

Tomasik, Martin; Helbling, Laura A.; Moser, Urs  
**Wirkungen frühkindlicher Bildung, Betreuung und Erziehung auf  
Lernverläufe in Mathematik und Deutsch während der Pflichtschulzeit**  
*Zeitschrift für Pädagogik 67 (2021) 5, S. 740-759*



**Quellenangabe/ Reference:**

Tomasik, Martin; Helbling, Laura A.; Moser, Urs: Wirkungen frühkindlicher Bildung, Betreuung und Erziehung auf Lernverläufe in Mathematik und Deutsch während der Pflichtschulzeit - In: Zeitschrift für Pädagogik 67 (2021) 5, S. 740-759 - URN: urn:nbn:de:0111-pedocs-287846 - DOI: 10.25656/01:28784; 10.3262/ZP2105740

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-287846>

<https://doi.org/10.25656/01:28784>

in Kooperation mit / in cooperation with:

**BELTZ JUVENTA**

<http://www.juventa.de>

**Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.  
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document.  
This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

**Kontakt / Contact:**

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipt.de](mailto:pedocs@dipt.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

# ZEITSCHRIFT FÜR PÄDAGOGIK

Heft 5

September/Oktober 2021

## ■ *Thementeil*

**Langfristige Auswirkungen von personalen und sozialen Ressourcen sowie von strukturellen Merkmalen des Bildungswesens auf schulische und berufliche Übergänge und Entwicklungsverläufe**

## ■ *Allgemeiner Teil*

**Zusammenarbeit von Lehramtstudierenden im Tandempraktikum – mehr als Austausch und Aufteilung?**

**Bildungsgerechtigkeit als gesellschaftskritische Kategorie. Zur jüngsten Kontroverse in der Erziehungswissenschaft über Leistung und Gerechtigkeit im Bildungswesen**

## Inhaltsverzeichnis

*Thementeil: Langfristige Auswirkungen von personalen und sozialen Ressourcen sowie von strukturellen Merkmalen des Bildungswesens auf schulische und berufliche Übergänge und Entwicklungsverläufe*

*Katharina Maag Merki/Fred Berger/Urs Grob*

Langfristige Auswirkungen von personalen und sozialen Ressourcen sowie von strukturellen Merkmalen des Bildungswesens auf schulische und berufliche Übergänge und Entwicklungsverläufe.

Einführung in den Thementeil ..... 671

*Michael Becker/Jürgen Baumert/Julia Tetzner/Jenny Wagner/*

*Kai Maaz/Olaf Köller*

Zum Zusammenspiel von Selbstwert, sozialer Herkunft und kognitiven Fähigkeiten für die Vorhersage des Bildungs- und Berufserfolgs im Erwachsenenalter .....

682

*Sandra Hupka-Brunner/Thomas Meyer*

Effekte von (bildungs-)institutionellen Rahmenbedingungen und familiären sowie individuellen Ressourcen im Jugendalter auf den weiteren schulischen und beruflichen Lebensverlauf: Befunde aus der TREE-Studie .....

703

*Gwendolin J. Blossfeld/Pia N. Blossfeld/Hans-Peter Blossfeld*

Erwerb der Hochschulzugangsberechtigung über das Gymnasium und den zweiten Bildungsweg: Lesekompetenz im Erwachsenenalter und sozioökonomischer Status beim Berufseinstieg .....

721

*Martin J. Tomasik/Laura A. Helbling/Urs Moser*

Wirkungen frühkindlicher Bildung, Betreuung und Erziehung auf Lernverläufe in Mathematik und Deutsch während der Pflichtschulzeit .....

740

## *Allgemeiner Teil*

*Minh-Ly Do/Tina Hascher*

Zusammenarbeit von Lehramtstudierenden im Tandempraktikum –  
mehr als Austausch und Aufteilung? ..... 760

*Krassimir Stojanov*

Bildungsgerechtigkeit als gesellschaftskritische Kategorie.  
Zur jüngsten Kontroverse in der Erziehungswissenschaft über Leistung  
und Gerechtigkeit im Bildungswesen ..... 784

## *Besprechungen*

*Daniel Erdmann*

Markus Rieger-Ladich/Anne Rohstock/Karin Amos (Hrsg.):  
Erinnern, Umschreiben, Vergessen. Die Stiftung des disziplinären Gedächtnisses  
als soziale Praxis ..... 803

*Laura Burkhardt*

Manfred Berger: Der Kindergarten im Nationalsozialismus.  
„Drum beten wir deutschen Kinder: Den Führer erhalte uns Gott“. Ein Beitrag  
zur Geschichte der öffentlichen Kleinkinder-/Kindergartenpädagogik  
in den Jahren 1933 bis 1945 ..... 806

*Milena Feldmann/Markus Rieger-Ladich*

Carsten Bünger,/Agnieszka Czejkowska (Hrsg.): Political Correctness  
und pädagogische Kritik (Jahrbuch für Pädagogik 2018)  
Nina Degele: Political Correctness – Warum nicht alle alles sagen dürfen ..... 809

*Berno Hoffmann*

Hans-Uwe Otto (Hrsg.): Soziale Arbeit im Kapitalismus.  
Gesellschaftstheoretische Verortungen – Professionspolitische Positionen –  
Politische Herausforderungen  
Hans Thiersch: Lebensweltorientierte Soziale Arbeit – revisited.  
Grundlagen und Perspektiven ..... 812

*Franz-Michael Konrad*

Rainer Bolle/Katja Grundig de Vazquez (Hrsg.): Herbart und  
der Herbartianismus in Jena ..... 816

*Erik Ode*

Simone Austermann/Georg Cleppien/Katharina Vogel (Hrsg.):

Strukturen der Erziehungswissenschaft – Erziehungswissenschaftliche  
Strukturen

Ulrich Binder/Wolfgang Meseth (Hrsg.):

Strukturwandel in der Erziehungswissenschaft. Theoretische Perspektiven

und Befunde ..... 819

*Wilfried Smidt*

Nicole Biedinger (Hrsg.): Was Eltern und Fachkräfte bewegt.

Ein Überblick über die vorschulische Bildung in Deutschland ..... 822

*Dokumentation*

Pädagogische Neuerscheinungen ..... 825

Impressum ..... U3

Table of Contents

*Topic: Long-Term Effects of Personal and Social Resources as well as Structural Characteristics of the Educational System on School and Vocational Transitions and Educational Trajectories*

*Katharina Maag Merki/Fred Berger/Urs Grob*  
Long-Term Effects of Personal and Social Resources as well as Structural Characteristics of the Educational System on School and Vocational Transitions and Educational Trajectories. An Introduction ..... 671

*Michael Becker/Jürgen Baumert/Julia Tetzner/Jenny Wagner/Kai Maaz/Olaf Köller*  
On the Interplay of Self-Esteem, Social Background, and Intelligence in Predicting Educational Attainment and Occupational Success in Adulthood ..... 682

*Sandra Hupka-Brunner/Thomas Meyer*  
How Institutional Opportunities and Individual Resources Affect Adolescents' Educational and Occupational Trajectories: Findings from the Swiss TREE Study ..... 703

*Gwendolin J. Blossfeld/Pia N. Blossfeld/Hans-Peter Blossfeld*  
Higher Education Entry Certificate Obtained via the Traditional Academic Track and Institutions of Second-Chance Education. Adult Reading Literacy and Socioeconomic Status at Labor Market Entry ..... 721

*Martin J. Tomasik/Laura A. Helbling/Urs Moser*  
Effects of Early Childhood Interventions on Learning Trajectories in Mathematics and German during Compulsory Schooling ..... 740

**Articles**

*Minh-Ly Do/Tina Hascher*  
Student Teachers' Cooperation in Paired Field Placement: More Than Just Exchange and Division of Work? ..... 760

*Krassimir Stojanov*  
Educational Justice as a Category of Social Critique. About the Current Educationalist Controversy on Achievement and Justice ..... 784

Book Reviews ..... 803

New Books ..... 825

Impressum ..... U3

# Wirkungen frühkindlicher Bildung, Betreuung und Erziehung auf Lernverläufe in Mathematik und Deutsch während der Pflichtschulzeit

**Zusammenfassung:** Wir untersuchen die Wirkungen frühkindlicher Bildung, Betreuung und Erziehung auf die Lernverläufe während der Schulzeit und differenzieren dabei nach deren Dauer und Intensität. Datengrundlage bilden die Zürcher Lernstandserhebungen ( $N = 1681$ ). Für die Auswertungen bedienen wir uns eines „propensity score“-Verfahrens. Es zeigen sich teilweise signifikante Unterschiede in den Lernfortschritten, vor allem in Deutsch und teilweise auch in Mathematik, wobei eine frühe, aber nicht so intensive Nutzung vorteilhaft ist.

**Schlagworte:** Deutsch, Frühförderung, Kindheit, Lernverlauf, Mathematik

## 1. Einleitung

Dieser Aufsatz soll zur empirischen Evidenz in Bezug auf die langfristigen Wirkungen frühkindlicher Bildung, Betreuung und Erziehung (FBBE) im deutschsprachigen Raum beitragen. Unter FBBE verstehen wir dabei sehr allgemein „all arrangements providing care and education for children under compulsory school age, regardless of setting, funding, opening hours, or programme content“ (OECD, 2001, S. 7), was nicht zuletzt gut zu unserer Datenlage passt, die eine weitergehende Differenzierung auch nicht erlaubt. Wenn man von Wirkungen der FBBE spricht, so lässt sich diese aus verschiedenen Perspektiven untersuchen, die wiederum unterschiedliche Akteure beinhalten. Zunächst lassen sich auf individueller Ebene direkte Wirkungen auf die Bildung und Entwicklung der betreuten Kinder (z. B. Belsky et al., 2007) sowie indirekte Wirkungen, welche durch die Entlastung der Eltern von ihren Betreuungsaufgaben entstehen (z. B. Schober & Schmitt, 2013), unterscheiden. Wenn wir weiter unten den Forschungsstand zu den Wirkungen von FBBE referieren, so werden wir uns ausschließlich auf die direkten Wirkungen auf das Kind konzentrieren. Bereits hier ist die Forschungslage sehr uneinheitlich, selbst wenn man die weit in die Zukunft reichenden Effekte auf die Lebensgestaltung und Lebenszufriedenheit (z. B. Heckman & Masterov, 2007) außer Acht lässt. Neben den individuellen Wirkungen gibt es aber auch eine Vielzahl (zumindest potenzieller) Wirkungen auf volkswirtschaftlicher und gesellschaftspolitischer Ebene. Hier kann man nach Fritschi und Oesch (2009) grob zwischen den (kurz- oder langfristigen) Einkommenseffekten auf die Eltern der betreuten Kinder (z. B. Schneider, Luptáčík & Schmidl, 2006), den (mittelfristigen) Einsparungseffekten auf das Bildungssystem (z. B. Anger, Plünnecke & Tröger, 2007) sowie den (kurz- bis langfristigen) finanzpoli-



tischen Effekten (z.B. Rauschenbach & Schilling, 2007) unterscheiden. Darüber hinaus werden eine Reihe allgemeiner gesellschaftlicher Rückwirkungen (z.B. Dörfler, 2007) und Effekte auf die Chancengleichheit und Bildungsgerechtigkeit (z.B. Keupp, 2011) diskutiert. Außerdem gibt es vermutlich auch einen demographischen Effekt (z.B. Bujard, 2016). Eine zusammenfassende Betrachtung und Modelle zur Wirkungsweise jenseits der individuellen Ebene finden sich in einem Bericht des Schweizerischen Wirtschaftsforschungsinstituts BAK Economics (2020).

Diese vielfältige und überaus interessante Makroperspektive können und wollen wir an dieser Stelle nicht einnehmen. Vor dem Hintergrund der uns zur Verfügung stehenden Daten interessiert uns vielmehr, ob sich direkte Wirkungen der FBBE auf die Bildung und Entwicklung der betreuten Kinder nachweisen lassen, wobei der Fokus auf den Lernverläufen in Mathematik und Deutsch liegt. Eine Stärke des zugrundeliegenden Datensatzes ist es, dass wir die Lernverläufe bis zum Ende der Schulpflicht nachverfolgen können. Des Weiteren wollten wir überprüfen, ob die etwaigen Wirkungen als Funktion der Dauer und der Intensität der FBBE variieren. Einerseits könnte man erwarten, dass die Zusammenhänge in die eine oder andere Richtung im Sinne eines Dosis-Wirkungs-Zusammenhangs mit steigender Dauer und Intensität grösser werden. Andererseits mag es nichtlineare Zusammenhänge geben oder aber die Dauer und die Intensität stehen möglicherweise für unterschiedliche Wirkmechanismen, die sich nicht komplementär zu einer ‚Dosis‘ zusammenfassen lassen.

## 2. Stand der Forschung

Während der Forschungsstand zu den Wirkungen von FBBE in Bezug auf die sozio-emotionale Entwicklung uneinheitlich ist, wobei negative Effekte überwiegen (Belsky, 2001; Jacob, 2009), scheinen die Effekte auf die kognitive und schulische Entwicklung insgesamt positiv zu sein. Diesen Schluss legt eine Reihe von Literaturübersichten und Metaanalysen nahe. Wenn wir im Folgenden die dort zusammengefassten Effektstärken berichten, so beziehen wir uns auf Cohen (1988) und die dort eingeführte Konvention, eine Effektstärke von  $d = .20$  als ‚gering‘, von  $d = .50$  als ‚mittel‘ und von  $d = .80$  als ‚hoch‘ zu bezeichnen.

### 2.1 Forschungsstand in Bezug Wirkungen in der Vorschulzeit

In einer sorgfältig durchgeführten Literaturanalyse, die allerdings ausschließlich Frühförderprogramme in den USA einschließt, gehen Anderson et al. der Frage nach, ob solche Programme „direkt die kognitiven und intellektuellen Leistungen in der frühen Kindheit verbessern“ (2003, S. 34). Berücksichtigt wurden dabei nur solche Programme, die sich an drei- bis fünfjährige Kinder aus armutsbetroffenen Familien richteten. Sie finden unter anderem positive Zusammenhänge mit der Schulreife ( $d = .38$ ), der Intelligenz ( $d = .43$ ) oder den sozialen Kompetenzen der Kinder ( $d = .38$ ).

Höhere Effektstärken liefern Nelson, Westhues und MacLeod (2003), und zwar für kognitive Ergebnisvariablen ( $d = .52$ ) und für Maße des familiären Wohlbefindens ( $d = .33$ ) während der Vorschulzeit. In ihrer Metaanalyse haben sie sowohl universelle als auch indizierte Programme aus dem englischsprachigen Raum aufgenommen. Bemerkenswert sind dabei einige ausgewählte Befunde zu bestimmten Programmmerkmalen. So waren Programme, die länger als ein Jahr dauerten ( $d = .53$ ) bzw. mehr als 300 Einheiten umfassten ( $d = .57$ ) weitaus wirksamer waren als kürzere Programme ( $d = .09$ ) bzw. solche mit wenigen Einheiten ( $d = .25$ ). Gleiche Befunde sind bei von Nores und Barnett (2010) nachzulesen. Damit stehen sowohl die Länge als auch die Intensität eines Programms in einem positiven Zusammenhang mit dessen Wirksamkeit. Dagegen fanden die Autoren keinen moderierenden Effekt des Alters des Kindes bei Programmbeginn.

Die Untersuchung von Nores und Barnett (2010) unterscheidet sich von den vorher genannten durch zwei Merkmale. Erstens werden nicht nur FBBE-Programme untersucht, sondern alle Formen von pädagogischen und sozialpolitischen Interventionen einschließlich Einkommens- und Ernährungsergänzungsmaßnahmen (wie etwa Familienzulagen oder vergünstigtes Schulesen). Diese breite Perspektive ist durchaus relevant, auch wenn man sich lediglich für den Einfluss von FBBE interessiert, liefert sie doch einen Vergleichsmaßstab („benchmark“), um die Effektstärken in einem größeren Zusammenhang zu sehen. Der zweite Unterschied ist, dass eine globale Perspektive eingenommen wird und ausschließlich Programme und Studien betrachtet werden, die *nicht* in den USA angesiedelt sind. Dabei werden insgesamt 30 Programme in 23 Ländern im Hinblick auf kognitive Zugewinne, Verhaltensänderungen, Gesundheitseffekte sowie das Ausmaß der Beschulung untersucht. Die Wirkungen auf diese drei Ergebnisvariablen sind durchaus vergleichbar (Effektstärken mit  $.27 < d < .31$ ), allerdings zeigen sich deutlichere Unterschiede, wenn man nach dem Typ der Intervention differenziert. Unterschieden wurde nach (a) Interventionen, die den Familien eine Einkommensergänzung zur Verfügung stellen, (b) Interventionen, die auf die Verbesserung der Ernährungssituation abzielen und (c) sogenannte gemischte Interventionen, die zumindest auch eine pädagogische Komponente beinhalten. Für den kognitiven Bereich zeigten sich Interventionen mit einer pädagogischen Komponente wirksamer ( $d = .35$ ) als solche, die nur auf die Ernährung ( $d = .26$ ) oder nur auf das Einkommen ( $d = .17$ ) abzielen. In Bezug auf Verhaltensänderungen sind ernährungsbezogene Interventionen ( $d = .46$ ) dagegen deutlich wirksamer als pädagogische ( $d = .27$ ) oder einkommensbezogene ( $d = .17$ ). Und auch die gesundheitlichen Effekte sind bei ernährungsbezogenen Interventionen ( $d = .38$ ) stärker als bei pädagogischen ( $d = .23$ ) oder einkommensbezogenen ( $d = .21$ ).

Zusammenfassend relativieren sich also die Effektstärken der pädagogischen Interventionen, insbesondere wenn man bedenkt, dass es sich bei diesen um die mit Abstand aufwändigsten handelt. Dieser Befund deckt sich aber auch mit einer neueren Metanalyse aus den Niederlanden (Fukkink, Jilink & Oostdam, 2017). Diese hat sich mit der Frage beschäftigt, ob sich FBBE-Programme mit besonderem pädagogischem Anspruch, häufig gerichtet an spezielle Zielgruppen, von regulären Programmen unter-

scheiden, die zumeist universell sind. Die Autoren finden dann auch keinen Mehrwert von pädagogischen Programmen im Hinblick auf kognitive und sozio-emotionale Ergebnisvariablen.

Dass der Inhalt eines FBBE-Programms keine Rolle spielen soll, verwundert zunächst. Aufklärung schafft hier eine von der OECD (2018) publizierte Metaanalyse, die verschiedene Merkmale der Struktur- und Prozessqualität (vgl. Pianta et al., 2005) als Moderatorvariablen berücksichtigt hat. Strukturelle Merkmale – etwa das Betreuungsverhältnis oder die Raumgröße – scheinen im Einklang mit früheren Untersuchungen (etwa Melhuish et al., 2015; Pianta et al., 2005) nur eine indirekte Rolle zu spielen, und zwar wenn sie sich wiederum positiv auf Prozessmerkmale auswirken. Eben diese Prozessmerkmale – etwa die Qualität der Interaktion zwischen Erzieherinnen und Kindern oder die Art der pädagogischen Aktivitäten – sind aber ausschlaggebend für die Wirksamkeit im kognitiven und sozio-emotionalen Bereich (Melhuish et al., 2015). Die indirekte Wirkung struktureller Merkmale kann zumindest zum Teil dafür verantwortlich sein, dass sich die Wirksamkeit der verschiedenen Arten von Förderprogrammen, die ja in den früheren Untersuchungen an strukturellen Merkmalen festgemacht worden sind, nicht sonderlich unterscheidet.

## *2.2 Forschungsstand in Bezug auf Wirkungen in der Schulzeit*

Anderson et al. (2003) finden eine Reihe positiver Wirkungen von FBBE-Programmen während der Schulzeit. Zum einen sind die Schulleistungen um  $d = .35$  höher, die Wahrscheinlichkeit einer sonderschulischen Maßnahme sinkt um 14 Prozentpunkte und die Zahl der Schulabschlüsse steigt um 17 Prozentpunkte. Zum andern sinkt die Delinquenz um  $d = .60$ , die Zahl der Verhaftungen während des Jugendalters um 20 Prozentpunkte und die Zahl der Teenagerschwangerschaften um 49 Prozentpunkte.

Nelson et al. (2003) finden vergleichbare Effekte bis zur achten Klassenstufe. Ähnlich wie schon für die vorschulischen Ergebnisvariablen werden diese Effektstärken durch die Dauer und die Intensität des Programms positiv moderiert. Zudem finden sich die stärksten Effekte für die am meisten benachteiligten Kinder.

Eine weitere, sehr umfangreiche Analyse liefern Mitchel, Wylie und Carr (2008). Diese unterscheidet zwischen Interventionsmaßnahmen (d. h. meistens gerichtet an eine spezielle Risikogruppe) und allgemeinbildenden Maßnahmen (d. h. meistens universelle Programme mit alltäglicher Betreuung) sowie zwischen kurzfristigen (d. h. meist bezogen auf die Grundschulzeit) und langfristigen (d. h. meistens bezogen auf das Alter von 14 oder 16 Jahren) Wirkungen. Für die Mathematikleistungen finden sich kurzfristig Effekte von  $.29 < d < .45$  bei Interventionsmaßnahmen und von  $.18 < d < .43$  bei Allgemeinmaßnahmen. Langfristig werden Effekte von Interventionsmaßnahmen in Höhe von  $.19 < d < .44$  berichtet. Bemerkenswert ist auch, dass die Effekte sowohl mit der Dauer als auch mit der Intensität der Maßnahme steigen. Für die Leseleistungen finden sich noch stärkere Effekte, wobei hier nur Zahlen aus Interventionsmaßnahmen berichtet wurden. Kurzfristig gibt es in diesem Bereich Effekte von  $.34 < d < .89$  und langfris-

tig von  $.17 < d < .44$ . Diese Effekte sind tendenziell stärker für eine längere Programmdauer, aber der Vorteil verliert sich auf lange Sicht. Hinsichtlich der Intensität waren die Befunde gemischt. In einer ähnlichen Größenordnung bewegen sich die Effektstärken in Bezug auf die sozio-emotionalen Ergebnisvariablen während der Schulzeit. Diese Zahlen sind übrigens konsistent mit denen von Nores und Barnett (2010), die auch gezielt nach der Art der Intervention unterschieden haben. Dort findet sich für schulische Ergebnisse ein Effekt von  $d = .38$  durch einkommensbezogene Maßnahmen, von  $d = .11$  durch ernährungsbezogene Maßnahmen und von  $d = .41$  durch pädagogische Maßnahmen. Nicht ganz so optimistisch stimmen dagegen die Ergebnisse der Metanalyse von Camili, Vargas, Ryan und Barnett (2010), die für schulbezogene Ergebnisvariablen lediglich Effektstärken von  $d = .14$  in isolierten Evaluationen und von  $d = .06$  in vergleichenden Evaluationen finden.

Sehr umfangreich im Hinblick auf Wirkungen in der Schulzeit und einschlägig im Hinblick auf die Berücksichtigung ausschließlich institutioneller FBBE-Programme ist die systematische Literaturübersicht von Burger (2010), in der ein asiatisches, acht europäische sowie elf nordamerikanische Programme berücksichtigt worden sind. Insgesamt zeigen sich positive und zuweilen beträchtliche Programmeffekte. Die größten Effektstärken finden sich für das Kennen von Buchstaben und Wörtern ( $d = .79$ ), von Büchern ( $d = .67$ ), von Farben ( $d = .60$ ) sowie für das frühe Zahlenverständnis ( $d = .47$ ).

Zusammenfassend lässt sich sagen, dass sich positiven Programmeffekte frühkindlicher Förderung bis in die Schulzeit finden lassen, wobei die Effekte für kognitive und sozio-emotionale Ergebnisvariablen ähnlich stark sind. Je nach Programminhalt, Zielgruppe und Studiendesign fallen diese Effekte mal höher, mal niedriger aus, so dass ein übertriebener Optimismus nicht angebracht ist. Zugleich findet man, dass die Persistenz der Effekte nicht unter allen Umständen gegeben ist. Vielmehr zeigt sich manchmal eine Domänen- und Geschlechtsspezifität (Hill, Gormley & Adelstein, 2015). Auch finden Camili et al. (2010), dass die Effekte auf allgemeine kognitive Variablen nachlassen, je mehr Zeit vergeht, was für die sozialen und schulleistungsbezogenen Ergebnisvariablen aber so nicht gilt. Vor dem Hintergrund dieser Befundlage argumentieren Melhuish et al. (2015), dass die Wirkungspfade von den verschiedenen Frühförderprogrammen auf die schulische Entwicklung der Kinder komplexer sind, als man das gemeinhin annehmen könnte.

### 3. Fragestellung

Die vorliegende Untersuchung möchte einen Beitrag zur Beantwortung der Frage leisten, wie unterschiedliche Nutzungsmuster von FBBE mit der Entwicklung von Schulleistungen in Mathematik und Deutsch zusammenhängen. Wir konzentrieren uns auf schulleistungsbezogene Ergebnisvariablen, weil hier die stärksten Effekte zu erwarten sind (Camili et al., 2010; Nelson et al., 2003) und erwarten für diese Ergebnisvariablen einen neutralen bis positiven, keinesfalls aber einen negativen Effekt. Gleichzeitig ist aus der Literatur bekannt, dass die Dauer und die Intensität einer frühpädagogischen

Maßnahme deren Wirksamkeit moderiert (Anderson et al., 2003; Mitchel et al., 2008). Aus diesem Grund möchten wir prüfen, ob sich dieser moderierende Effekt auch für FBBE-Programme im deutschsprachigen Raum finden lässt. Allerdings möchten wir uns hier nicht auf eine gerichtete Hypothese festlegen, weil uns keine Angaben zum Inhalt und zur Qualität der hier genutzten FBBE-Angebote haben, kulturspezifische Besonderheiten existieren könnten und die Dauer mit der Intensität in einer komplexen Wechselwirkung zueinander stehen könnten (Loeb, Bridges, Bassok, Fuller & Rumberger, 2007).

## 4. Methode

### 4.1 Grundgesamtheit, Stichprobe und Studiendesign

Die Grundgesamtheit der hier verwendeten Zürcher Lernstandserhebungen bildeten alle 11 118 Schülerinnen und Schüler, die im Sommer 2003 im Kanton Zürich in eine der 650 neu gebildeten Regelklassen eingeschult worden sind. In den 120 zufällig ausgewählten Klassen wurden sämtliche 2 043 Schülerinnen und Schüler zur Teilnahme eingeladen. Über die vier Erhebungswellen nahmen dann im Schuljahr 2003/04 wenige Wochen nach der Einschulung  $N_1 = 1\,970$ , im Schuljahr 2005/06 drei Jahre nach der Einschulung  $N_2 = 1\,935$ , im Schuljahr 2008/09 sechs Jahre nach der Einschulung  $N_3 = 1\,679$  und im Schuljahr 2011/12 neun Jahre nach der Einschulung  $N_4 = 1\,635$  Schülerinnen und Schüler teil. Die Reduktion in der Stichprobengröße war in erster Linie auf die geographische Mobilität (Wegzug aus dem Kanton Zürich) zurückzuführen und es zeigte sich nur ein geringes Ausmaß an Strichprobenselektivität (vgl. Tomasik, Oostlander & Moser, 2018). Zu allen Messzeitpunkten wurden die Teilnehmerinnen und Teilnehmer individuell nachverfolgt und unabhängig von Klassenstufe, besuchtem Schultyp usw. untersucht.

### 4.2 Operationalisierung der untersuchten Variablen

Wenige Wochen nach der Einschulung wurde ein Fragebogen ausgeteilt, in dem die Eltern Angaben dazu machen sollten, ob, wie lange und in welchem Umfang ihr Kind eine Kinderkrippe oder eine vergleichbare Einrichtung besucht hat. Die Elternfragebögen wurden dabei in verschiedene Sprachen übersetzt. Der Rücklauf belief sich auf 82 Prozent. In Tabelle 1 findet sich eine Zusammenstellung aller sich ergebenden Kombinationen nach Intensität und Dauer. Bei der Intensität wurde unterschieden nach „1–3 Tage“ und „4 und mehr Tage“, bei der Dauer nach „für 1 Jahr“, „für 2 Jahre“ und „für 3 oder mehr Jahre“. Weil Kinder in den seltensten Fällen aus einer Einrichtung wieder austreten, wenn sie sie einmal besucht haben, korrespondiert eine kurze Dauer mit einem späten Eintritt in die FBBE. Eine lange Dauer dagegen impliziert gleichzeitig einen frühen Eintritt in die FBBE. Da die Zellenbesetzung zu klein war, um weitere Analysen zu

Intensität	Dauer				Total
	nie	1 Jahr	2 Jahre	3+ Jahre	
nie	1234 <sup>a</sup> (73 %)	–	–	–	1234 (73 %)
1–3 Tage/Woche	–	84 <sup>b</sup> (5 %)	91 <sup>b</sup> (5 %)	127 <sup>c</sup> (8 %)	302 (18 %)
4–5 Tage/Woche	–	22 <sup>d</sup> (1 %)	44 <sup>d</sup> (3 %)	79 <sup>d</sup> (5 %)	145 (9 %)
Total	1234 (73 %)	106 (6 %)	135 (8 %)	206 (12 %)	1681 (100 %)

Anmerkungen: <sup>a</sup> zusammengefasst zu „nie“; <sup>b</sup> zusammengefasst zu „spät“; <sup>c</sup> zusammengefasst zu „früh“; <sup>d</sup> zusammengefasst zu „intensiv“

Tab. 1: Verteilung des Krippenbesuches nach Dauer und Intensität

rechnen, wurden die einzelnen Kategorien zu vier Gruppen zusammengefasst. Die erste Gruppe umfasste die 1 234 Kinder, welche die FBBE „nie“ besucht haben. Die „späte“ Gruppe umfasste die 175 Kinder, welche die FBBE an bis zu drei Tagen die Woche während einer relativ kurzen Dauer von einem oder zwei Jahren besucht haben. Die „frühe“ Gruppe umfasste die 127 Kinder, welche die FBBE an bis zu drei Tagen besuchten, dieses aber für drei oder mehr Jahre taten. Die „intensive“ Gruppe schließlich umfasste alle 145 Kinder, welche die FBBE an vier oder mehr Tagen die Woche besuchten, und zwar unabhängig davon, wie lange sie es taten. Die Wahl dieser Gruppenbildung war vor allem dadurch motiviert, möglichst gleich große Gruppen der Kinder, die die FBBE überhaupt besucht haben, zu erhalten. Alternative Gruppenbildungen lieferten aber praktisch gleiche Ergebnisse, wenn man einmal von den Signifikanzen absieht, die ja auch eine Funktion der Gruppengröße sind. Hinweise auf die soziodemographische Zusammensetzung der Gruppen finden sich in den jeweils oberen Zeilen von Tabelle 2. Insgesamt fällt auf, dass das „frühe“ Nutzungsmuster eher mit Indikatoren einer privilegierten Herkunft einhergeht, während das „intensive“ Nutzungsmuster eher bei den benachteiligten Kindern anzutreffen ist.

Der *Verlauf der Schulleistungen* in Deutsch und Mathematik wurde drei, sechs und neun Jahre nach der Einschulung erfasst. Auf Grundlage des Lehrplans und der Lehrmittel des Kantons Zürich wurde dafür mit Unterstützung durch qualifizierte Lehrpersonen jeweils ein Deutschtest und ein Mathematiktest entwickelt. Die Ergebnisse der Schülerinnen und Schüler in allen Leistungstests wurden nach dem Rasch-Modell skaliert und dann recht aufwändig mit Hilfe von zusätzlichen Stichproben in den Klassenstufen zwischen den eigentlichen Erhebungen auf einer einheitlichen Skala abgebildet. Für die anschließenden Analysen modellierten wir die schulischen Leistungen in Deutsch und Mathematik in jeweils einem latenten Wachstumskurvenmodell und bestimmten somit jeweils einen Anfangswert (,intercept‘) und eine Steigung (,slope‘). Eine Korrelation zwischen Anfangswert und Steigung wurde nicht zugelassen, (1) weil sie mit

	nie	spät	früh	intensiv	$\eta^2$
Alter	6.96	6.92	6.84	6.89	.009
	6.94	6.95	6.91	6.96	.002
Geschlecht	50.0	50.9	57.5	49.7	.002
	50.3	51.0	64.9	54.7	.011
Fremdsprachigkeit	25.0	21.5	11.0	41.4	.020
	25.7	21.5	19.2	38.1	.023
Schweizerdeutsch	81.2	75.9	89.1	67.8	.021
	80.2	80.2	84.2	76.5	.006
Schweizer Mutter	60.6	47.9	68.5	41.5	.019
	58.7	56.1	65.1	44.1	.018
Schweizer Vater	61.3	50.2	76.0	33.6	.035
	60.2	50.9	60.7	48.1	.012
Geschwister	1.39	1.18	1.01	0.90	.034
	1.33	1.22	1.13	1.18	.007
Haushaltsgrösse	4.29	3.94	3.82	3.60	.052
	4.22	4.04	4.02	4.00	.009
Zimmer pro Person	1.17	1.26	1.34	1.17	.019
	1.19	1.21	1.27	1.17	.007
Eigenes Zimmer	74.6	74.4	71.4	75.9	.000
	74.4	74.1	78.3	74.7	.001
Bildung der Eltern	12.4	13.7	14.7	12.3	.040
	12.6	13.2	13.7	12.8	.015
Reg. Arbeitslosigkeit	4.59	4.55	4.51	4.95	.007
	4.61	4.65	4.64	4.62	.000
Reg. Ausländerquote	22.8	23.2	23.3	27.4	.022
	23.2	23.4	23.1	24.8	.006
Reg. Eigenheimquote	19.9	17.1	13.0	8.8	.061
	18.4	17.7	16.5	13.3	.021
Reg. Sesshaftigkeit	55.5	54.7	55.5	55.0	.003
	55.4	55.6	55.6	54.6	.009

Tab. 2: Zusammensetzung der Nutzungsgruppen vor (jeweils obere Zeile) und nach (jeweils untere Zeile) der Gewichtung



$r = -.27$  ( $p = .35$ ) für Deutsch und mit  $r = .00$  ( $p = 1.00$ ) für Mathematik sowieso nicht signifikant gewesen wäre, (2) um den für die Schätzung der Modellanpassungsgüte notwendigen einen Freiheitsgrad zu gewinnen und (3) um das Modell letztlich so sparsam wie möglich zu parametrisieren. Um eine nichtlineare Steigerung der schulischen Leistung zu ermöglichen, haben wir die jeweils dritte Ladung der Steigungskomponente freigegeben. Sowohl für Deutsch ( $\Delta\chi^2 = 288.36$ ,  $df = 1$ ,  $p < .001$ ) als auch für Mathematik ( $\Delta\chi^2 = 3789.26$ ,  $df = 1$ ,  $p < .001$ ) passte ein lineares Modell nämlich signifikant schlechter als nicht-lineares, so dass wir in beiden Fällen jeweils das nicht-lineare Modell verwendeten. Die nicht-standardisierten Ladungen auf die Steigungsfaktoren waren  $\lambda_3 = .00$ ,  $\lambda_6 = 1.00$  und  $\lambda_9 = 1.37$  für Deutsch und  $\lambda_3 = .00$ ,  $\lambda_6 = 1.00$  und  $\lambda_9 = 1.14$  für die Mathematik. Die Passung der so parametrisierten nicht-linearen Modelle zu den Daten war dabei sowohl für Deutsch ( $\chi^2 = 3.58$ ,  $df = 1$ ,  $p = .06$ ; RMSEA = .039; CFI = .995; TLI = .986) als auch für Mathematik ( $\chi^2 = 0.01$ ,  $df = 1$ ,  $p = .94$ ; RMSEA = .000; CFI = 1.000; TLI = 1.000) als gut zu bezeichnen (auch wenn man bei der Interpretation dieser Koeffizienten berücksichtigen sollte, dass hier ein beinahe saturiertes Modell auf seine Passung zu den Daten hin überprüft worden ist).

Angaben zum *Alter* und zu *Geschlecht* der Schülerinnen und Schüler wurden aus der offiziellen „Statistik der Lernenden“ der Bildungsdirektion des Kantons Zürich entnommen. Aus dem bereits erwähnten Elternfragebogen wurde entnommen, ob zu Hause eine *Fremdsprache* gesprochen wurde, welche *Nationalität* und welchen höchsten *Bildungsabschluss* der Vater und die Mutter des Kindes hatte. Außerdem wurde die Anzahl der *Geschwister* sowie die Zahl der Personen im Haushalt und die Anzahl der Zimmer in der Wohnung bestimmt sowie erfragt, ob das Kind ein *eigenes Zimmer* hat oder nicht. Aus der Zahl der Personen im Haushalt und der Anzahl der Zimmer in der Wohnung wurden die *Zimmer pro Person* als Indikator für den sozio-ökonomischen Status bestimmt. Alle diese Variablen wurden ausgewählt, weil wir gute Gründe für die Annahme haben konnten, dass diese zumindest größtenteils vor dem Besuch einer FBBE-Einrichtung feststanden bzw. sich seit dem Besuch einer FBBE-Einrichtung nicht verändert hatten. Neben den Angaben auf der Ebene der Schülerinnen und Schüler sowie ihrer Familien wurden drei Indikatoren auf der Ebene des Schulbezirks in den Analysen berücksichtigt, nämlich die *regionale Arbeitslosenquote* (Anteil der Erwerbspersonen an allen Erwerbspersonen), die *regionale Ausländerquote* (Anteil der Einwohner ohne Schweizer Staatsbürgerschaft an allen Einwohnern), die *regionale Eigenheimquote* (Anteil der Wohnungen in Einfamilienhäusern an allen Wohnungen) sowie die *regionale Sesshaftigkeit* (Index für die Häufigkeit von Umzügen außerhalb der Grenzen des Schulbezirks).

### 4.3 Statistisches Verfahren

Bei fehlender Randomisierung zielen „propensity score“-Verfahren darauf ab, möglichst ein experimentelles Setting nachzustellen (Austin & Stuart, 2015; Gangl & di Prete, 2004). Der „propensity score“ ist definiert als die konditionale Wahrscheinlich-



keit, mit der eine Person aufgrund ihrer Merkmale ein bestimmtes Treatment erfährt. Diese Wahrscheinlichkeit wird auf der Basis von beobachteten Hintergrundmerkmalen geschätzt, welche einen potenziell verzerrenden Effekt haben könnten. Es wird versucht, möglichst einen Ausgleich auf potenziell konfundierenden Hintergrundmerkmalen zwischen Personengruppen, die unterschiedliche Formen des Treatments erfahren haben, zu erzielen.

Für die folgenden Analysen wurde für die Schätzung der „propensity scores“ auf „generalized boosted models“ (GBMs) aus dem Bereich des maschinellen Lernens zurückgegriffen (vgl. McCaffrey, Ridgeway & Morral, 2004). Bei diesen Modellen wird keine funktionale Form des Zusammenhangs vorgegeben und alle möglichen Interaktionen zwischen den Kovariaten zugelassen. GBMs bieten also viel Flexibilität bei der Schätzung der „propensity scores“ und funktionieren auch wenn mehrere Treatments verglichen werden sollen (McCaffrey et al., 2013). Basierend auf den geschätzten „propensity scores“ für die vier unterschiedlichen Nutzungsformen der FBBE haben wir nach dem Verfahren des „inverse probability treatment weighting“ (IPTW; Austin & Stuart, 2015; McCaffrey et al., 2013) Gewichte berechnet, die es erlauben, Unterschiede in den ausgewählten Merkmalen nach Angebotsnutzung auszugleichen. Wir haben dafür das Statistikpaket „twang“ (Ridgeway, McCaffrey, Morral, Griffin & Burgette, 2017) verwendet.

Tabelle 2 gibt einen Überblick über das erreichte ‚balancing‘ zwischen den vier Nutzungsgruppen. Dabei wird deutlich, dass viele Gruppenunterschiede nach der Gewichtung kleiner werden oder sogar ganz verschwinden. Die größten verbleibenden Gruppenunterschiede betreffen die Merkmale Fremdsprachigkeit und regionale Eigenheimquote, hier gelingt ein ‚balancing‘ nicht ganz. Diese Gewichte wurden in allen Analysen einbezogen, um die Unterschiede zwischen den Nutzergruppen zu korrigieren. Zusätzlich wurden verschiedene Kontrollvariablen in die Modelle aufgenommen („doubly robust“-Ansatz; McCaffrey et al., 2013). Wir berichten dabei immer sogenannte „average treatment effects“ (ATEs; Gangl & di Prete, 2004; McCaffrey et al., 2013). ATEs entsprechen dem durchschnittlichen Effekt eines bestimmten Nutzungsmusters im Vergleich zu einem anderen Nutzungsmuster für eine zufällig gezogene Person aus der Gesamtpopulation.

## 5. Ergebnisse

Für die Schulleistungen in Deutsch finden sich in Tabelle 3 zunächst die Ergebnisse für das Ausgangsniveau, gemessen drei Jahre nach der Einschulung. Die meisten signifikanten Effekte verlieren sich nach Kontrolle der Kovariaten. Es zeigt sich dann aber, dass ein „intensives“ Nutzungsmuster mit niedrigeren Schulleistungen in Deutsch einhergeht als ein „Frühes“. Das Ausgangsniveau ist außerdem höher für Mädchen im Vergleich zu Jungen. Mehr Zimmer pro Person im Haushalt korreliert mit besseren Deutschleistungen, während ein eigenes Zimmer mit schlechteren Deutschleistungen einhergeht.

	Modell ohne Kovariaten			Modell mit Kovariaten		
	B	SE	p	B	SE	p
Regressionsmodell						
Konstante	505.01			505.84		
Kinderkrippe „spät“ vs. „nie“	11.93	12.17		3.73	13.07	
Kinderkrippe „früh“ vs. „nie“	42.44	14.22	***	23.57	14.61	
Kinderkrippe „intensiv“ vs. „nie“	-16.99	10.48		-14.20	11.66	
Alter				-6.92	13.64	
Geschlecht				30.84	9.59	***
Fremdsprache				-22.16	15.44	
Schweizer Mutter				6.89	11.80	
Schweizer Vater				0.94	14.12	
Geschwister				-0.28	5.81	
Zimmer pro Person				10.90	18.00	
Eigenes Zimmer				-14.34	13.43	**
Bildung der Eltern				9.39	1.66	***
Reg. Ausländerquote				0.34	0.94	
Reg. Eigenheimquote				0.56	0.55	
Reg. Sesshaftigkeit				1.15	0.90	
Vollständiger Kontrastvergleich						
Kinderkrippe „spät“ vs. „nie“	11.93	12.17		3.73	13.07	
Kinderkrippe „früh“ vs. „nie“	42.44	14.22	***	23.57	14.61	
Kinderkrippe „intensiv“ vs. „nie“	-16.99	10.48		-14.20	11.66	
Kinderkrippe „früh“ vs. „spät“	30.51	18.04	†	19.84	18.84	
Kinderkrippe „intensiv“ vs. „spät“	-28.92	15.18	†	-17.93	16.28	
Kinderkrippe „intensiv“ vs. „früh“	-59.43	16.63	***	-37.77	17.36	*

Anmerkungen:  $R^2 = .052$  (ohne Kovariaten) bzw.  $R^2 = .285$  (mit Kovariaten)  
\*\*\*  $p < .001$ ; \*\*  $p < .01$ ; \*  $p < .05$ ; †  $p < .10$

Tab. 3: Gruppenunterschiede in den Schulleistungen in Deutsch (Ausgangsniveau in Klasse 3 im Wachstumskurvenmodell)

	Modell ohne Kovariaten			Modell mit Kovariaten		
	<i>B</i>	<i>SE</i>	<i>p</i>	<i>B</i>	<i>SE</i>	<i>p</i>
<b>Regressionsmodell</b>						
Konstante	277.02			276.30		
Kinderkrippe „spät“ vs. „nie“	-5.34	6.66		-3.98	7.13	
Kinderkrippe „früh“ vs. „nie“	14.07	8.19	†	15.64	8.44	†
Kinderkrippe „intensiv“ vs. „nie“	-7.84	10.94		-7.89	11.61	
Alter				-16.53	9.16	†
Geschlecht				-14.60	6.21	*
Fremdsprache				5.04	8.73	
Schweizer Mutter				-5.18	7.31	
Schweizer Vater				9.81	7.10	
Geschwister				-4.35	3.66	
Zimmer pro Person				-6.27	9.61	
Eigenes Zimmer				8.02	7.80	
Bildung der Eltern				0.03	0.88	
Reg. Ausländerquote				-0.98	0.60	†
Reg. Eigenheimquote				-0.75	0.36	*
Reg. Sesshaftigkeit				1.34	0.93	
<b>Vollständiger Kontrastvergleich</b>						
Kinderkrippe „spät“ vs. „nie“	-5.34	6.66		-3.98	7.13	
Kinderkrippe „früh“ vs. „nie“	14.07	8.19	†	15.64	8.44	†
Kinderkrippe „intensiv“ vs. „nie“	-7.84	10.94		-7.89	11.61	
Kinderkrippe „früh“ vs. „spät“	19.41	10.07	†	19.62	10.48	†
Kinderkrippe „intensiv“ vs. „spät“	-2.50	12.54		-3.92	13.51	
Kinderkrippe „intensiv“ vs. „früh“	-21.91	13.09	†	-23.54	14.13	†

Anmerkungen:  $R^2 = .090$  (ohne Kovariaten) bzw.  $R^2 = .356$  (mit Kovariaten)

\*\*\*  $p < .001$ ; \*\*  $p < .01$ ; \*  $p < .05$ ; †  $p < .10$

Tab. 4: Gruppenunterschiede in den Schulleistungen in Deutsch (Lernfortschritt im Wachstums-kurvenmodell)

	Modell ohne Kovariaten			Modell mit Kovariaten		
	B	SE	p	B	SE	p
Regressionsmodell						
Konstante	495.67			497.08		
Kinderkrippe „spät“ vs. „nie“	6.77	12.71		4.87	14.42	
Kinderkrippe „früh“ vs. „nie“	13.15	14.70		10.73	15.76	
Kinderkrippe „intensiv“ vs. „nie“	-24.48	12.73	†	-13.83	13.93	
Alter				5.84	15.51	
Geschlecht				-32.37	10.30	**
Fremdsprache				-10.26	15.45	
Schweizer Mutter				-3.27	12.70	
Schweizer Vater				26.20	14.23	†
Geschwister				0.02	5.14	
Zimmer pro Person				33.48	18.39	†
Eigenes Zimmer				-7.38	13.50	
Bildung der Eltern				7.35	1.64	***
Reg. Ausländerquote				1.16	0.79	
Reg. Eigenheimquote				0.93	0.62	
Reg. Sesshaftigkeit				-0.18	1.07	
Vollständiger Kontrastvergleich						
Kinderkrippe „spät“ vs. „nie“	6.77	12.71		4.87	14.42	
Kinderkrippe „früh“ vs. „nie“	13.15	14.70		10.73	15.76	
Kinderkrippe „intensiv“ vs. „nie“	-24.48	12.73	†	-13.83	13.93	
Kinderkrippe „früh“ vs. „spät“	6.38	18.79		5.86	20.01	
Kinderkrippe „intensiv“ vs. „spät“	-31.25	17.31	†	-18.70	18.36	
Kinderkrippe „intensiv“ vs. „früh“	-37.63	18.87	*	-24.56	19.69	

Anmerkungen:  $R^2 = .023$  (ohne Kovariaten) bzw.  $R^2 = .252$  (mit Kovariaten)

\*\*\*  $p < .001$ ; \*\*  $p < .01$ ; \*  $p < .05$ ; †  $p < .10$

Tab. 5: Gruppenunterschiede in den Schulleistungen in Mathematik (Ausgangsniveau in Klasse 3 im Wachstumskurvenmodell)

	Modell ohne Kovariaten			Modell mit Kovariaten		
	<i>B</i>	<i>SE</i>	<i>p</i>	<i>B</i>	<i>SE</i>	<i>p</i>
<b>Regressionsmodell</b>						
Konstante	383.01			384.14		
Kinderkrippe „spät“ vs. „nie“	-6.26	8.21		-9.28	8.86	
Kinderkrippe „früh“ vs. „nie“	22.41	10.67	*	16.68	11.23	
Kinderkrippe „intensiv“ vs. „nie“	-19.21	20.20		-27.13	21.32	
Alter				-35.16	14.01	*
Geschlecht				5.94	9.64	
Fremdsprache				3.64	12.82	
Schweizer Mutter				-17.99	10.77	†
Schweizer Vater				0.43	11.43	
Geschwister				-6.49	4.81	
Zimmer pro Person				-18.99	14.81	
Eigenes Zimmer				-9.67	9.51	
Bildung der Eltern				2.59	1.34	†
Reg. Ausländerquote				-0.94	0.67	
Reg. Eigenheimquote				-0.62	0.42	
Reg. Sesshaftigkeit				1.42	1.79	
<b>Vollständiger Kontrastvergleich</b>						
Kinderkrippe „spät“ vs. „nie“	-6.26	8.21		-9.28	8.86	
Kinderkrippe „früh“ vs. „nie“	22.41	10.67	*	16.68	11.23	
Kinderkrippe „intensiv“ vs. „nie“	-19.21	20.20		-27.13	21.32	
Kinderkrippe „früh“ vs. „spät“	28.67	12.80	*	25.96	13.73	
Kinderkrippe „intensiv“ vs. „spät“	-12.95	21.61		-17.86	23.04	
Kinderkrippe „intensiv“ vs. „früh“	-41.62	22.03	†	-43.82	24.64	†

Anmerkungen:  $R^2 = .091$  (ohne Kovariaten) bzw.  $R^2 = .348$  (mit Kovariaten)

\*\*\*  $p < .001$ ; \*\*  $p < .01$ ; \*  $p < .05$ ; †  $p < .10$

Tab. 6: Gruppenunterschiede in den Schulleistungen in Mathematik (Lernfortschritt im Wachstumskurvenmodell)

Für den Lernzuwachs in Deutsch, dargestellt in Tabelle 4, zeigt sich ein steilerer Anstieg bei „früher“ Nutzung im Vergleich zu der Gruppe, welche die Krippe „nie“, „spät“ oder „intensiv“ besucht hat, allerdings nur mit einer Irrtumswahrscheinlichkeit von 10 Prozent. Über die FBBE-Nutzung hinaus sagt nur noch das Geschlecht sowie die regionalen Eigenheimquote den Anstieg der Deutschleistungen signifikant vorher.

Wie aus Tabelle 5 ersichtlich, zeigen sich für die Schulleistungen in Mathematik keine signifikanten Gruppenunterschiede im Ausgangsniveau, wenn man die Kovariaten statistisch kontrolliert. Zwar scheint zunächst und je nach Vergleichsgruppe ein „intensives“ Nutzungsmuster mit einem niedrigeren Ausgangsniveau zu korrelieren, aber dieser Effekt verschwindet, wenn man die Kovariaten kontrolliert. Ein höheres Ausgangsniveau in Mathematik zeigt sich außerdem eher bei Jungen im Vergleich zu Mädchen sowie bei Kindern, die gebildete Eltern haben.

Dagegen zeigt sich im Lernzuwachs, dargestellt in Tabelle 6, dass eine „frühe“ FBBE-Nutzung mit steileren Fortschritten einhergeht als eine FBBE-Nutzung „nie“ oder „spät“. Allerdings verlieren sich diese Effekte, wenn man die Kovariaten kontrolliert. Übrig bleibt eine Tendenz, dass das „intensive“ Nutzungsmuster gegenüber dem „frühen“ nachteilig ist.

## 6. Diskussion

Die folgende Diskussion gliedert sich in drei Teile. Nach einer Zusammenfassung der wichtigsten Befunde vor dem Hintergrund des Standes der Forschung werden wir die von uns verwendete statistische Methode kritisch diskutieren, um anschließend Implikationen für Forschung und Anwendung zu ziehen sowie einen Ausblick zu wagen.

Wenn man den Forschungsstand gesamthaft betrachtet, so sind Effekte von FBBE am ehesten auf die kognitiven, schulnahen Variablen zu erwarten (Camili et al., 2010; Nelson et al., 2003). Das war für uns der wichtigste Grund gewesen, eben solche Variablen in den Mittelpunkt unserer eigenen Untersuchung zu stellen. Dennoch sind unsere Befunde eher gemischt. Das Ausgangsniveau in Deutsch scheint für Kinder mit einem „frühen“ Nutzungsmuster zunächst deutlich höher zu sein, der Effekt verliert sich aber nach statistischer Kontrolle der Kovariaten. Dagegen zeigt sich für diese Kinder ein steilerer Lernfortschritt sowohl ohne als auch mit Kontrolle der Kovariaten. Allerdings ist dieser lediglich marginal signifikant. Auch gibt es für Mathematik signifikante oder knapp signifikante Effekte auf das Ausgangsniveau und den Lernfortschritt, aber diese verlieren sich fast vollständig, wenn man die verschiedenen Kovariaten (und hier insbesondere die Bildung der Eltern) statistisch kontrolliert. Wir können also schlussfolgern, dass am ehesten die sprachliche Entwicklung der Kinder mit FBBE zusammenhängt. Dabei zeigen sich am ehesten für das „frühe“ Nutzungsmuster vorteilhafte Lernverläufe, während das „intensive“ Nutzungsmuster am ehesten mit nachteilhaften Lernverläufen einhergeht.

Für eine angemessene Interpretation dieser Befunde muss man sich die Frage stellen, wie viel an diesen Ergebnissen tatsächlich ein Effekt der Nutzung von FBBE ist und wie

viel der unterschiedlichen Gruppenzusammensetzung geschuldet ist. Das ist ein prinzipielles Problem bei der kausalen Interpretation nicht-experimenteller Daten (Gangl & di Prete, 2004) und an dieser Stelle deswegen besonders wichtig, weil die Nutzungsgruppe mit den vorteilhaftesten Entwicklungsergebnissen, also der „frühe“ Typ, gleichzeitig auch diejenige ist, in der die Eltern einen eher privilegierten sozioökonomischen Status aufweisen und die Fremdsprachigkeit (und die damit häufiger in Verbindung stehenden schulischen Probleme) relativ selten ist. Gleichzeitig ist die Nutzungsgruppe mit den unvorteilhaftesten Entwicklungsergebnissen, also der „intensive“ Typ, auch eine, in der die Eltern niedrige Bildungsabschlüsse haben und die Zweisprachigkeit besonders häufig ist. Trotz sorgfältiger Bemühung, ein ‚balancing‘ zwischen den vier Nutzungsgruppen zu erreichen, haben sich diese am Ende immer noch geringfügig unterschieden und zwar in eine Richtung, welche mit den Unterschieden in den Entwicklungsergebnissen plausibel vereinbar ist. Das spricht für starke und offenbar schichtabhängige Selbstselektionseffekte in der Angebotsnutzung, die wir mit den vorhandenen Daten nicht herausrechnen können. Das an sich ist vielleicht auch ein zentrales Ergebnis dieser Untersuchung, das wir zwar nicht als Hypothese formuliert haben, das aber deutlich macht, in welchem historischen Kontext diese Daten erhoben wurden und interpretiert werden müssen. Natürlich kann es auch sein, dass am Ende unsere Gruppen nicht groß genug waren, um eine ausreichende Überlappung auf den vielen Kovariaten zu ermöglichen. Hinzu kommt, dass die retrospektive Erfassung des Nutzungsmusters die Möglichkeit zur Kontrolle zeitvariabler Kovariaten deutlich einschränkt.

Bedeutsamer ist jedoch, dass wir nicht wissen, ob es nicht weitere Unterschiede auf nicht beobachteten Variablen gibt, auch wenn wir eine sehr breite Auswahl an Variablen auf unterschiedlichen Untersuchungsebenen verwendet haben. Diese nicht beobachteten Variablen lassen sich nämlich auch nicht durch die Berücksichtigung von Kovariaten ausgleichen, was Rosenbaum (1995) ausführlich diskutiert. Somit steht und fällt die Interpretation mit der Frage, ob es plausibel ist anzunehmen, dass es solche Variablen gibt und dass diese für die heterogenen Effekte verantwortlich sind. Diese Frage lässt sich nur spekulativ beantworten. Unseres Erachtens haben wir die soziodemographischen und kontextuellen Variablen sehr ausführlich erfasst, sodass durch weitere Indikatoren kaum mehr Effekte auf die Entwicklungsergebnisse nachgewiesen werden. Wir haben Merkmale des Kindes und des Elternhauses berücksichtigt, den kulturellen Hintergrund der Familie, Migrationserfahrung sowie die Wohn- und Familienverhältnisse. Außerdem haben wir mit den Regionalindikatoren ein recht genaues Bild über das soziale Umfeld, die vorherrschenden sozialen Normen und auch die ökonomischen Merkmale der Region.

Was wir jedoch gar nicht berücksichtigt haben sind verhaltensnahe, psychologische Variablen, sowohl auf der Seite der Eltern als auch auf der Seite der Kinder, oder sogar genetische Faktoren, welche die Nutzungseffekte moderieren (Belsky & Pluess, 2013) oder aber ein bestimmtes Nutzungsmuster überhaupt erst verursacht haben könnten (Kiff, Lengua & Zalewski, 2011). Und hier ist es wahrscheinlich anzunehmen, dass diese sowohl mit der Nutzung von FBBE als auch mit verschiedenen Entwicklungsergebnissen einhergehen. So kann es beispielsweise durchaus sein, dass Eltern mit tem-

peramentvollen oder gar verhaltensauffälligen Kindern sich eher dafür entscheiden, ihr Kind „intensiv“ in eine Kindertagesstätte zu schicken, weil sie sich vielleicht zu Hause überfordert fühlen mit ihrer Erziehung. Gleichzeitig werden solche Kinder später eher Schwierigkeiten in der Schule haben, was mit dem Besuch der Kindertagesstätte vielleicht gar nichts zu tun hat. Im Endeffekt aber würde man einen Zusammenhang zwischen „intensivem“ FBBE-Nutzungsmuster und schlechteren Schulleistungen finden. In diesem Sinne ist eine Reihe von Merkmalen der Eltern, des Erziehungsverhaltens und der Kinder denkbar, die hier als ‚Drittvariablen‘ relevant sein könnten. Diese ließen sich aber im Prinzip nur über ein experimentelles Design (also die zwangsweise Zuweisung der Kinder in bestimmte Formen der FBBE) untersuchen, was nicht nur ethisch unvertretbar wäre, sondern auch die externe Validität der Ergebnisse in Zweifel ziehen würde (Melhuish, 2004).

Diese beiden Einschränkungen vorausgeschickt kann man aber aufgrund des Datensatzes und den methodischen Maßnahmen zur Herstellung der Äquivalenz der vier Vergleichsgruppen sowie des Einbezugs relevanter Kovariaten mit einer gewissen Vorsicht gültige Schlüsse aus den Befunden ziehen. Zugleich muss man bei der Interpretation berücksichtigen, dass wir wegen des langen Längsschnitts immer von einer Nutzung von FBBE um das Jahr 2000, also vor etwa 20 Jahren, sprechen. Ob sich die pädagogische Qualität der damaligen Angebote mit den heutigen vergleichen lässt, muss an dieser Stelle offenbleiben, zumal in unseren Daten sowieso keine Informationen zu Qualitätsmerkmalen (wie etwa zur Ausbildung der Betreuungspersonen, zur Gruppengröße oder zu Prozessmerkmalen wie Interaktion und Kommunikation in der Einrichtung) vorliegen. Dass aber die pädagogische Qualität einen zentralen und nicht zu ignorierenden Moderator darstellt, ist unbestritten (OECD, 2018).

Die vorliegende Untersuchung hat auch trotz der eben genannten methodischen Einschränkungen Implikationen für Forschung und Anwendung. Für die Forschung ergibt sich die Konsequenz, dass allgemeine Aussagen darüber, ob die Nutzung von FBBE generell vorteilhaft oder nachteilhaft sei, nur einen bedingten Nutzen haben. Vielmehr kommt es auf die Nutzungsdauer und die Nutzungsintensität an, die beide in Kombination betrachtet werden müssen, weil sich die Effekte gegenseitig aufheben können (Loeb et al., 2007). Eine Interpretation von Studien, welche diese beiden Aspekte nicht berücksichtigen, ist vor dem Hintergrund der vorliegenden Daten nicht möglich.

Aus der Anwendungsperspektive stellt sich die in dieser Untersuchung gar nicht angegangene Frage nach einer etwaigen Heterogenität der Effekte dar. Es ist wahrscheinlich, dass FBBE nicht für alle Kinder gleichermaßen wirksam ist. So konnte beispielsweise Becker (2010) zeigen, dass die Wirkung der Frühförderung auf den Wortschatz gerade bei den Kindern besonders groß ist, die mit einem geringen Wortschatz in die Frühförderung eintreten. Obwohl solche Effekte hier nicht untersucht wurden, so liegt doch die Vermutung nahe, dass sie eine wichtige moderierende Rolle spielen könnten.



## Literatur

- Anderson, L. M., Shinn, C., Fullilove, M. T., Scrimshaw, S. C., Fielding, J. E., Normand, J., Carande-Kulis, V. G., & Task Force on Community Preventive Services (2003). The effectiveness of early childhood development programs: A systematic review. *American Journal of Preventive Medicine*, 24, 32–46.
- Anger, C., Plünnecke, A., & Tröger, M. (2007). *Renditen der Bildung: Investitionen in den frühkindlichen Bereich*. Köln: Institut der deutschen Wirtschaft.
- Austin, P. C., & Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine*, 34, 3661–3679.
- BAK Economics (Hrsg.) (2020). *Volkswirtschaftliches Gesamtmodell für die Analyse zur „Politik der frühen Kindheit“: Bericht im Auftrag der Jacobs Foundation*. Basel: BAK Economics.
- Becker, B. (2010). Wer profitiert mehr vom Kindergarten? Die Wirkung der Kindergartenbesuchsdauer und Ausstattungsqualität auf die Entwicklung des deutschen Wortschatzes bei deutschen und türkischen Kindern. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62, 139–163.
- Belsky, J. (2001). Developmental risks (still) associated with early child care. *Journal of Child Psychology and Psychiatry*, 42, 845–859.
- Belsky, J., & Pluess, M. (2015). Genetic moderation of early child-care effects on social functioning across childhood: A developmental analysis. *Child Development*, 84, 1209–1225.
- Belsky, J., Vandell, D. L., Burchinal, M., Clarke-Stewart, K. A., McCartney, K., Owen, M. T., & NICHD Early Child Care Research Network (2007). Are there long-term effects of early child care? *Child Development*, 78, 681–801.
- Bujard, M. (2016). Wirkungen von Familienpolitik auf die Geburtenentwicklung. In Y. Niephaus, M. Kreyenfeld & R. Sackmann (Hrsg.), *Handbuch Bevölkerungssoziologie* (S. 619–646). Wiesbaden: Springer.
- Burger, K. (2010). How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. *Early Childhood Research Quarterly*, 25, 140–165.
- Camili, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*, 112, 579–620.
- Cohen, J. (1998). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Routledge.
- Dörfler, S. (2007). *Kinderbetreuungskulturen in Europa: Ein Vergleich vorschulische Kinderbetreuung in Österreich, Deutschland, Frankreich und Schweden* (Working Paper Nr. 57/2007). Wien: Österreichisches Institut für Familienforschung.
- Fritschi, T., & Oesch, T. (2009). *Volkswirtschaftlicher Nutzen von frühkindlicher Bildung in Deutschland: Eine ökonomische Bewertung langfristiger Bildungseffekte bei Krippenkindern*. Gütersloh: Bertelsmann Stiftung.
- Fukkink, R., Jilink, L., & Oostdam, R. (2017). A meta-analysis of the impact of early childhood interventions on the development of children in the Netherlands: An inconvenient truth? *European Early Childhood Education Research Journal*, 25, 656–666.
- Gangl, M., & di Prete, T. A. (2004). Kausalanalyse durch Matchingverfahren. In A. Diekmann (Hrsg.), *Methoden der Sozialforschung* (S. 396–420). Opladen: Westdeutscher Verlag.
- Heckman, J. J., & Masterov, D. V. (2007). The productivity argument for investing in young children. *Applied Economic Perspectives and Policy*, 29, 446–493.
- Hill, C. J., Gormley, W. T., Jr., & Adelstein, S. (2015). Do the short-term effects of a high-quality preschool program persist? *Early Childhood Research Quarterly*, 32, 60–79.

- Jacob, J.I. (2009). The socio-emotional effects of non-maternal childcare on children in the USA: A critical review of recent studies. *Early Child Development and Care*, 179, 559–570.
- Keupp, H. (2011). Verwirklichungschancen von Anfang an: Frühe Förderung als Beitrag zur Befähigungsgerechtigkeit. In R. Günther, K. Pfeifer & T. Drössler (Hrsg.), *Aufwachsen in Dialog und sozialer Verantwortung: Bildung – Risiken – Prävention in der frühen Kindheit* (S. 49–70). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kiff, C. J., Lengua, L. J., & Zalewski, M. (2011). Nature and nurturing: Parenting in the context of child temperament. *Clinical Child and Family Psychology Review*, 14, 251–301.
- Loeb, S., Bridges, M., Bassok, D., Fuller, B., & Rumberger, R.W. (2007). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economic of Education Review*, 26, 52–66.
- McCaffrey, D., Ridgeway, G., & Morral, A. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, 9, 403–425.
- McCaffrey, D.F., Griffin, B.A., Almirall, D., Slaughter, M.E., Ramchand, R., & Burgette, L. (2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in Medicine*, 32, 3388–3414.
- Melhuish, E. C. (2004). *A literature review of the impact of early years provision on young children, with emphasis given to children from disadvantages backgrounds*. London: National Audit Office.
- Melhuish, E., Ereky-Stevens, K., Petrogiannis, K., Ariescu, A., Penderi, E., Rentzou, K., Tawell, A., Slot, P., Broekhuizen, M., & Leseman, P. (2015). *A review of research on the effects of early education and childcare (ECEC) upon child development*. Brussels: European Commission.
- Mitchell, L., Wylie, C., & Carr, M. (2008). *Outcomes of early childhood education: Literature review*. Wellington: Ministry of Education.
- Nelson, G., Westhues, A., & MacLeod, J. (2003). A meta-analysis of longitudinal research on pre-school prevention programs for children. *Prevention and Treatment*, 6, 31.
- Nores, M., & W.S. Barnett (2010). Benefits of early childhood interventions across the world: (Under) investing in the very young. *Economics of Education Review*, 29, 271–282.
- OECD (2001). *Starting strong: Early childhood education and care*. Paris: OECD.
- OECD (2018). *Engaging young children: Lessons from research about quality in early childhood education and care*. Paris: OECD.
- Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, 9, 144–159.
- Rauschenbach, T., & Schilling, M. (2007). *Erwartbare ökonomische Effekte durch den Ausbau der Betreuungsangebote für unter Dreijährige auf 750 000 Plätze bis 2013*. München: Deutsches Jugendinstitut.
- Ridgeway, G., McCaffrey, D., Morral, A., Griffin, B.A., & Burgette, L. (2017). *Package „twang“: Toolkit for weighting and analysis of nonequivalent groups* [Software]. <https://cran.r-project.org/web/packages/twang/twang.pdf> [21.06.2021].
- Rosenbaum, P.R. (1995). *Observational studies*. New York: Springer.
- Schneider, U., Luptáćik, M., & Schmidl, B. (2006). *Volkswirtschaftliche Effekte außerhäuslicher Kinderbetreuung: Eine Input-Output-Analyse der Produktions-, Wertschöpfungs- und Beschäftigungseffekte der Wiener Kindertageseinrichtungen* (Forschungsbericht 01/2006). Wien: Institut für Sozialpolitik.

Schober, P.S., & Schmitt, C. (2013). *Ausbau der Kinderbetreuung: Entlastung der Eltern erhöht Zufriedenheit* (DIW Wochenbericht Nr. 50/2013). Berlin: Deutsches Institut für Wirtschaftsforschung.

Tomasik, M.J., Oostlander, J., & Moser, U. (2018). *Von der Schule in den Beruf: Wege und Umwege in der nachobligatorischen Ausbildung*. Zürich: Bildungsdirektion Zürich.

**Abstract:** We examine the effects of early childhood education and care on learning trajectories across compulsory school, differentiating between duration and intensity. The data are based on the Zurich Learning Progress Study and comprise a sample of  $N = 1681$  students. We use a propensity score procedure for evaluating the effects. We find close to significant differences in learning progress, especially in German and in part also in mathematics, whereby early but not intensive childcare turns out most advantageous.

**Keywords:** Childhood, German, Early Education, Learning Trajectory, Mathematics

#### **Anschrift der Autor\_innen**

Prof. Dr. Martin J. Tomasik, Institut für Bildungsevaluation,  
Wilfriedstrasse 15, 8032 Zürich, Schweiz  
E-Mail: martin.tomasik@ibe.uzh.ch

Dr. Laura A. Helbling, Institut für Bildungsevaluation,  
Wilfriedstrasse 15, 8032 Zürich, Schweiz  
E-Mail: laura.helbling@ibe.uzh.ch

Prof. Dr. Urs Moser, Institut für Bildungsevaluation,  
Wilfriedstrasse 15, 8032 Zürich, Schweiz  
E-Mail: urs.moser@ibe.uzh.ch