ein neuer Zusammenhalt für unser Land: Koalitionsvertrag zwischen CDU, CSU und SPD*.

https://www.cdu.de/system/tdf/media/dokumente/koalitionsvertrag_2018.pdf?file=1.

Kuger, S., Klieme, E., Jude, N., & Kaplan, D. (Hrsg.) (2016). *Assessing contexts of learning*. Cham: Springer International Publishing.

Leibniz-Institut für Bildungsverläufe (2018a). *Information on competence testing: NEPS starting cohort 3 – grade 5: Paths through lower secondary school – Educational pathways of students in grade 5 and higher: Wave 1: grade 5*.

https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/1-0-0/NEPS_SC3_Compences_W1_en.pdf.

Leibniz-Institut für Bildungsverläufe (2018b). *Information on competence testing: NEPS starting cohort 3 – grade 5: Paths through lower secondary school – Educational pathways of students in grade 5 and higher: Wave 6: grade 9*.

https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/6-0-0/NEPS_SC3_Compences_W6_en.pdf.

Leibniz-Institut für Bildungsverläufe (2018c). *Information on competence testing: NEPS starting cohort 3 – grade 5: Paths through lower secondary school – Educational pathways of students in grade 5 and higher: Wave 3: grade 7*. https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/3-1-0/NEPS_SC3_Compences_W3_en.pdf.

Linberg, T., Struck, O., & Bäumer, T. (2018). Vorzug Ganztagschule? Zusammenhänge mit der Kompetenzentwicklung im Bereich Lesen und Mathematik. *Zeitschrift für Erziehungswissenschaft*, 14(2), 1–23.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2. Aufl.). Hoboken, NJ: Wiley.

Lossen, K., Tillmann, K., Holtappels, H. G., Rollett, W., & Hannemann, J. (2016). Entwicklung der naturwissenschaftlichen Kompetenzen und des sachunterrichtsbezogenen Selbstkonzepts bei Schüler/-innen in Ganztagsgrundschulen: Ergebnisse der Längsschnittstudie StEG-P zu Effekten der

- Schülerteilnahme und der Angebotsqualität. *Zeitschrift für Pädagogik*, 62(6), 760–779.
- Mahoney, J. L., Eccles, J. S., & Larson, R. W. (Hrsg.) (2005). *Organized activities as contexts of development: Extracurricular activities, after-school, and community programs*. Mahwah, NJ: Lawrence Erlbaum.
- Miller, B. M., & Truong, K. A. (2009). The role of afterschool and summer in achievement. The untapped power of afterschool and summer to advance student achievement. *Zeitschrift für Pädagogik*, 55(54. Beiheft), 124–142.
- OECD (2007). *Babies and bosses: Reconciling work and family life: A synthesis of findings for OECD countries* (4. Aufl.). Paris: Organisation for Economic Co-operation and Development.
- OECD (2011a). *Doing better for families*. <http://www.oecd.org/social/family/doingbetter>.
- OECD (2011b). *Quality time for students: Learning in and out of school*. https://www.oecd-ilibrary.org/quality-time-for-students-learning-in-and-out-of-school_5kmdfm0f17ln.pdf?itemId=%2Fcontent%2Fpublication%2F9789264087057-en&mimeType=pdf.
- Plantenga, J., & Remery, C. (2013). *Childcare services for school age children: A comparative review of 33 countries*. Luxembourg: Publications Office of the European Union.
- Radisch, F. (2009). *Qualität und Wirkung ganztägiger Schulorganisation: Theoretische und empirische Befunde*. Studien zur ganztägigen Bildung. Weinheim, München: Juventa.
- Radisch, F., Klemm, K., & Tillmann, K.-J. (2017). Gelingensfaktoren guter Ganztagschulen: eine qualitative Studie bewährter Schulpraxis. In Bertelsmann Stiftung (Hrsg.), *Mehr Schule wagen. Empfehlungen für guten Ganztag* (S. 20–38). Gütersloh.
- Radisch, F., Klieme, E., & Bos, W. (2006). Gestaltungsmerkmale und Effekte ganztägiger Angebote im Grundschulbereich. *Zeitschrift für Erziehungswissenschaft*, 9(1), 30–50.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.
- Scheerens, J. (Hrsg.) (2014). *Effectiveness of time investments in education*. SpringerBriefs in Education. Cham: Springer International Publishing.
- Spillebeen, L., Holtappels, H. G., & Rollett, W. (2011). Schulentwicklungsprozesse an Ganztagschulen: Effekte schulischer Entwicklungsarbeit im Längsschnitt. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche*

- Befunde der Studie zur Entwicklung von Ganztagschulen (StEG)* (S. 120–138). Weinheim, Basel: Beltz Juventa.
- Stecher, L., & Maschke, S. (2013). Research on extended education in Germany – A general model with all-day schooling and private tutoring as two examples. *International Journal for Research on Extended Education*, 1(1), 31–52.
- Stecher, L., Radisch, F., Fischer, N., & Klieme, E. (2007). Bildungsqualität außerunterrichtlicher Angebote in der Ganztagschule. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 27(4), 346–366.
- StEG-Konsortium (2013). *Ganztagschule 2012/2013: Deskriptive Befunde einer bundesweiten Befragung*. http://www.projekt-steg.de/sites/default/files/Bundesbericht_Schulleiterbefragung_2012_13.pdf.
- StEG-Konsortium (2015). *Ganztagschule 2014/2015: Deskriptive Befunde einer bundesweiten Befragung*. http://www.projekt-steg.de/sites/default/files/StEG_Bundesbericht%202015_final_0.pdf.
- StEG-Konsortium (2016). *Ganztagschule: Bildungsqualität und Wirkungen außerunterrichtlicher Angebote: Ergebnisse der Studie zur Entwicklung von Ganztagschulen 2012-2015*. http://www.projekt-steg.de/sites/default/files/StEG_Brosch_FINAL.pdf.
- Steiner, C. (2009). Mehr Chancengleichheit durch die Ganztagschule? *Zeitschrift für Pädagogik*, 55(54. Beiheft), 81–105.
- Steiner, C. (2011a). Ganztagseteiligung und Klassenwiederholung. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG)* (S. 187–206). Weinheim, Basel: Beltz Juventa.
- Steiner, C. (2011b). Teilnahme am Ganztagsbetrieb: Zeitliche Entwicklung und mögliche Selektionseffekte. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG)* (S. 57–75). Weinheim, Basel: Beltz Juventa.
- Steinhauer, H. W., & Zinn, S. (2016). *NEPS technical report for weighting: Weighting the sample of starting cohort 3 of the National Educational Panel Study (waves 1 to 3)*

- (NEPS Working Paper Nr. 63). Bamberg: Leibniz Institute for Educational Trajectories, National Educational Panel Study.
- Strietholt, R., Gustafsson, J.-E., Hogrebe, N., Rolfe, V., Rosén, M., Steinmann, I. et al. (in Druck). The impact of education policies on socioeconomic inequality in student achievement: A review of comparative studies. In L. Volante, S. Schnepf, J. Jerrim & D. Klinger (Hrsg.), *Socioeconomic inequality and student outcomes: National trends, policies, and practices*. New York, NY: Springer.
- Strietholt, R., Manitius, V., Berkemeyer, N., & Bos, W. (2015). Bildung und Bildungsungleichheit an Halb- und Ganztagschulen. *Zeitschrift für Erziehungswissenschaft*, 18(4), 737–761.
- Tillmann, K., Sauerwein, M., Hannemann, J., Decristan, J., Lossen, K., & Holtappels, H. G. (2018). Förderung der Lesekompetenz durch Teilnahme an Ganztagsangeboten? – Ergebnisse der Studie zur Entwicklung von Ganztagschulen (StEG). In M. Schüpbach, L. Frei & W. Nieuwenboom (Hrsg.), *Tagesschulen. Ein Überblick* (S. 289–307). Wiesbaden: Springer Fachmedien.
- Tillmann, K.-J. (2011). Die Steuerung von Ganztagschulen. *Zeitschrift für Erziehungswissenschaft*, 14(S3), 11–24.
- Wendt, H., Goy, M., Walzebug, A., & Valtin, R. (2016). Bildungsangebote an Ganz- und Halbtagsgrundschulen in Deutschland. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 225–246). Münster, New York: Waxmann Verlag.
- Werquin, P. (2010). *Recognising non-formal and informal learning: Outcomes, policies and practices*. Paris: OECD.
- Willems, A. S., Wendt, H., & Radisch, F. (2015). Domänenspezifische Kompetenzen und Chancengerechtigkeit im Vergleich von Ganz- und Halbtagsgrundschultypen: Zur Rolle individueller Herkunftsmerkmale und der Komposition der Schülerschaft. In H. Wendt, T. C. Stubbe, K. Schwippert & W. Bos (Hrsg.), *10 Jahre international vergleichende Schulleistungsforschung in der Grundschule. Vertiefende Analysen zu IGLU und TIMSS 2001 bis 2011* (S. 219–238). Münster, New York: Waxmann.
- Wissenschaftlicher Beirat für Familienfragen (2006). *Ganztagschule: Eine Chance für Familien: Gutachten für das Bundesministerium für Familie, Senioren, Frauen und Jugend*. Wiesbaden: VS Verlag für Sozialwissenschaften.

- Züchner, I. (2011). Ganztagschulen und Familienleben. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG)* (S. 291–311). Weinheim, Basel: Beltz Juventa.
- Züchner, I., & Arnoldt, B. (2011). Schulische und außerschulische Freizeit- und Bildungsaktivitäten: Teilhabe und Wechselwirkungen. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG)* (S. 267–290). Weinheim, Basel: Beltz Juventa.
- Züchner, I., & Fischer, N. (2011). Ganztagschulentwicklung und Ganztagschulforschung: Eine Einleitung. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG)* (S. 9–17). Weinheim, Basel: Beltz Juventa.

4 Anhang

4.1 Veröffentlichungsstatus der Einzelbeiträge

Beitrag I:

Steinmann, I. (2018). Ganztagschulen als neue Organisationsformen – Entwicklungen und Evaluation. *Enzyklopädie Erziehungswissenschaft Online*.
<https://doi.org/10.3262/EEO08180405>

Beitrag II:

Steinmann, I. & Strietholt, R. (in Druck). Effekte nachmittäglicher Bildungsangebote auf sprachliche Kompetenzentwicklungen. *Zeitschrift für Pädagogik*

Der Beitrag wurde am 06.06.2018 zur Veröffentlichung in der Zeitschrift für Pädagogik angenommen und wird in dem zweiten Heft 2019 erscheinen.

Beitrag III:

Steinmann, I. & Strietholt, R. (in Druck). Student achievement and educational inequality in half- and all-day schools: Evidence from Germany. *International Journal for Research on Extended Education*

Der Beitrag wurde am 17.01.2019 zur Veröffentlichung bei dem International Journal for Research on Extended Education angenommen.

Beitrag IV:

Steinmann, I., Strietholt, R., & Caro, D. (2018). Participation in extracurricular activities and student achievement: Evidence from German all-day schools. *School Effectiveness and School Improvement*. <https://doi.org/10.1080/09243453.2018.1540435>

4.2 Eigenanteile bei den Einzelbeiträgen

Im Folgenden werden die Anteile der Doktorandin an den Beiträgen (bereits angenommene sowie weitere Analysen) beschrieben. Hierbei wird auf die Bereiche der wissenschaftlichen Konzeption der Beiträge, der Auswertungen, der Planungen und Durchführungen der schriftlichen Abfassungen sowie auf die Revisionsprozesse eingegangen.

4.2.1 Beitrag I

Steinmann, I. (2018). Ganztagschulen als neue Organisationsformen – Entwicklungen und Evaluation. *Enzyklopädie Erziehungswissenschaft Online*.
<https://doi.org/10.3262/EEO08180405>

- Wissenschaftliche Konzeption der Untersuchung und des Beitrags: Isa Steinmann konzipierte sowohl die Untersuchung als auch den Beitrag.
- Wissenschaftliche Auswertung: Isa Steinmann führte die wissenschaftliche Auswertung durch.
- Konzeption und Durchführung der schriftlichen Abfassung des Beitrags: Isa Steinmann konzipierte die schriftliche Abfassung und führte sie durch.

4.2.2 Beitrag II

Steinmann, I. & Strietholt, R. (in Druck). Effekte nachmittäglicher Bildungsangebote auf sprachliche Kompetenzentwicklungen. *Zeitschrift für Pädagogik*

- Wissenschaftliche Konzeption der Untersuchung und des Beitrags: Isa Steinmann konzipierte die Untersuchung und den Beitrag. Rolf Strietholt beriet hierbei.
- Statistische Auswertung: Isa Steinmann führte die statistische Auswertung durch und wurde hierbei von Rolf Strietholt beraten.
- Konzeption der schriftlichen Abfassung des Beitrags: Isa Steinmann konzipierte die Verschriftlichung des Beitrags hauptverantwortlich, Rolf Strietholt gab hierzu Rückmeldungen.
- Schriftliche Abfassung und Revision nach Rückmeldungen: Der Text des Beitrags wurde von Isa Steinmann verfasst. Sie überarbeitete und finalisierte das Manuskript nach Rückmeldungen von Rolf Strietholt und der Reviews der Zeitschrift.

4.2.3 Beitrag III

Steinmann, I. & Strietholt, R. (in Druck). Student achievement and educational inequality in half- and all-day schools: Evidence from Germany. *International Journal for Research on Extended Education*

- Wissenschaftliche Konzeption der Untersuchung und des Beitrags: Isa Steinmann konzipierte die Untersuchung und den Beitrag. Rolf Strietholt beriet sie hierbei.
- Statistische Auswertung: Isa Steinmann führte die statistische Auswertung durch, während Rolf Strietholt beratend tätig war.
- Konzeption der schriftlichen Abfassung des Beitrags: Die Hauptverantwortung der schriftlichen Abfassung lag bei Isa Steinmann. Rolf Strietholt gab hierzu Rückmeldungen.
- Schriftliche Abfassung und Revision nach Rückmeldungen: Isa Steinmann verfasste den Text des Beitrags und überarbeitete ihn nach Rückmeldungen von Rolf Strietholt und der Reviews der Zeitschrift.

4.2.4 Beitrag IV

Steinmann, I., Strietholt, R., & Caro, D. (2018). Participation in extracurricular activities and student achievement: Evidence from German all-day schools. *School Effectiveness and School Improvement*. <https://doi.org/10.1080/09243453.2018.1540435>

- Wissenschaftliche Konzeption der Untersuchung und des Beitrags: Isa Steinmann konzipierte den Beitrag maßgeblich. Sie wurde hierbei durch Rolf Strietholt beraten.
- Statistische Auswertung: Die statistische Auswertung wurde von Isa Steinmann hauptverantwortlich durchgeführt. Sie wurde hierbei von Rolf Strietholt und Daniel Caro beraten.
- Konzeption der schriftlichen Abfassung des Beitrags: Isa Steinmann konzipierte die schriftliche Abfassung und wurde hierbei von Rolf Strietholt beraten.
- Schriftliche Abfassung und Revision nach Rückmeldungen: Der Beitrag wurde von Isa Steinmann verfasst. Sie überarbeitete den Text nach Rückmeldungen von Rolf Strietholt, Daniel Caro sowie den Reviews der Zeitschrift.

4.3 Eidesstattliche Erklärung

Hiermit versichere ich schriftlich und eidesstattlich gemäß § 11 Abs. 2 PromO v. 08.02.2011:

- (1) Die von mir vorgelegte Dissertation ist selbstständig verfasst und alle in Anspruch genommenen Quellen und Hilfen sind in der Dissertation vermerkt worden.
- (2) Die von mir eingereichte Dissertation ist weder in der gegenwärtigen noch in einer anderen Fassung an der Technischen Universität Dortmund oder an einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegt worden.

Ort, Datum	Unterschrift
------------	--------------

- (3) Weiterhin erkläre ich schriftlich und eidesstattlich, dass mir der „Ratgeber zur Verhinderung von Plagiaten“ und die „Regeln guter wissenschaftlicher Praxis der Technischen Universität Dortmund“ bekannt und von mir in der vorgelegten Dissertation befolgt worden sind.

Ort, Datum	Unterschrift
------------	--------------