

Komponenten und Auswirkungen der Qualität mathematischer Bildung in
frühkindlichen Bildungs- und Betreuungseinrichtungen

Dissertation

zur Erlangung des akademischen Grades

Doktorin der Philosophie (Dr. phil.)

Am Fachbereich Erziehungswissenschaft und Psychologie
der Freien Universität Berlin

vorgelegt von

Hannah Ulferts

Berlin, 2017

Erstgutachterin: Prof. Dr. Yvonne Anders

Zweitgutachterin: Prof. Dr. Bettina Hannover

Tag der Disputation: 15.02.2017

Danksagung

Herzlich bedanken möchte ich mich bei Prof. Dr. Yvonne Anders für die Betreuung der Dissertationsschrift, für die vielen anregenden Diskussionen sowie die Möglichkeit, Einblicke in spannende Forschungsarbeiten im Rahmen von Projekten und Tagungen zu erhalten.

Prof. Dr. Bettina Hannover danke ich für die Bereitschaft als Zweitgutachterin zu fungieren und die Möglichkeit, mir in ihrem Kolloquium wertvolle Rückmeldungen und anregende Diskussionen zum Zwischenstand der Arbeit einzuholen.

Weiteren Dank möchte ich an meine Kollegen/innen und die studentischen Hilfskräfte richten, die mich auf unterschiedliche Weise bei den Arbeiten im Rahmen des Promotionsvorhabens unterstützt haben. Dank gilt an dieser Stelle insbesondere Prof. Dr. Axinja Hachfeld für die intensive Betreuung in der Anfangsphase sowie Ana Susac, Nadine Wieduwilt und Dr. Franziska Wilke für die unglaublich tolle Unterstützung im weiteren Verlauf.

Ein ganz besonderer Dank richtet sich an meine Eltern, die mir den Weg zur Dissertation erst geebnet haben und mir ständig unterstützend zur Seite stehen. Schließlich möchte ich mich bei meinen Geschwistern und Freunden für die vielen motivierenden Gespräche bedanken.

Inhaltsverzeichnis

Zusammenfassung	1
1 Einleitung.....	7
2 Grundlagen früher Lern- und Anregungsprozesse	10
2.1 Bedeutung der frühkindlichen Phase und der spezifischen Anregung in Mathematik ...	10
2.2 Ökosystemische Sicht auf frühe Lern- und Anregungsprozesse	11
2.3 Entwicklungsrelevante Merkmale der Familien.....	12
3 Frühkindliche Einrichtungen als Bildungsinstitutionen der Mathematik.....	14
3.1 Entstehungsgeschichte frühkindlicher mathematischer Bildung.....	14
3.2 Politische Verankerung durch Bildungs- und Rahmenlehrpläne.....	15
3.3 Erwartungen an die mathematische Bildungsarbeit in der FIBB	16
3.4 Die hierarchische Struktur der FIBB und ihre Bedeutung	17
3.5 Debatte um universelle vs. zielgruppenspezifische Bildungsangebote in Europa und den USA	18
4 Wissenschaftliche Auseinandersetzung mit mathematischer Bildung in der FIBB	20
4.1 Fachdidaktische Auseinandersetzung mit mathematischer Bildung	20
4.2 Forschungsstrang zu mathematischen Förderprogrammen	21
4.3 Forschungsstrang zu mathematischen Bildungsoffensiven bzw. -initiativen	22
4.4 Forschungsstrang zur mathematischen Bildung in der FIBB-Regelversorgung	25
5 Komponenten der Qualität mathematischer Bildung.....	28
5.1 Komponenten etablierter Qualitätsmodelle.....	28
5.2 Bildungseffektivität als Ergänzung für etablierte Qualitätsmodelle.....	32
5.3 Ausdifferenzierung fachkraftbezogener Qualitätskomponenten in Mathematik	34
6 Evidenz zu Qualitätskomponenten und Auswirkungen in der FIBB.....	37
6.1 Befundlage zur Prozessqualität und ihrer Auswirkungen auf Bildungsergebnisse.....	37
6.2 Evidenz zur mathematischen Bildungseffektivität frühkindlicher Einrichtungen.....	40
6.3 Evidenz zu fachkraftbezogenen Qualitätskomponenten	41
7 Anliegen des Promotionsvorhabens	43
7.1 Die zugrundeliegenden Desiderata	43
7.2 Die Ziele des Dissertationsvorhabens und der Teilprojekte (TP).....	45
8 Auswirkungen der Prozessqualität FIBB auf Bildungsergebnisse in Mathematik und Sprache/Literacy – Eine Metaanalyse europäischer Längsschnittstudien (TP1).....	50
8.1 Erkenntnisse und Limitierungen bisheriger Forschungssynthesen	50
8.2 Herleitung der Moderatoren der Prozesseffekte	51
8.3 Beschreibung der Metaanalyse zu Prozesseffekten des TP1	53

8.4 Methode der Metaanalyse des TP1.....	54
8.5 Ergebnisse der Metaanalyse zu den Prozesseffekten des TP1	56
8.6 Diskussion zu Auswirkungen der Prozessqualität des TP1.....	60
9 Effektivität FIBB in Deutschland – Diskussion und empirische Untersuchung des Ansatzes im Bildungsbereich Mathematik (TP2)	62
9.1 Herausforderungen des Qualitätsansatzes	62
9.2 Herausforderungen des Effektivitätsansatzes	65
9.3 Anforderungen von Effektivitätsanalysen an Studiendesigns.....	67
9.4 Beschreibung der Effektivitätsanalysen für die BiKS-Einrichtungen des TP2	68
9.5 Methode der mathematischen Effektivitätsanalysen des TP2	68
9.6 Ergebnisse zur mathematischen Effektivität FIBB des TP2	70
9.7 Diskussion zur mathematischen Effektivität der BiKS-Einrichtungen des TP2	72
10 Zentrale fachkraftbezogene Komponenten der Struktur- und Orientierungsqualität in der FIBB in Deutschland und ihr Zusammenhang (TP3)	74
10.1 Strukturkomponenten – Qualifikation und Professionalisierung in Mathematik.....	74
10.2 Orientierungskomponenten – Motivationale Orientierungen.....	75
10.3 Zusammenhang zwischen den Struktur- und Orientierungskomponenten	78
10.4 Beschreibung der Studie zu fachkraftbezogenen Qualitätskomponenten des TP3	80
10.5 Methode der Studie zu den fachkraftbezogenen Qualitätskomponenten des TP3.....	81
10.6 Ergebnisse zu den fachkraftbezogenen Qualitätskomponenten des TP3.....	82
10.7 Diskussion zu fachkraftbezogenen Qualitätskomponenten des TP3.....	86
11 Gesamtdiskussion.....	89
11.1 Diskussion der Teilprojekte des Dissertationsvorhabens.....	89
11.2 Implikationen für die Forschung zur Qualität mathematischer Bildung in der FIBB....	93
11.3 Implikationen für das zukünftige Bündeln von Evidenz unterschiedlicher Studien ...	98
11.4 Bildungspolitische und praktische Implikationen für die FIBB-Regelversorgung.....	99
Literaturverzeichnis.....	104
Anhang A – Manuskript zur Metaanalyse (TP1)	137
Anhang B - Zusätzliches Online-Material zur Metaanalyse (TP1).....	165
Anhang C – Manuskript zur mathematischen Bildungseffektivität (TP2).....	178
Anhang C – Manuskript zur Struktur- und Orientierungsqualität (TP3).....	207
Anhang E - Lebenslauf.....	236
Anhang F – Erklärung.....	240

Zusammenfassung

Komponenten der Qualität spielen eine besondere Rolle, um die mathematischen Bildungspotenziale der Regelversorgung frühkindlicher institutioneller Bildung und Betreuung (FIBB) zu erklären. In etablierten Modellen wird eine Unterteilung der Komponenten nach drei Dimensionen vorgenommen: Struktur-, Orientierungs- und Prozessqualität, welche sich weiter in globale und bereichsspezifische Prozessqualität aufgliedert. Im Promotionsvorhaben wird sich in *drei Teilprojekten (TP)* theoretisch und empirisch mit den Komponenten und Auswirkungen der Qualität mathematischer Bildung in der FIBB in Europa insbesondere Deutschland auseinandergesetzt. Ziel ist zum einen zur Weiterentwicklung von Qualitätsmodellen und -forschung anzuregen zum anderen den Erkenntnisgewinn aufzuzeigen, der aus einem Bündeln der Evidenz unterschiedlicher Studien mit methodisch-statistischer Vielfalt resultieren kann.

Das *TP1* aggregiert mit einer Längsschnittmehrebenenmetaanalyse die Evidenz zur Auswirkung der Prozessqualität aus 17 Längsschnittstudien und neun Ländern Europas für 2 887 Einrichtungen und 16 461 Kinder. Es resultierten persistente Effekte globaler und bereichsspezifischer Prozessqualität ($ES = .11$ und $.10$, Cohens $d = .22$ und $.20$) für Bildungsergebnisse in Mathematik und Sprache/Literacy. Mixed-Effects-Moderatoranalysen zeigen, dass die Effekte bereichsspezifisch ausfallen. Zudem erklären das gewählte Prozessmaß und das statistisch-analytische Vorgehen variierende Befunde.

Im *TP2* wird der Ansatz der Bildungseffektivität für die FIBB in Deutschland diskutiert, um den mathematischen Bildungsoutput von Einrichtungen zu beurteilen. Ferner werden Ergebnisse verschiedener Random-Intercepts-Modelle zur mathematischen Effektivität der 97 deutschen Einrichtungen der Studie BiKS-3-10 vorgestellt, welche auf der Leistung der 554 Kinder im K-ABC Rechnen basieren. Es ergaben sich keine Effektivitätsunterschiede in fairen Vergleichen, welche die Kompetenzen zu einem früheren Messzeitpunkt oder entwicklungsrelevante Merkmale der Kinder berücksichtigten. Obschon das kindliche Kompetenzniveau in den Einrichtungen von Beginn bis zum Ende des Betreuungszeitraums voneinander abwich, scheinen sich die Unterschiede durch eine andere soziokulturelle Zusammensetzung in den Einrichtungen zu erklären. Die Gegenüberstellung von Pro- und Kontraargumenten spricht für einen komplementären Einsatz von Analysen zur Prozessqualität und Effektivität in der FIBB, auch weil die Befunde beider Ansätze in BiKS voneinander abwichen.

Qualitätskomponenten, die sich auf die Fachkräfte beziehen, stehen aktuell im Fokus von Debatten, sind jedoch unscharf konzeptualisiert und unterforscht. Das *TP3* zeigt daher exemplarisch auf, wie fachkraftbezogene Komponenten der Struktur- und Orientierungsqualität ausdifferenziert und erforscht werden können. In explorativen Analysen einer Fragebogenerhebung wurden Schwächen in der mathematikspezifischen Strukturqualität der untersuchten 29 deutschen Regeleinrichtungen aufgedeckt, weil die 221 deutschen Fachkräften nur von wenig besuchten Aus- und Fortbildungsangeboten in Mathematik berichteten. Mehrheitlich scheinen die formalen Erfahrungen mit Mathematik auf Schulerfahrungen begrenzt, die Fachkräfte weder als besonders positiv noch negativ erinnerten. Ein positives Bild ergab sich für fachkraftbezogene Orientierungskomponenten mit durchschnittlich relativ hoher Ausprägung der mathematischen Selbstwirksamkeit und Wertorientierung der Fachkräfte. Eine Ausnahme bildete das vergleichsweise niedrige mathematische Selbstkonzept. Die Regressionsanalysen zu den Zusammenhängen zwischen Struktur- und Orientierungskomponenten zeigten, dass die Orientierungen (v. a. das Selbstkonzept) mit positiven und negativen Schulerfahrungen zusammenhängen. Zusätzlich zeigte sich ein Zusammenhang zwischen Selbstwirksamkeit und einem höheren Abschluss sowie mehr besuchten Ausbildungsangeboten in Mathematik. Eine längere Berufserfahrung stand in positiver Beziehung zur Selbstwirksamkeit und Wertorientierung der Fachkräfte.

In der Zusammenschau betont die Diskussion, dass in Zukunft Kompetenz- und Prozessmaße der Mathematik zu ergänzen und ökonomische Fragebogenmaße der Prozessqualität zu entwickeln sind. Zudem wird auf das Erkenntnispotenzial weiterer Effektivitätsanalysen mit verbesserten Studiendesigns und vergleichender Analysen zur Qualität und Effektivität FIBB in anderen Entwicklungsbereichen hingewiesen. Viele ungeklärte Fragen werfen die Rolle der materiell-räumlichen Umgebung für die Bildungsqualität und die Wirkweise von Struktur- und Orientierungskomponenten auf. Generell besteht großer Forschungsbedarf zu den fachkraftbezogenen Komponenten für die FIBB-Regelversorgung in Anbetracht der abweichenden Befunde aus Fachkräftetrainings- und Initiativen sowie für US-Fachkräfte. Die Evidenzbündelung ermöglichte es nur teilweise Limitierungen einzelner Studien zu überwinden, wobei zukünftig aufgrund der wachsenden Verfügbarkeit von Daten und elaborierteren Verfahren komplexere Fragestellungen untersucht werden können. Die Befunde implizieren, dass bildungspolitische Maßnahmen und Investitionen zur Steigerung der Prozessqualität beispielsweise eine intensiviertere Professionalisierung der Fachkräfte in Mathematik das Bildungspotenzial FIBB maximieren können. Als Steuerungsparameter für die flächendeckende Implementation mathematischer Bildung außerhalb von Programmen und Offensiven bzw. Initiati-

ven wird ein kontinuierliches Monitoring der Prozessqualität (mit interaktionsfokussierten Beobachtungsverfahren) sowie ggfs. des Bildungoutputs diskutiert. Zudem gilt es die soziokulturelle Durchmischung in FIBB voranzutreiben und pädagogische Konzepte für den Umgang mit Diversität zu erarbeiten.

Schlüsselwörter: Frühkindliche Bildung und Betreuung, Mathematik, Qualität, Bildungseffektivität

Abstract

Quality components play a key role in explaining the educational impact of regular provision of centre-based Early Childhood Education and Care (ECEC). The components fall into three main categories: structural, orientation, and process quality, which is further subdivided into global and domain-specific. *Three studies (TP for Teilprojekte)* comprising this dissertation theoretically and empirically examine the quality components and impact of math education in centre-based ECEC in Europe, and especially in Germany. The studies aim, firstly, at providing incentive for the development of quality concepts and related research, and, secondly, at demonstrating the importance and capabilities of a mixed methodological-statistical approach for pooling evidence from multiple studies in reaching meaningful scientific insights.

The *first study (TP1)* is a longitudinal multilevel meta-analysis, combining evidence on the developmental impact of process quality from 17 longitudinal studies and nine European countries for 2,887 ECEC centres and 16,461 children. The results show persistent effects of global and domain-specific process quality ($ES = .11$ and $.10$, Cohen's $d = .22$ and $.20$) on children's outcomes in mathematics and language/literacy. Mixed-effects moderator analyses show that the effects differ by the outcome domain, and choice of process measure and statistical-analytical approach.

In the *second study (TP2)* the approach of educational efficacy is discussed to evaluate the mathematical output of ECEC centres in Germany. Random intercept models analysed the mathematical effectiveness of 97 German centres of the BiKS-3-10 study based on the assessment of 554 children with the K-ABC subscale arithmetics. No differences in effectiveness emerged in fair comparisons, which took into account the children's previous math competences or developmentally relevant characteristics. Though the centres differed in children's level of mathematical competences from the beginning to the end of the childcare phase, this is probably linked to the divergent socio-cultural compositions of the centres. Comparing the pros and cons and the divergent results for process quality and effectiveness in BiKS points to a complementary use of quality and effectiveness analyses in ECEC.

Quality components relating to the educators are currently in the centre of quality debates, but are imprecisely conceptualized and understudied. The *third study (TP3)* provides an example of how structural and orientation quality components relating to the educators can be further differentiated and investigated. An exploratory analysis of a questionnaire survey revealed weaknesses in math-specific structural quality of the examined 29 German centres, as the 221 studied educators reported a low number of math courses attended in their vocational and fur-

ther training. Thus, for the majority of educators, formal experience with mathematics was restricted to school experience, which they recalled as neither particularly positive nor negative. A positive picture emerged for orientation quality components with relatively high mean levels of mathematical self-efficacy and value orientation. An exception was the comparably low mathematical self-concept of the educators. The regression analyses of the interrelations between structural and orientation components showed that orientations (self-concept in particular) were related to positive and negative school experiences. Additionally, a relationship emerged between educators' self-efficacy and a higher degree and more math courses in vocational training. Job experience was positively related to self-efficacy and value orientation.

The discussion summarizes the findings to emphasize the need for augmentation of competency and process measures in mathematics, and for development of economic questionnaire measures of process quality. It also outlines the potential of further efficacy analyses with improved study designs, and comparative analyses of quality and effectiveness of ECEC in other developmental areas. The role of the material-spatial environment for the quality of education and the effects of structural and orientation components raise many open questions. More research is also needed into the math-specific components relating to educators in regular provision in the light of the divergent findings for US teachers and from evaluations of professional trainings and initiatives. Pooling evidence allowed only partially to overcome limitations of individual studies, whereby the growing availability of data and more elaborate procedures open way to the investigation of more complex questions. The findings imply that policy measures and investments to improve process quality, e.g., strengthening the mathematical professionalization of educators, are able to maximize the educational impact of ECEC. A target parameter for a nationwide implementation of mathematical education, outside of programs and initiatives, is identified in the continuous monitoring of process quality with interaction-focused observation methods, as well as, if applicable, of the educational output. Furthermore, it is also necessary to promote the socio-cultural mix in centre-based ECEC and to develop educational concepts for pedagogical work with diversity.

Key words: Early Childhood Education and Care, mathematics, quality, educational effectiveness

1 Einleitung

Die Mathematik gilt als Schlüsselwissenschaft für Leistungen und Fortschritt in vielen Bereichen wie Medizin, Informatik, Technik sowie Wirtschafts- und Sozialwissenschaften (Baptist & Raab, 2007). Zudem sind mathematische Kompetenzen zentral für den individuellen akademischen und beruflichen Erfolg (Clements & Sarama, 2011; Heckman, 2006; Rammstedt, 2013; Sammons et al., 2014). Mathematische Lernprozesse beginnen bereits in der frühen Kindheit und finden unter anderem in frühkindlichen Einrichtungen statt. Die Fremdbetreuung bildet heutzutage einen üblichen Bestandteil in der Biographie der meisten Kinder in Europa (European Commission/EACEA/Eurydice/Eurostat, 2014; OECD, 2015a). Mehrheitlich handelt es sich hierbei um Formen der regulären frühkindlichen institutionellen Bildung und Betreuung (FIBB) in Kindertagesstätten, Vorschulen, Krippen oder ähnlichen Einrichtungen der Regelversorgung.

Studien unterschiedlicher Forschungsstränge haben das mathematische Bildungspotenzial frühkindlicher Einrichtungen aufgezeigt (Anders, 2013; Burger, 2010; Camilli, Vargas, Ryan & Barnett, 2010; Clements & Sarama, 2011; Deutscher & Selter, 2013; Melhuish et al., 2015; Pauen & Herber, 2009). Während etliche Befunde zum Bildungspotenzial FIBB aus Förderprogrammen oder Bildungsoffensiven bzw. -initiativen stammen, wurde in Europa im Rahmen mehrere Längsschnittstudien das Bildungspotenzial der FIBB-Regelversorgung untersucht (vgl. Anders, 2013; Burger, 2010; Melhuish et al., 2015). Dabei kristallisierte sich die pädagogische Qualität als Forschungsschwerpunkt heraus. In Deutschland und in den meisten Ländern Europas haben frühkindliche Einrichtungen generell einen offiziellen Bildungsauftrag in Mathematik (OECD, 2006; Sylva, Ereky-Stevens & Ariescu, 2015). Damit steht die Bildungspolitik vor der Aufgabe, flächendeckend das Bildungspotenzial in der FIBB sicherzustellen. Einen besonderen Erkenntniswert besitzen in diesem Zusammenhang die Befunde zur regulären FIBB und damit Modelle pädagogischer Qualität.

Der Qualitätsansatz und die angestoßene Qualitätsdebatte haben national und international wesentlich zur Weiterentwicklung der mathematischen Bildung in der FIBB beigetragen und bedeutende Forschungsvorhaben hervorgebracht. Insgesamt betrachtet ergeben sich trotzdem zentrale Forschungslücken und Ansatzpunkte, um gängige Qualitätsmodelle, Messverfahren und Forschungsdesigns zu verfeinern. Im vorliegenden Promotionsvorhaben wird sich daher in *drei Teilprojekten (TP)* den Komponenten und Auswirkungen der Qualität mathematischer Bildung in frühkindlichen Einrichtungen der Regelversorgung gewidmet. Zudem ist ein Ziel des Vorhabens den Erkenntnisgewinn aufzeigen, der aus der Bündelung der Evidenz unter-

schiedlicher Studien entstehen kann. In der Arbeit wird bei der Beschreibung des theoretischen Hintergrunds und des empirischen Forschungsstands ausschließlich auf Formen der vergleichsweise gut erforschte FIBB fokussiert ohne dabei anderen Formen der Fremdbetreuung wie Spielgruppen oder Tagesmüttern/-vätern ein mathematisches Bildungspotenzial abzuerkennen. Außerdem wird sich im Vorhaben zwar auf die internationale Forschung zu mathematischen Lern- und Bildungsprozessen berufen, es wird jedoch ein Schwerpunkt auf die europäische Evidenz gesetzt und dabei an mehreren Stellen vertiefend auf die FIBB in Deutschland eingegangen.

Die Arbeit beginnt im *zweiten Kapitel* mit wichtigen Grundlagen früher mathematischer Lern- und Anregungsprozesse in der Familie und FIBB. Anschließend geht das *dritte Kapitel* auf frühkindliche Einrichtungen als Bildungsinstitutionen der Mathematik ein. Einen Überblick über die wissenschaftliche Auseinandersetzung mit mathematischer Bildung in der FIBB gibt das *vierte Kapitel* mit einem Fokus auf Forschungsprojekte zur pädagogischen Qualität. Das *fünfte Kapitel* bietet die theoretischen Grundlagen für die TP des Promotionsvorhabens, indem es zentrale Komponenten der Qualität mathematischer Bildung und ihre Erfassung behandelt. Das Kapitel beginnt mit einer Beschreibung von Komponenten etablierter Qualitätsmodelle insbesondere Prozesskomponenten, welche zentral für die Erklärung der Auswirkung FIBB auf die kindliche Entwicklung sind (TP1). Im Anschluss wird die mathematische Bildungseffektivität als ergänzende Komponente der Qualität vorgestellt (TP2), um zuletzt auf wichtige Ansatzpunkte für die Ergänzung fachkraftbezogener Komponenten einzugehen (TP3). Das *Kapitel sechs* fasst den für die TP relevanten empirischen Forschungsstand zur Qualität und Effektivität mathematischer Bildung in der FIBB zusammen. Das erste Unterkapitel beschreibt Befunde zum Durchschnittsniveau der Prozessqualität und illustriert exemplarisch die europäische Befundlage zu deren Auswirkung auf die mathematische Entwicklung (TP1). Die folgenden Unterkapitel umfassen die Befunde zur mathematischen Bildungseffektivität der Studie Effective Pre-School, Primary and Secondary Education (EPPSE; Sylva, Melhuish, Sammons, Siraj-Blatchford & Taggart, 2010) und die vorliegende Evidenz zu fachkraftbezogenen Qualitätskomponenten (TP3). Das *Kapitel sieben* geht auf die zugrundeliegenden Desiderata und Ziele der drei TP sowie das methodische Anliegen des Promotionsvorhabens ein. Es folgen drei Kapitel, welche die TP und den jeweiligen spezifischen theoretischen Hintergrund zusammenfassend beschreiben: Das *Kapitel acht* stellt die Metaanalyse europäischer Längsschnittstudien des TP1 zur Auswirkung der Prozessqualität in der FIBB vor. Im *Kapitel neun* zum TP2 wird der Ansatz der mathematischen Bildungseffektivität für die FIBB in Deutschland diskutiert und mit Daten der Studie Bildungsprozesse, Kompetenzentwicklung

und die Formation von Selektionsentscheidungen (BiKS 3–10; von Maurice et al., 2007) erprobt. Schließlich werden im *Kapitel zehn* wertvolle Ergänzungsmöglichkeiten für fachkraftbezogene Qualitätskomponenten erwogen. Dies erfolgt anhand der Ergebnisse einer explorativen Studie des TP3 zur Qualifikation, zur Professionalisierung und zu Schulerfahrungen deutscher frühpädagogischer Fachkräfte. Zudem werden die motivationalen mathematischen Orientierungen der Fachkräfte untersucht. Zuletzt bietet *Kapitel elf* eine zusammenfassende Diskussion der empirischen Beiträge des Promotionsvorhabens und leitet Implikationen für die Forschung zur Qualität mathematischer Bildung in der FIBB ab sowie bildungspolitische und praktische Implikationen für die FIBB-Regelversorgung.

2 Grundlagen früher Lern- und Anregungsprozesse

Das Kapitel 2 widmet sich den theoretischen Vorstellungen und empirischen Erkenntnissen über frühe mathematische Lernprozesse und die Anregung in der Familie und FIBB, welche die Grundlage für die Auseinandersetzung mit mathematischen Bildungsprozessen in frühkindlichen Einrichtungen bilden. Im Kapitel 2.1 werden wichtige Schritte der mathematischen Entwicklung in der frühkindlichen Phase thematisiert, um die Relevanz früher Anregung und Stimulation in Mathematik aufzeigen. Zudem wird erläutert, dass die mathematische Entwicklung besonders von spezifischer Anregung und Stimulation in Mathematik profitiert. Im Anschluss wird in Kapitel 2.2 das ökosystemische Verständnis erklärt, das der Arbeit zugrunde liegt, während im Kapitel 2.3 die Rolle familialer Merkmale für frühe mathematische Lernprozesse in den Blick genommen werden.

2.1 Bedeutung der frühkindlichen Phase und der spezifischen Anregung in Mathematik

Die frühkindliche Forschung und Praxis weist seit einiger Zeit auf die Bereichsspezifität früher Lernprozesse hin (Burger, 2014; OECD, 2006; Roßbach & Weinert, 2008; Schneider, Küspert & Krajewski, 2013; Stamm & Edelmann, 2012; Sylva et al., 2015). Zwar gibt es bereichsunspezifische bzw. -übergreifende Merkmale von Kindern, welche mit dem Erwerb von Kompetenzen in unterschiedlichen Entwicklungsbereichen zusammenhängen, z. B. Alter, Intelligenz, Geschlecht oder die Kapazität des Arbeitsgedächtnisses (Anders, 2013; Anders et al., 2012; Kaufmann, 2011; Schneider et al., 2013), aber in der Forschung wird vermehrt die besondere Relevanz spezifischen Wissens sowie spezieller Fertigkeiten und Fähigkeiten im jeweiligen Entwicklungsbereich betont.

Der Erwerb mathematischer Kompetenz beginnt in der frühen Kindheit (Kaufmann, 2011; Pauen & Herber, 2009; Roßbach & Weinert, 2008). Bereits bei Säuglingen sind erste Grundlagen für die Auseinandersetzung mit Mathematik beobachtbar: So können sie unter bestimmten Bedingungen Unterschiede zwischen Mengen oder Veränderungen in Mengen erkennen. Dehaene (2011) geht davon aus, dass Säuglinge mit einem angeborenen Number Sense bzw. Zahlensinn ausgestattet sind. Insbesondere für den numerischen Bereich gibt es Untersuchungen zum weiteren Entwicklungsverlauf in den ersten Lebensjahren beispielsweise zum schrittweisen Erwerb der Zahlreihe und Zählprinzipien und zum Meistern einfacher Additions- bzw. Subtraktionsaufgaben (Cross, Woods & Schweingruber, 2009; Kaufmann, 2011; Pauen & Herber, 2009). Bereits im Vorschulalter lassen sich Disparitäten in der mathematischen Kompetenz nachweisen (Anders et al., 2012; Heckman, 2006; Kaufmann, 2011; Pauen &

Herber, 2009; Roßbach & Weinert, 2008; Sammons et al., 2002). Früh erworbene Kompetenzen bilden die Grundlage für die weitere mathematische Entwicklung in der Schulzeit und gelten daher als sogenannte Vorläuferfähigkeiten (Anders, Grosse, Roßbach, Ebert & Weinert, 2013; Aunola, Leskinen, Lerkkanen & Nurmi, 2004; Jordan, Kaplan, Ramineni & Locuniak, 2009; Lehl, Kluczniok & Roßbach, 2016; Locuniak & Jordan, 2008; Schneider et al., 2013). Existierten vornehmlich Belege für den Zusammenhang von schulischen Leistungen mit den frühen Mengen-, Zahl- und Zählkompetenzen (Baroody, Lai & Mix, 2006; Burger, 2014; Kaufmann, 2011; Pauen & Herber, 2009; Schneider et al., 2013), berücksichtigen aktuellere Untersuchungen ebenfalls Kompetenzen aus anderen Bereichen der Mathematik, z. B. Messen, Raum und Form oder Häufigkeiten und Wahrscheinlichkeiten (z. B. Clements & Sarama, 2011; Fried, Hoeft, Isele, Stude & Wexler, 2012; Neumann et al., 2013). Bei der Untersuchung von Einflüssen der Lernumwelten wie der Familien, der frühkindlichen Einrichtungen oder Schulen rückten ebenfalls vermehrt die spezifische Anregung und Stimulation in Mathematik in den Vordergrund.

2.2 Ökosystemische Sicht auf frühe Lern- und Anregungsprozesse

Aus ökosystemischer Perspektive findet Lernen in der Interaktion und Auseinandersetzung mit der Umwelt statt und die Umwelten, in die mathematische Lern- und Bildungsprozesse eingebettet sind, spielen für individuelle Entwicklungsverläufe eine große Rolle (Bronfenbrenner, 1986; Bronfenbrenner & Morris, 2006; Evangelou, Sylva, Kyriacou, Wild & Glenny, 2009; Roßbach & Weinert, 2008; Scarr & Eisenberg, 1993). Am Beginn der Lebensspanne stellt nach der Familie die frühkindliche Einrichtung für die meisten Kinder Europas die primäre Umwelt dar, in der Kinder Erfahrungen, Wissen und Kompetenzen sammeln (Bronfenbrenner, 1986; European Commission/EACEA/Eurydice/Eurostat, 2014; Krüger & Grunert, 2010; OECD, 2015a). Mit den primären Lernumwelten, auch Mikrosysteme genannt, befindet sich das lernende Kind in unmittelbarer Interaktion, weshalb sie die kindliche Entwicklung in besonderem Maße beeinflussen. Dabei stehen Kind und Mikrosysteme in wechselseitiger Abhängigkeit und bedingen sich gegenseitig. Als Mesosystem bezeichnet die Theorie die Wechselwirkungen zwischen den Systemen. So können mitunter auch Systeme Einfluss auf die kindliche Entwicklung ausüben, mit denen das Kind selbst nicht direkt interagiert. Beispielsweise können Stress am Arbeitsplatz und lange Arbeitszeiten der Mutter sich ungünstig auf die Mutter-Kind-Interaktion und somit die Anregung und Stimulation von Entwicklungsprozessen in der Familie auswirken. Das Gesamt des Beziehungsgeflechts wird Exosystem genannt. Die gesamtgesellschaftliche Situation, in der mathematische Lern- und Anregungs-

prozesse eingebettet sind, wird im Makrosystem festgehalten. Dazu zählen die Normen, Werte, Traditionen, Konventionen oder Ideologien der Gesellschaft oder relevanter Gesellschaftsgruppen sowie die Gesetze und Vorschriften. Bronfenbrenner (1986) umschreibt die Dynamik des Bedingungsgefüges bzw. zeitliche Veränderungen als Chronosystem. Ein veränderter Stellenwert mathematischer Bildung im gesellschaftlichen Bewusstsein, neue, gesetzlich festgelegte Qualitätsstandards oder die Einführung von Bildungsplänen für die FIBB wirken sich etwa auf die Lern- und Bildungserfahrungen von Kindern aus.

Gemäß dem ökologischen Verständnis weisen Merkmale der Umwelten einen unterschiedlichen Grad an Proximität zum kindlichen Lernprozess auf (Bronfenbrenner & Morris, 2006; Ditton, 2000). Merkmale der Kinder, der frühpädagogischen Fachkräfte sowie der Gruppen in den Einrichtungen – z. B. Gruppengröße oder die Zusammensetzung hinsichtlich verschiedener Merkmale – gelten als proximale Bedingungsfaktoren, weil sie die Lern- und Bildungssituation in der FIBB direkt betreffen und somit zentral für den kindlichen Lernfortschritt sind (Cryer, Tietze, Burchinal, Leal & Palacios, 1999; Pianta et al., 2005). Distale Bedingungsfaktoren wie Einrichtungsmerkmale oder die Bildungspläne der Länder befinden sich weiter entfernt vom pädagogischen Geschehen und der Wirkungskette, beeinflussen aber indirekt das Lern- und Bildungsgeschehen in den Gruppen.

2.3 Entwicklungsrelevante Merkmale der Familien

Ein Wechselspiel aus individuellen, familialen und institutionellen Bedingungsfaktoren erklärt die mathematischen Entwicklungsverläufe von Kindern in FIBB, weshalb viele Studien zur Auswirkung FIBB zusätzlich individuelle Merkmale der Kinder und entwicklungsrelevante Merkmale der Familie einbeziehen. In Studien werden mehrheitlich strukturelle Merkmale d. h. relativ stabile, überdauernde Merkmale der Familie betrachtet (Lehrl, 2013). Insbesondere werden Indikatoren der Bildungsbenachteiligung berücksichtigt, die mit frühen Disparitäten in der mathematischen Kompetenz zusammenhängen, wie ein Migrationshintergrund bzw. eine andere ethnische Abstammung des Kindes, ein geringes Einkommen der Eltern oder ein niedriger Bildungsabschluss sowie beruflicher Status, oder ein spezifischer Indikator für den sozioökonomischen Status der Familie (SES für Socioeconomic Status; vgl. Abreu-Lima, Leal, Cadima & Gamelas, 2013; Anders et al., 2013; Anders et al., 2012; Arnold & Doctoroff, 2003; Lehrl et al., 2016; Lehrl, Kuger & Anders, 2014; Sammons et al., 2002; Sammons et al., 2007; Schneider et al., 2013; Sylva et al., 2010). In einige Studien werden landesspezifische Indikatoren verwendet: So wurde in EPPSE die Berechtigung auf subventionierte Mahlzeiten in den Einrichtungen als Indikator einer Benachteiligung genutzt (FSM-Indikator für free

school meals; Sammons et al., 2002; Sammons et al., 2007). Vermehrt wird ebenfalls die familiäre Anregung bzw. das Home Learning Environment abgebildet (HLE; Kluczniok, Lehl, Kuger & Roßbach, 2013; Lehl, 2013). Das HLE bezieht sich zum einen auf die globale Qualität familialer Anregung, wie ein unterstützendes, fürsorgliches und warmes elterliches Verhalten, zum anderen auf die spezifische Anregung und Förderung in Bereichen wie Mathematik und Sprache/Literacy durch Aktivitäten, Interaktionen und Materialien (Kluczniok et al., 2013; Lehl, 2013). Ein Zusammenhang zwischen HLE bzw. familiäre Förderung und mathematischer Kompetenz ließ sich in mehreren Studien nachweisen (Anders et al., 2013; Anders et al., 2012; Blevins-Knabe & Musun-Miller, 1996; LeFevre et al., 2009; Lehl et al., 2016; Niklas & Schneider, 2012; Niklas, Segerer, Schmiedeler & Schneider, 2012; Saxe et al., 1987; Susperreguy Jorquera, 2013; Sylva et al., 2013), wenn sich auch in der BiKS-Studie für den Kindergartenzeitraum teilweise keine Belege fanden (Anders et al., 2013; Anders et al., 2012; vgl. Lehl, 2013).

3 Frühkindliche Einrichtungen als Bildungsinstitutionen der Mathematik

Dem Promotionsvorhaben liegt ein Verständnis von frühkindlichen Einrichtungen als Bildungsinstitutionen der Mathematik zugrunde, auf welches im Folgenden eingegangen wird. Das Kapitel 3.1 beginnt mit einer Beschreibung der Entstehungsgeschichte der mathematischen Bildung in frühkindlichen Einrichtungen. Im Folgenden wird im Kapitel 3.2 die politische Verankerung der Mathematik in der FIBB durch Bildungs- und Rahmenlehrpläne thematisiert. Anschließend beschreibt das Kapitel 3.3, welche Erwartungen an die mathematische Bildungsarbeit in der FIBB gestellt werden und welche Rolle hierbei die frühpädagogischen Fachkräfte einnehmen. Im Kapitel 3.4 werden die Konsequenzen für die Umsetzung des mathematischen Bildungsauftrags erläutert, die sich aus der hierarchischen Struktur FIBB ergeben. In Kapitel 3.5 erfolgt eine Gegenüberstellung der Ansätze früher Bildung in Europa und den USA, welche mit der Debatte um universelle vs. zielgruppenspezifische Bildungsangebote verbunden ist.

3.1 Entstehungsgeschichte frühkindlicher mathematischer Bildung

Bereits in den historischen Anfängen der frühkindlichen Bildung und Betreuung in Deutschland im 19. Jahrhundert galten die Kindergärten bei Friedrich Fröbel und seinen Vertretern als erste Stufe des Bildungssystems (Baroody et al., 2006; Hasemann & Gasteiger, 2014; Henschen & Teschner, 2013; Kaufmann, 2011) und die mathematische Bildung gehörte zum Bildungskanon frühkindlicher Einrichtungen, wie die Spielgaben Fröbels verdeutlichen (Fröbel, 1850). Am Beginn des 20. Jahrhunderts gehörten mathematische Grunderfahrungen für Maria Montessori ebenfalls zum elementarpädagogischen Alltag. So entwickelte sie zusätzlich zu Sinnes-, Alltags- und Sprachmaterialien für die frühe Förderung in Einrichtungen mathematische Materialien (Kaufmann, 2011).

Heutzutage ist die FIBB als erste Phase institutionellen Lernens allgemein anerkannt und die mathematische Bildung gehört mehrheitlich offiziell zum Bildungsauftrag frühkindlicher Einrichtung weltweit, allerdings mit einem Fokus auf die Drei- bis Sechsjährigen (Oberhumer, 2012; OECD, 2006, 2011; Sylva et al., 2015). Denn in vielen Ländern vollzieht sich eine Trennung von Angeboten frühkindlicher Bildung, welche vornehmlich älteren Kindern vorbehalten sind, und Angeboten für jüngere Kinder, die eher auf die Betreuung ausgerichtet sind (z. B. hinsichtlich Ausmaß der Förderung, curricularer Richtlinien oder Qualitätsregulationen und -standards; Diskowski, 2009; European Commission/EACEA/Eurydice/Eurostat, 2014; OECD, 2011; Sylva et al., 2015). In den verschiedenen Ansätze und Forschungsstränge zur

mathematischen Bildung in der FIBB wurde sich daher lange Zeit weitgehend auf diese Altersgruppe beschränkt (Baroody et al., 2006; Burger, 2014; Hellmich, 2007). In letzter Zeit ist allerdings ein Trend in Forschung und Praxis erkennbar, Wege aufzuzeigen, wie auch jüngere Kinder in ihrer mathematischen Entwicklung von FIBB profitieren können (Anders, 2013; Dalli et al., 2011; OECD, 2011).

3.2 Politische Verankerung durch Bildungs- und Rahmenlehrpläne

Bildungspolitisch bildet die Verankerung der Mathematik in den elementarpädagogischen Bildungsplänen der meisten Länder die Grundlage der frühen mathematischen Bildung in der FIBB in Europa. Dem Gedanken der Bereichsspezifität von Lern- und Anregungsprozessen folgend sind die Bildungspläne mittlerweile nach Entwicklungsbereichen unterteilt (Anders, 2014; OECD, 2006, 2011; Sylva et al., 2015), wobei nur manche Pläne Lernziele, andere eher Grunderfahrungen für die Bereiche formulieren. Die Bildungspläne verweisen größtenteils auf die Besonderheiten frühkindlicher Lern- und Bildungsprozesse, z. B. eine an den kindlichen Bedürfnissen und Interessen orientierte Pädagogik oder den Grundgedanken *holistischer Förderung*, welcher eine umfassende und integrierte Förderung in verschiedenen Entwicklungsbereichen vorsieht.

Trotz der historisch langen Tradition früher mathematischer Bildung ist die offizielle Implementation in Deutschland im internationalen Vergleich z. B. zu England oder den USA spät erfolgt (Oberhuemer, 2012; OECD, 2011; Sylva et al., 2015). Im Zeitraum von 2002 bis 2006 führten die deutschen Bundesländer Bildungspläne für den Elementarbereich ein (Diskowski, 2009). 2004 beschlossen die Jugend- und Kultusministerkonferenz (JMK & KMK) einen „Gemeinsamen Rahmen der Länder für die frühe Bildung in Kindertageseinrichtungen“, welcher die die Bildungsfunktion FIBB hervorhebt, u. a. im Bereich „Mathematik, Naturwissenschaft, (Informations-)Technik“ (S. 4). Zudem wird betont, dass die Anschlussfähigkeit der Bildungsprozesse beim Übergang vom Elementar- in den Primarbereich gewährleistet sein muss. Einige Kritik gab es an der Ausgestaltung des mathematischen Bildungsbereichs in den Bildungsplänen der einzelnen Bundesländer der ersten Generation (vgl. Peter-Koop, 2009; Royar, 2007). Beispielsweise fehlte der Bereich gänzlich in den Plänen von Thüringen und Bremen und die Aussagen zur Mathematik in den Plänen aus Baden-Württemberg, Nordrhein-Westfalen und Schleswig-Holstein fielen sehr allgemein aus. Bei der Überarbeitung mehrere Bildungspläne ist eine Konkretisierung, Differenzierung und Ausweitung des mathematischen Bildungsauftrags erkennbar. In den aktuellen Fassungen ist die Mathematik in allen Bildungsplänen der einzelnen Bundesländer enthalten (vgl. Diskowski, 2009; Ministerium für Bildung,

Jugend & Sport, 2015; Peter-Koop, 2009; Royar, 2007; Senatorin für Soziales, Kinder, Jugend & Frauen der Freien Hansestadt Bremen, 2012).

Ebenfalls gestiegen ist der Stellenwert der Mathematik in der Aus- und Fortbildung frühpädagogischer Fachkräfte, z. B. mit Blick auf die Zulassungsvoraussetzungen und Abschlussprüfungen, in Form mathematischer Professionalisierungsangebote. Insgesamt wird die mathematikspezifische Qualifikation und Professionalisierung jedoch noch immer als weitgehend unzureichend kritisiert (Baumeister & Grieser, 2011; Beher & Walter, 2012; Deppe, 2011; Gasteiger, 2017; Janssen, 2010; KMK, 2009; Ledig, 2011).

3.3 Erwartungen an die mathematische Bildungsarbeit in der FIBB

Allgemein wird in der Elementarpädagogik die Einzigartigkeit der frühkindlichen Entwicklungsphase betont, in der eine holistische, ganzheitliche Förderung und das spielbasierte Lernen im Vordergrund stehen sollen (Anders, 2014; Gasteiger, 2012; König, 2007; OECD, 2011; Schuler, Pelzer, Wittkowski & Wittmann, 2015; Sylva et al., 2015). In Deutschland ist der *Situationsansatz*, in welchem die ganzheitliche Förderung in Alltagssituationen angestrebt wird, stark im elementarpädagogischen Bildungsverständnis verwoben. Im Vergleich zum schulischen ist der frühpädagogische Alltag weniger strukturiert, weil sich Bildungs- und Lernsituationen nicht nach festgelegten Zeiten oder Stundenplänen richten und selten in Form strukturierter Lernarrangements stattfinden. Stattdessen sollen sich Bildungsangebote an den individuellen Bedürfnissen und Interessen der Kinder, der Gruppendynamik sowie dem Alltagsgeschehen in den Einrichtungen orientieren.

Der Situationsansatz steht in der Kritik, weil primär die Persönlichkeitsentfaltung im Vordergrund stünde und die Bildungsförderung dadurch vernachlässigt würde (Carle & Wenzel, 2007). Kritik gibt es ebenfalls an der Umsetzung des spielbasierten Lernens (vgl. Slot, 2014; van Oers, 2002). Es existieren allerdings Vorschläge verschiedener Bildungsoffensiven und -initiativen sowie aus der Fachdidaktik, die eine in den Alltag eingebettete mathematische Bildung umsetzbar werden lassen (Burger, 2014; Gasteiger, 2012; Ginsburg, Lee & Boyd, 2008; Kaufmann, 2011; Sommerlatte, Lux, Meiering & Führlich, 2009). Unter gewissen Voraussetzungen können Spielsituationen ebenso gelungene Bildungsmöglichkeit repräsentieren (Burger, 2014; Dooley, Dunphy & Shiel, 2012; NAEYC & NCTM, 2010; Rechsteiner, Hauser & Vogt, 2012; Siraj-Blatchford, 2009; van Oers, 2010).

Wichtig erscheinen jedoch eine ausreichende Unterstützung und Anregung durch die Fachkraft für das Ausschöpfen des Bildungspotenzials von wenig strukturierten Lernsituationen und -angeboten. Die *Ko-Konstruktion* hat sich als didaktische Leitidee weitgehend durch-

gesetzt. Sie bezeichnet einen sozialkonstruktiven Lernprozess, welcher durch einen wechselseitigen Austausch der frühpädagogischen Fachkräfte und Kinder gekennzeichnet ist (Fthenakis, Schmitt, Daut, Eitel & Wendell, 2009; Gasteiger, 2007, 2012; Hardy & Steffensky, 2014; König, 2007). Die Beteiligten handeln gemeinsam die Bedeutungen und den Sinn von Lernerfahrungen aus und konstruieren das Wissen gemeinschaftlich. Momente, in denen alle Beteiligten gemeinsam versuchen Gedankengänge zu entwickeln und Probleme zu lösen, in denen sie ihre Gedanken und Ideen miteinander austauschen und Aktivitäten aufeinander abstimmen, werden als *geteilte Denkprozesse* bzw. *Sustained Shared Thinking* bezeichnet und gelten als besonders wichtig (Gasteiger, 2007; Hardy & Steffensky, 2014; Siraj-Blatchford, 2009; Siraj-Blatchford, Sylva, Muttock, Gilden & Bell, 2002). Fachkräfte begleiten Bildungsprozesse in *Scaffolding-Prozessen*, indem sie an geeigneter Stelle Kindern bei der Bewältigung des nächsten Entwicklungsschritts in unterstützend zur Seite stehen (Fthenakis et al., 2009; Gasteiger, 2007, 2012; Hardy & Steffensky, 2014; König, 2007). Ginsburg et al. (2008) betonen die Bedeutung von Bildungsmomenten bzw. *Teachable Moments* für die mathematische Bildung in der FIBB (vgl. Hyun & Marshall, 2003): Fachkräfte sollen Momente, die sich für die mathematische Förderung eignen und während des Spiels und anderer Aktivitäten spontan entstehen, beobachten und nutzen. Deutlich wird, dass das beschriebene didaktische Verständnis früher mathematischer Bildung den Fachkräften außerordentliche Motivation, Flexibilität und Spontaneität abverlangt (Carle & Wenzel, 2007).

3.4 Die hierarchische Struktur der FIBB und ihre Bedeutung

Wie schulische Bildungssysteme weisen FIBB-Systeme üblicherweise eine hierarchische Struktur auf (Creemers & Kyriakides, 2009; Cryer et al., 1999; Ditton, 2000; Pianta et al., 2005). Gerade das deutsche FIBB-System ist durch eine komplexe Verschachtelung gekennzeichnet (Carle & Wenzel, 2007; Diskowski, 2009; Larra, 2005; Roux & Tietze, 2007): So findet die Bildungsarbeit in Deutschland im Allgemeinen in Gruppen statt, die in Einrichtungen organisiert sind, die wiederum zumeist verschiedenen Trägern zugeordnet sind. Auf übergeordneter Ebene gehören Einrichtungen verschiedenen Kommunen und Bundesländern und unterschiedlichen sozial-regionalen Kontexten an. Es gibt pädagogische Konzepte, in denen die Gruppenebene in dem Maße nicht existiert wie das offene Konzept, welches in Deutschland weite Verbreitung genießt: Das Prinzip der Stammgruppen wird vollständig oder teilweise aufgelöst und durch flexible, teils frei wählbare Lernangebote und -gruppen ersetzt (vgl. Tietze et al., 2012).

In anderen Ländern sind die FIBB-Systeme ebenfalls hierarchisch organisiert (Cryer et al., 1999; Pianta et al., 2005), was neben Implikationen für die frühkindliche Bildungsforschung (z. B. für die Analysen in allen TP; siehe Kapitel 8 bis 10) wichtige steuerungspolitische Implikationen hat (Diskowski, 2009; Larra, 2005). Höhere Ebenen dienen als sogenannte Unterstützungssysteme hierarchisch untergeordneter Ebenen (supporting environments, Ditton, 2000; Scheerens & Creemers, 1989). Beispielsweise bilden rechtliche Vorgaben auf Landesebene (z. B. Vorgaben zum Erzieher-Kind-Schlüssel oder Vorgaben der Bildungspläne) und deren Umsetzung durch Träger und Einrichtungen, den Rahmen für die tägliche pädagogische Arbeit der frühpädagogischen Fachkräfte in den Gruppen (Cryer et al., 1999; Pianta et al., 2005; Tietze, 2008). Es ist jedoch nicht davon auszugehen, dass eine Einführung von Bildungsplänen, welche Leitlinien für die mathematische Bildungsarbeit und ggfs. Bildungsziele formulieren, ausreicht um die frühe mathematische Bildung in der FIBB flächendeckenden zu implementieren (vgl. Diskowski, 2009; Ditton, 2000; Roux & Tietze, 2007). Stattdessen sind Veränderungsprozesse auf allen Ebenen erforderlich: Erfolgreich wird die Implementation nur dann sein, wenn der bildungspolitische Einzug mathematischer Bildung in der FIBB seinen Niederschlag in den schriftlichen Träger- bzw. Einrichtungskonzeptionen findet und von einer Stärkung des mathematischen Bildungsauftrags in dem Bewusstsein und in der pädagogischen Arbeit des Personals begleitet wird. Nur dann wird das *offizielle Curriculum* zu einem *gelebten Curriculum* im pädagogischen Alltag der Einrichtungen und Gruppen (vgl. Sylva et al., 2015).

3.5 Debatte um universelle vs. zielgruppenspezifische Bildungsangebote in Europa und den USA

Obschon mathematische Bildung in den meisten Ländern zum Bildungsauftrag FIBB zählt bestehen wichtige Unterschiede im Bildungsansatz und hinsichtlich weiterer Charakteristika FIBB innerhalb Europas und insbesondere im Vergleich zu den USA (Anders, 2014; EACEA/Eurydice, 2009; European Commission/EACEA/Eurydice/Eurostat, 2014; OECD, 2006, 2013). In Europa handelt es sich bei FIBB mehrheitlich um universelle Angebote, die prinzipiell allen Kindern offen stehen. Spätestens mit ungefähr vier Jahren befinden sich durchschnittlich 91 % der Kinder in Europa und nahezu alle Kinder in Deutschland in Fremdbetreuung, davon der überwiegend Anteil in FIBB (European Commission/EACEA/Eurydice/Eurostat, 2014; OECD, 2015a). Die Betreuungsrate liegt damit über dem OECD-Durchschnitt, wobei die Rate bei jüngeren Kindern deutlich niedriger ist mit einem allgemeinen Trend zu einer früheren Fremdbetreuung. Die Betreuungsrate für Vierjährige in den USA liegt demgegenüber gerade einmal bei 66 %.

Wenngleich es ein generelles Ziel FIBB ist, allen Kindern gleiche Startchancen für eine akademische Karriere und Lebenserfolg zu sichern und herkunftsbedingte Disparitäten in der kindlichen Entwicklung abzubauen (sogenannte *kompensatorische Effekte*, vgl. Anders, 2013; EACEA/Eurydice, 2009; OECD, 2006), unterscheiden sich die Strategien zur Bekämpfung der Benachteiligung durch frühe Bildung in Europa und den USA. In den USA richtet sich ein vergleichsweise großer Anteil staatlich subventionierter Angebote, insbesondere förderorientierte Maßnahmen, selektiv an sozial benachteiligte Kinder (sogenannte *zielgruppenspezifische Angebote*; Barnett, 2011; Barnett, Brown & Shore, 2004; OECD, 2006). Vorteil zielgruppenspezifischer Angebote ist es, dass sie besser an die speziellen Bedürfnisse benachteiligter Kinder und Familien ausgerichtet werden können. Mehrere Autoren kritisieren jedoch, dass eine institutionalisierte Selektion benachteiligter Kinder die Diskriminierung benachteiligter Gruppen und Segregationsprozesse verstärkt (Barnett, 2011; Barnett et al., 2004; OECD, 2006; Penn, 2011). Außerdem führen Bradbury, Corak, Waldfogel und Washbrook (2015) an, dass die großflächige Ausstattung von FIBB-Systemen mit staatlich bezuschussten Angeboten für Geringverdiener zum Nachteil der Mittelschicht sei.

Im Gegensatz dazu verfolgt die europäische FIBB mehrheitlich die Idee gleiche Entwicklungschancen durch einen *universellen Zugang zu qualitativ hochwertiger Bildung* zu gewährleisten (OECD, 2006). An die FIBB wird ein inklusiver oder zumindest integrativer Anspruch gestellt: Die kompensatorische Förderung benachteiligter Kinder soll innerhalb des Regelsystems stattfinden. EACEA/Eurydice (2009) kritisieren allerdings, dass trotz des nominell, universellen Zugangs FIBB bildungsbenachteiligten Kindern faktisch gleiche Zugangs- und Teilhabechancen in den meisten Ländern verwehrt bleiben (vgl. Becker, 2010; Becker & Schober, 2015; Geier & Riedel, 2009; Leseman & Slot, 2014; OECD, 2006; Schober & Spieß, 2012); (Lehrl et al., 2014). Das betrifft einerseits die Inanspruchnahme FIBB, andererseits den Zugang zu qualitativ hochwertigen Angeboten. Zudem mehren sich die Anzeichen soziokultureller Segregation in der FIBB: Kinder mit und ohne Migrationshintergrund und unterschiedlichem sozioökonomischem Status scheinen tendenziell seltener die gleichen Einrichtungen zu besuchen.

4 Wissenschaftliche Auseinandersetzung mit mathematischer Bildung in der FIBB

Das Verständnis der Mathematik als elementarpädagogischem Bildungsbereich fußt auf der wissenschaftlichen Auseinandersetzung verschiedener Forschungsrichtungen, über das Kapitel 4 einen Überblick bietet. Zu Beginn wird der Beitrag der Fachdidaktik im Kapitel 4.1 diskutiert, welcher auf wichtige Inhalts- und Prozesskompetenzen der elementarpädagogischen Mathematik verweist, die in Forschung und Praxis nur unzureichend abgedeckt werden. Während entwicklungspsychologische und hirnpyschologische Beiträge wichtige Entwicklungsschritte, -theorien und -prinzipien in der frühkindlichen Phase empirisch untersuchten (vgl. Baroody et al., 2006; Cross et al., 2009; Kaufmann, 2011; Krajewski, Grübing & Peter-Koop, 2009; Pauen & Herber, 2009; Roßbach & Weinert, 2008; Schneider et al., 2013), erforschten verschiedene Forschungsstränge das mathematische Bildungspotenzial frühkindlicher Einrichtungen mit jeweils unterschiedlichen Implikationen für die Umsetzung des Bildungsauftrags in der pädagogischen Praxis, auf welche in folgenden Kapiteln eingegangen wird. Im Kapitel 4.2 und 4.3 werden die Forschung zu mathematischen Förderprogrammen und zu Bildungsoffensiven bzw. -initiativen vorgestellt und schließlich wird im Kapitel 4.4 auf die Forschung zur mathematischen Bildung in der FIBB-Regelversorgung eingegangen, welche im Fokus des Promotionsvorhabens steht.

4.1 Fachdidaktische Auseinandersetzung mit mathematischer Bildung

Die Fachdidaktik stellt Überlegungen zu Teilkompetenzen und Grunderfahrungen des Bildungsbereichs Mathematik in der FIBB an, welche sich stark an die Bildungsziele für den Primarbereich anlehnen (Stiftung Bildungspakt Bayern, 2007; Gasteiger, 2010; Hellmich, 2007; Kaufmann, 2011; NAEYC & NCTM, 2010). Denn aus fachdidaktischer Perspektive ist es wichtig, die Anschlussfähigkeit der Lernprozesse beim Übergang von der FIBB in die Schule sicherzustellen. Die zu erwerbenden Kompetenzen werden einerseits gemäß *Inhaltsbereichen* des Primarbereichs in folgende fünf Bereiche unterteilt (KMK, 2004; NCTM, 2000): (a) Zahlen und Operationen, (b) Raum und Form, (c) Muster und Strukturen, (d) Größen und Messen sowie (e) Daten und Häufigkeiten. Zusätzlich zu den Inhaltskompetenzen werden *Prozesskompetenzen* benannt, welche Kinder in Auseinandersetzung mit der Mathematik erwerben und welche entscheidend dafür sind, wie Kinder mathematische Kenntnisse, Fähigkeiten und Fertigkeiten erwerben und nutzen. Hierzu zählen (a) das Problemlösen, (b) das Kommunizieren, (c) das Modellieren, (d) das Argumentieren und (e) das Darstellen von Mathema-

tik. Prozessbezogene Kompetenzen gelten als zentral für eine positive Grundhaltung zur Mathematik und für den weiteren mathematischen Werdegang. Viele der empirischen Erkenntnisse über das Lernen und Fördern in den ersten Lebensjahren beschränken sich auf den Bereich Zahlen und Operationen. Für diesen Bereich existieren viele Testverfahren sowie zahlreiche Belege für seine Förderbarkeit und Vorhersagekraft für den weiteren Entwicklungsverlauf. Erst allmählich geraten weitere Kompetenzbereiche in den Blick der Forscher (z. B. Fried, et al., 2012; Gasteiger, 2012; Neumann et al., 2013; Zacharos & Ravanis, 2000). Das mangelnde Wissen über viele Kompetenzbereiche stellt aus Sicht der Fachdidaktik eine wichtige zu schließende Forschungslücke dar (vgl. Deutscher & Selter, 2013; Kaufmann, 2011).

4.2 Forschungsstrang zu mathematischen Förderprogrammen

Ein erster Forschungsstrang bezieht sich auf mathematische Förderprogramme (siehe Deutscher & Selter, 2013; Hasemann & Gasteiger, 2014; Hellmich, 2007; Quaiser-Pohl, 2008 zu deutschen Programmen sowie Clements & Sarama, 2011; Ginsburg et al., 2008 zu prominenten US-amerikanischer Programme). In Deutschland sind einige Programme recht verbreitet wie verschiedene Varianten von Zahlenland (Friedrich & Munz, 2004, 2011; Pauen & Herber, 2009), Mengen, Zählen, Zahlen (Krajewski, Nieding & Schneider, 2008; Krajewski, Renner, Nieding & Schneider, 2009), Spielend Mathe (Quaiser-Pohl, 2008; Rademacher, Trautewig, Günther, Lehmann & Quaiser-Pohl, 2005) oder Mathe 2000 (Pauen & Herber, 2009; Wittmann, 2004). Obschon die konkrete Umsetzung und das dahinterstehende Grundverständnis mathematischer Bildung voneinander abweicht, ähneln sich die Angebote in ihrem programmatischen Ansatz mathematischer Bildung: In der Regel geht es darum, Kinder zeitlich begrenzt in spezifischen Teilkompetenzen der Mathematik zu trainieren. Die Ziele und Inhalte der Programme sind in der Regel auf einen vergleichsweise kleinen Ausschnitt mathematischer Kompetenz begrenzt (Zahlen und Operationen sowie teilweise Raum und Form). Außerdem bezieht sich die Förderung lediglich auf einen Ausschnitt von einigen Wochen (selten mehr als zwei Monate) in der Vorschulphase (meist Vier- bis Sechsjährige).

Dabei arbeiten die Programme oft mit sich wiederholenden, ähnlichen Aufgaben und Aktivitäten sowie klaren Vorgaben für die mathematische Förderung, teilweise in manualisierter Form. So sind Materialien zumeist vorgegeben und Rahmenbedingungen klar definiert, beispielsweise die Gruppengröße sowie der zeitliche Umfang und Ablauf jeder Fördereinheit. Die einzelnen Fördereinheiten bauen im Allgemeinen sequentiell aufeinander auf. Der Fachkraft kommt eine eher instruktive Rolle zu und sie erhält genaue Vorgaben für das Umsetzen mathematischer Übungen und Aktivitäten.

Wirksamkeitsstudien deuten für viele dieser Programme in den jeweils definierten Zielbereichen kurzfristige und teilweise auch langfristige Effekte an (z. B. Arnold, Fisher, Doctoroff & Dobbs, 2002; Clements & Sarama, 2007, 2011; Friedrich & Munz, 2004, 2011; Ginsburg et al., 2008; Griffin, 2004; Griffin & Case, 1997; Krajewski et al., 2008; Krajewski, Renner, et al., 2009; Pauen & Herber, 2009; Sarama & Clements, 2004; Sarama, Clements, Starkey, Klein & Wakeley, 2008). Die Evaluationsstudien finden teilweise in experimentellen oder zumindest quasi-experimentellen Designs mit Prä-Post- und Treatment-Kontrollgruppenvergleichen statt, sind aber hinsichtlich Stichprobengröße und Follow-Up-Zeiträumen relativ beschränkt. Einige dieser Programme sind speziell für Risikogruppen konzipiert, z. B. Kinder mit Verdacht auf oder diagnostizierter Dyskalkulie bzw. Rechenschwäche. Insgesamt liefern die Ergebnisse aus Evaluationsstudien Hinweise darauf, wie mit spezifischen Materialien und didaktischen Methoden in eng umschriebenen Inhaltsbereichen unter speziellen zeitlichen und strukturellen Rahmenbedingungen wirksam gefördert werden kann. Die Möglichkeit aus den Förderprogrammen Handlungsimplicationen für die mathematische Bildungsarbeit in der FIBB-Regelversorgung abzuleiten, welche mit den Vorstellungen über die Bildungsarbeit und den Vorgaben in den Bildungsplänen übereinstimmen, ist daher stark eingeschränkt.

4.3 Forschungsstrang zu mathematischen Bildungsoffensiven bzw. –initiativen

In einem weiteren Forschungsstrang wurde die Wirksamkeit von relativ breit angelegten Bildungsoffensiven bzw. –initiativen untersucht, welche darauf abzielen, die frühkindliche Bildung als festen Bestandteil des pädagogischen Alltags in den Einrichtungen zu verankern. Bei den meisten Bildungsoffensiven und -initiativen in Deutschland, z. B. Stärkung der Bildungs- und Erziehungsqualität in Kindertageseinrichtungen und Grundschule – Gestaltung des Übergangs (Fried, et al., 2012), Natur-Wissen schaffen (Fthenakis et al., 2009) oder Kindergarten der Zukunft in Bayern (Roßbach, Sechtig & Freund, 2010; Stiftung Bildungspakt Bayern) handelt es sich um universelle Angebote, die allen Kinder im entsprechenden Alter in den Einrichtungen offenstehen. Viele Beispiele aus den USA, z. B. Big Math for Little Kidz (Presser, Clements, Ginsburg & Ertle, 2015) oder das High/Scope Perry Preschool Programm (Schweinhart et al., 2005; Schweinhart & Weikart, 1997) sowie die niederländischen Voor- en Vroegschoolse Educatie Programme (VVE, Pre- and early primary school education; de Haan, Elbers, Hoofs & Leseman, 2013; de Haan, Leseman & Elbers, 2011; Slot, 2014), sind als zielgruppenspezifische Angebote konzipiert und hauptsächlich bildungsbenachteiligten Kinder vorbehalten (Barnett et al., 2004; OECD, 2006). Einige Offensiven und Initiativen formen

einen festen Bestandteil der FIBB-Regelversorgung, z. B. VVE in den Niederlanden oder die Bildungsangebote von Head Start in den USA (siehe Kapitel 3.5). Einige Bildungsmaßnahmen in frühkindlichen Einrichtungen beispielsweise das Perry Preschool Programm finden im Rahmen von Multikomponenten-Interventionen statt, welche zusätzliche Unterstützungen für die Familien anbieten wie Hausbesuche durch geschulte Fachkräfte oder Elterntrainings.

Teilweise wird in den Ansätze mit eigenen Bildungsplänen bzw. Curricula oder Konzepten gearbeitet, die theoretische Grundannahmen zum Lernen und zur frühen Bildung enthalten, Handlungsleitlinien für die pädagogische Arbeit bieten und auf zentrale Bildungsinhalte eingehen (z. B. Stiftung Bildungspakt Bayern, 2007; Fthenakis et al., 2009; Greenes, Ginsburg & Balfanz, 2004; Weikart & Schweinhart, 2013). Die Bildungsziele erstrecken sich mehrheitlich auf verschiedene Entwicklungsbereiche (z. B. Mathematik, Naturwissenschaften und Sprache/Literacy), wobei oft eine bereichsübergreifende Förderung innerhalb einzelner Bildungsangebote anvisiert wird (z. B. Fried, et al., 2012; Fthenakis et al., 2009; Roßbach et al., 2010; Weikart & Schweinhart, 2013). Das Ausmaß, in welchem Mathematik im Bildungsangebot berücksichtigt ist, variiert und es gibt weitere Unterschiede in der Ausgestaltung der Ansätze (siehe Clements & Sarama, 2011; Deutscher & Selter, 2013; Ginsburg et al., 2008; Hasemann & Gasteiger, 2014; Hellmich, 2007; Kaufmann, 2011 für einen Einblick in mathematikspezifische Ansätze und Bestandteile sowie Anderson et al., 2003; Barnett, 1995, 2011; Camilli et al., 2010; Chambers, Cheung, Slavin, Smith & Laurenzano, 2010; Melhuish et al., 2015 für Einblicke in generelle Ansätze früher Bildung). Die mathematischen Bildungsinhalte sind im Vergleich zu Förderprogrammen breiter und umschließen beispielsweise auch die Bereiche Größen und Messen oder Daten und Häufigkeiten (Stiftung Bildungspakt Bayern, 2007; Clements & Sarama, 2011; High/Scope, 2008; Gasteiger, 2010, 2012; Greenes et al., 2004).

Erkennbar ist zudem oft ein Bestreben die Anschlussfähigkeit der mathematischen Bildungsprozesse sicherzustellen, indem frühe Kompetenzen angebahnt werden, an die sich die Bildungsinhalte der Primarstufe anschließen (Gasteiger, 2012; Roßbach et al., 2010). Ziel ist allerdings nicht, schulische Bildungsinhalte in die FIBB zu verlagern, sondern den Kindern Angebote und Erfahrungen zu unterbreiten, die dem Entwicklungsniveau und den Bedürfnissen der Kinder sowie dem pädagogischen Alltag in den frühkindlichen Einrichtungen entsprechen (Gasteiger, 2010, 2012; Greenes et al., 2004; Presser et al., 2015).

Statt mit sich wiederholenden, ähnlichen Aufgaben zu arbeiten steht die Vielfalt der mathematischen Grunderfahrungen im Vordergrund (Fthenakis et al., 2009; Gasteiger, 2010, 2012; Greenes et al., 2004; Weikart & Schweinhart, 2013). Teilweise ist ebenfalls ein sequen-

tieller Aufbau einzelner Bildungsblöcke erkennbar, wobei die Sequenzierung weniger rigide, starr und kleinschrittig als in den meisten Förderprogrammen ausfällt.

Die Bildungsmaßnahmen beziehen sich auf Zeiträume von einem Jahr bis zu mehreren Jahren oder sie umfassen den gesamten Kindergartenzeitraum, wobei selten die Förderung von Kindern unter drei Jahren vorgesehen ist. Allgemein sind die Maßnahmen relativ offen gestaltet und formulieren weniger starre Vorgaben für die Bildungsarbeit als die Förderprogramme. So findet die Bildungsarbeit unter flexiblen Rahmenbedingungen statt, z. B. unterschiedliche Gruppengrößen und Zeiten, um die Umsetzbarkeit für eine Vielfalt an pädagogischen Konzeptionen und Rahmenbedingungen zu gewährleisten. Die Implementation vieler Offensiven und Initiativen fand allerdings unter günstigen strukturellen Rahmenbedingungen statt wie kleinere Gruppen oder Erzieher-Kind-Schlüssel und speziell ausgebildete Fachkräfte.

Im Rahmen mancher Offensiven und Initiativen gibt es eigens entwickelte Materialien oder es werden Materialvorschläge unterbreitet, beispielsweise für die mathematische Bildung mit Alltagsgegenständen und -materialien. Hierbei handelt es sich weniger um Anweisungen sondern vielmehr um Unterstützungsangebote an die Fachkraft. So sollen die Anregungsmaterialien von Sommerlatte et al. (2009) oder die Beispiele des Stiftung Bildungspakt Bayern (2007) als Anregung für die Fachkräfte dienen, eigene Ideen für mathematische Bildungsangebote zu entwerfen. Die Verfahren zur Dokumentation von Lernschritten sollen den Fachkräften erleichtern, die Bildungsangebote an die individuellen Bedürfnisse und Interessen der Kinder anzupassen (Stiftung Bildungspakt Bayern, 2007; Steinweg, 2009; vgl. Gasteiger, 2010, 2012). Die Kinder sollen dazu angeregt werden, sich mathematische Inhalte eigenständig zu erschließen und sich aktiv forschend mit mathematischen Herausforderungen auseinanderzusetzen. Obschon teils instruktive Momente vorhanden sind, besteht die Rolle der Fachkraft hauptsächlich darin, die Lernprozesse zu begleiten und zu unterstützen. Mehrere Offensiven und Initiativen beinhalten eigens ausgearbeitete Schulungskonzepte, in welchen die Fachkräfte wichtige Entwicklungsschritte der Mathematik sowie zentrale Inhalte, Methoden und theoretische Grundlagen der mathematischen Bildung erlernen (z. B. Clements & Sarama, 2011; High/Scope, 2008; Gasteiger, 2010, 2012; Presser et al., 2015).

Die vorliegenden Evaluationsstudien zeigen mehrheitlich, dass die Teilnahme an Offensiven und Initiativen mit Entwicklungsfortschritten u. a. in Mathematik über kurze, mittlere und sogar lange Zeiträume zusammenhängt (vgl. Anders, 2013; Blok et al., 2005; Burger, 2010; Camilli et al., 2010; Clements & Sarama, 2011; Gasteiger, 2014; Ginsburg et al., 2008; Melhuish et al., 2015; Nores & Barnett, 2010; Roßbach et al., 2010). Dabei scheinen einzelne Angebote erfolgreich Disparitäten abzubauen: Bildungsbenachteiligte Kinder holten Entwick-

lungsrückstände zumindest teilweise auf. Die Studien weisen meist ein quasi-experimentelles Design auf, wobei ebenfalls mehrere kleinere Studien mit experimentellem Design existieren. Die Stichprobengrößen und Untersuchungszeiträume der Studien variieren (von einigen bis mehreren Jahren, in einigen Fällen sogar Jahrzehnten), fallen aber in der Regel größer als in Evaluationsstudien zu den Förderprogrammen aus. Im Forschungsstrang wurde insbesondere das Potenzial von qualitativ hochwertigen und intensiven frühen Bildungsangeboten in der FIBB hervorgehoben, welche teilweise besonders auf die Bedürfnisse von Kindern und Familien abgestimmt waren. Weil verschiedene Maßnahmen mit spezifischen Curricula, speziell aus- bzw. fortgebildeten Fachkräften und nur mit einer bestimmten Zielgruppe an Kindern arbeiteten sowie teilweise zusätzliche Unterstützung für die Familien anboten und die Wirksamkeit der einzelnen Komponenten der Offensiven und Initiativen nicht getrennt aufschlüsselten, ist eine Übertragung der Befunde auf den Kontext der Regelversorgung nur eingeschränkt möglich (vgl. Kuger, Sechtig & Anders, 2012; Roßbach, Kluczniok & Kuger, 2008).

4.4 Forschungsstrang zur mathematischen Bildung in der FIBB-Regelversorgung

In einem letzten Forschungsstrang wurde sich der Regelversorgung FIBB und ihrer Auswirkung auf die kindliche Entwicklung gewidmet (vgl. Anders, 2013; Burger, 2010; Dalli et al., 2011; Kuger et al., 2012; Melhuish et al., 2015; Roßbach et al., 2008). Nicht nur sammeln Kinder ggfs. andere Erfahrungen im Rahmen eines spezifischen Förderprogramms oder einer Bildungsoffensive bzw. -initiative als im pädagogischen Alltag regulärer Einrichtungen, auch innerhalb der FIBB-Regelversorgung weichen die Betreuungserfahrungen voneinander ab (Anders, 2013; OECD, 2013). Für die Frage nach den Effekten FIBB sind Längsschnittstudien von großer Bedeutung, die untersuchen, wie sich die existierenden Unterschiede in der Regelversorgung auf die kindliche Entwicklung auswirken. Während in älteren Studien dieses Forschungsstrangs ausschließlich die Effekte institutioneller Betreuung im Vergleich zur ausschließlich familialen Betreuung erforscht und dabei ggfs. Unterschiede im Ausmaß der Fremdbetreuung untersucht wurden (z. B. Andersson, 1989, 1992; Broberg, Hwang & Chace, 1993; Broberg, Wessels, Lamb & Hwang, 1997; Wessels, Lamb & Hwang, 1996), bieten neuere Studien ein differenzierteres Bild (vgl. Anders, 2013). In Anbetracht der hohen und wachsenden Betreuungsraten in Europa, repräsentiert Fremdbetreuung eine fast universelle Erfahrung von Kindern und die Frage nach dem Für und Wider einer Fremdbetreuung stellt sich kaum noch (Anders, 2013; OECD, 2013). Daher wird in neueren Studien weniger der Frage nachgegangen, ob sich institutionelle Betreuung per se positiv oder negativ auswirkt, sondern

vielmehr der Frage, unter welchen Bedingungen FIBB eine möglichst positive Wirkung entfaltet (Melhuish, 2011; OECD, 2013).

Zunächst wurden in Studien quantitative Unterschiede FIBB exploriert, beispielsweise Umfang und Intensität oder Stabilität der Betreuung (vgl. Anders, 2013; Burger, 2010; Melhuish et al., 2015). Die Erfahrungen von Kindern unterscheiden sich in weiteren Charakteristika, z. B. Anzahl, Vielfalt und Qualität von Bildungsangeboten, Größe der Gruppen und Klima in den Gruppen sowie Umgang der fröhpädagogischen Fachkräfte mit den Kindern. Der Qualitätsansatz ermöglicht es solch qualitative Unterschiede abzubilden.

Das wohl prominenteste Beispiel für eine Längsschnittstudie dieses Forschungsstrangs, in der sowohl quantitative als auch qualitative Unterschiede erforscht wurden, ist die groß angelegte, prospektive US-Studie des National Institute of Child Health and Human Development des Early Child Care Research Networks (ECCRN, 2002a, 2002b, 2003, 2005; Vandell, Belsky, Burchinal, Vandergrift & Steinberg, 2010). In Europa entstanden mit der Zeit ähnliche längsschnittliche Forschungsprojekte (siehe Anhang B Tabelle 1 für einen Überblick). Zu den einflussreichsten und längsten Studien Europas zählt EPPSE, die 1990 in England begann (Sylva et al., 2014; Sylva et al., 2010). Als deutsche Beispiele lassen sich die 2005 begonnene Studie BiKS 3–10 (Anders et al., 2013; Anders et al., 2012; von Maurice et al., 2007) und die Kindergartenkohorte des 2010 ins Feld geföhrten, bundesweiten Nationalen Bildungspanels anführen (National Educational Panel Study [NEPS]; Bäumer, Aßmann, von Maurice & Blossfeld, 2013; Durda, 2015).

Bei Untersuchungen zur regulären FIBB ist eine Randomisierung der Stichprobe in der Regel nicht möglich. Außerdem werden in Studien normalerweise keine Bildungsangebote verglichen und selten noch Stichproben von Kindern ohne Betreuungserfahrung herangezogen, weshalb kaum Kontroll- oder Vergleichsgruppendedesigns in diesem Forschungsstrang vorliegen (vgl. Anders, 2013; Burger, 2010; Dalli et al., 2011; Melhuish et al., 2015). In den meisten Studien wird jedoch mit komplexen Forschungsdesigns gearbeitet, in welchen verschiedene Merkmale der Kinder und der Familien sowie institutionelle und außerinstitutionelle Bedingungsfaktoren vergleichsweise umfangreich erfasst werden, damit die Analysen zu Effekten FIBB das komplexe Zusammenspiel der Bedingungsfaktoren abbilden können (Burger, 2010; Nores & Barnett, 2010).

Die Befunde dieses Forschungsstrangs sind besonders wertvoll, weil sie ggfs. Implikationen für die flächendeckende Umsetzung des mathematischen Bildungsauftrags und die Gestaltung der FIBB-Regelversorgung bieten, in der sich die Mehrheit der Kinder in Europa befinden (Anders, 2013; OECD, 2013). Als zentrales Forschungsthema kristallisierte sich die pädagogische Qualität heraus insbesondere die Qualität pädagogischer Prozesse, welche zentrale für das TP1 des Promotionsvorhabens zu den Auswirkungen FIBB auf die kindliche Entwicklung ist und auf die in einem eigenen Kapitel genauer eingegangen wird (siehe Kapitel 6.1).

5 Komponenten der Qualität mathematischer Bildung

Im Promotionsvorhaben sollen Modelle pädagogischer Qualität im Hinblick auf die Eignung zur Untersuchung der mathematischen Bildung in der FIBB diskutiert und ergänzt werden (Katz, 1996; Kluczniok & Roßbach, 2014; OECD, 2006; Pianta et al., 2005). Im Qualitätsansatz wird der Frage nachgegangen, was unter guter FIBB zu verstehen ist bzw. was für die Qualität von Angeboten spricht. Der Ansatz setzte wichtige Impulse für die pädagogische und bildungspolitische Praxis: Nach dem erfolgten und weiterhin forcierten quantitativen Ausbau der frühkindlichen Betreuung ist die Sicherstellung und Verbesserung der Qualität der Betreuung das zentrale Thema bildungspolitischer Debatten der FIBB in Deutschland (Larra, 2005; Roux & Tietze, 2007) und weltweit (OECD, 2011, 2013). In der Forschung erlangte der Ansatz weite Verbreitung, weil die Qualitätsmodelle hilfreich sind, um Unterschiede in der Regelversorgung abzubilden und die Auswirkung regulärer FIBB zu untersuchen (siehe Kapitel 4.4; vgl. Anders, 2013; Burger, 2010; Melhuish et al., 2015).

Das Kapitel 5 geht auf zentrale Komponenten der Qualität mathematischer Bildung und ihre Erfassung ein, wobei zentrale Ansatzpunkte zur Ergänzung und weiteren Differenzierung etablierter Modelle aufgezeigt werden. Das Kapitel 5.1 beginnt mit einer Beschreibung der Komponenten etablierter Modelle. Dabei werden Prozesskomponenten und ihre Erfassung eingehender behandelt, weil sie als zentral für die Auswirkung FIBB gelten und einen zentralen Forschungsgegenstand des TP1 darstellen. Im Anschluss wird im Kapitel 5.2 die Diskussion zur mathematischen Bildungseffektivität als ergänzende Qualitätskomponente eingeleitet, welche im TP2 fortgeführt wird. Ebenfalls wird erklärt, wie Effektivitätsanalysen empirisch umgesetzt werden. Zuletzt werden in Kapitel 5.3 in Vorbereitung des TP3 wichtige Ansatzpunkte zur Ergänzung von Qualitätsmodellen aufgezeigt, die sich durch das weitere Differenzieren von Komponenten ergeben, die sich auf die Fachkräfte beziehen.

5.1 Komponenten etablierter Qualitätsmodelle

In Wissenschaft und Praxis haben sich unterschiedliche Modelle pädagogischer Qualität etabliert, welche sich aber im Wesentlichen im Grundverständnis und den Komponenten ähneln (Kluczniok & Roßbach, 2014; Pianta et al., 2005). Diese basieren oftmals auf Expertenmeinungen (Dalli et al., 2011; Melhuish, 2001; Pianta & Hamre, 2009), in die zahlreiche Forschungsbefunde sowie langjährige Praxiserfahrungen eingeflossen sind und die gemäß neuer Anforderungen und Erkenntnisse weiterentwickelt wurden. Bei der Beurteilung der Qualität von Einrichtungen stehen Komponenten im Vordergrund, welche die Förderung der kindli-

chen Entwicklung und das Wohlbefinden von Kindern in der FIBB gewährleisten (Roux & Tietze, 2007). Die einzelnen Komponenten werden in der Regel verschiedenen Dimensionen zugeordnet, wobei eine Unterteilung nach Struktur- und Prozessqualität gängig ist (siehe Abbildung 1; Kluczniok & Roßbach, 2014; Pianta et al., 2005).

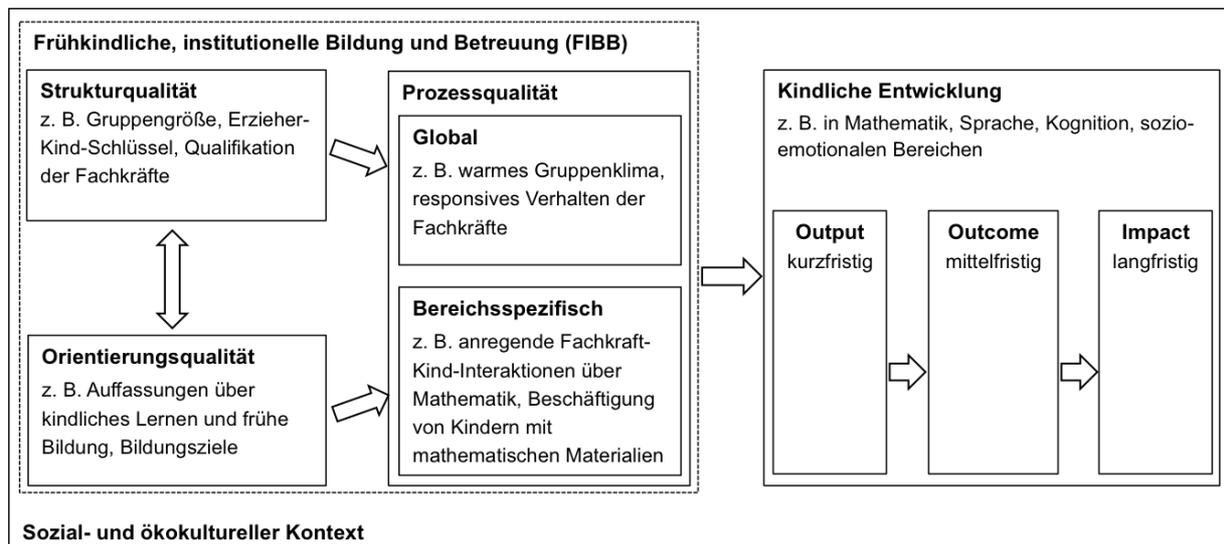


Abbildung 1. Strukturell-prozessuales Modell pädagogischer Qualität (nach Kluczniok & Roßbach, 2014; Roux & Tietze, 2007; Tietze et al., 1998).

Unter *Strukturqualität* werden die strukturellen Rahmenbedingungen FIBB zusammengefasst wie die Einrichtungs- oder Gruppengröße, der Erzieher-Kind-Schlüssel oder die Qualifikation der Fachkräfte. Strukturkomponenten sind relativ stabil und politisch regulierbar (Kuger & Kluczniok, 2008). Unter *Prozessqualität* wird das pädagogische Geschehen in den Einrichtungen verstanden z. B. die Interaktionen der Kinder mit den frühpädagogischen Fachkräften, der räumlich-materiellen Umgebung und anderen Kindern (Kluczniok & Roßbach, 2014; La Paro, Thomason, Lower, Kintner-Duffy & Cassidy, 2012; Pianta et al., 2005; Tietze, Schuster, Grenner & Roßbach, 2007). Darüber hinaus gibt es in manchen Modellen weitere Dimensionen, beispielsweise die *Orientierungsqualität*, welche sich in Deutschland weitgehend durchgesetzt hat und die Auffassungen, Überzeugungen, Ziele und Wertvorstellungen in den Einrichtungen umfasst (Kluczniok & Roßbach, 2014; Roux & Tietze, 2007; Tietze, 2008). Hierzu zählen zum einen die offiziellen, schriftlich dokumentierten Orientierungen, z. B. in den Bildungsplänen oder Einrichtungskonzeptionen (Anders, Ballaschk & Tietze, 2014; Roux & Tietze, 2007; Tietze et al., 2012), zum anderen die Orientierungen der Fachkräfte (Anders et al., 2014; Kluczniok, Anders & Ebert, 2011; Kluczniok & Roßbach, 2014).

Dem gestiegenen Bewusstsein für die Bereichsspezifität von Lern- und Anregungsprozessen tragen Qualitätsmodelle in Ansätzen Rechnung (Roßbach et al., 2008): Beispielsweise bezog sich die Prozessqualität anfangs ausschließlich auf Komponenten *globaler Prozessqualität*, welche bereichsübergreifend für die Entwicklung und das Wohlbefinden von Kindern wichtig sind wie die Responsivität der Fachkräfte sowie ein warmes Klima und ein Angebot von allgemein anregenden und entwicklungsangemessenen Aktivitäten in den Gruppen (Kluczniok & Roßbach, 2014; La Paro et al., 2012; Pianta et al., 2005; Tietze et al., 2007). Zunehmend wurden Komponenten *bereichsspezifischer Prozessqualität* in die Beurteilung einbezogen, die das Ausmaß und die Art der Förderung in unterschiedlichen Entwicklungsbereichen wie Mathematik, Sprache/Literacy oder Naturwissenschaften charakterisieren (Halle, Vick Whittaker & Anderson, 2010; Kuger & Kluczniok, 2008; Sylva et al., 2006).

Gemäß strukturell-prozessualer Wirkannahme kommt der Prozessqualität eine Schlüssel-funktion zu, weil von einem direkten Zusammenhang zur kindlichen Entwicklung ausgegangen wird (ECCRN, 2002a; Kuger & Kluczniok, 2008; Mashburn et al., 2008; Pianta et al., 2005). Für die Komponenten der Struktur- und Orientierungsqualität wird ein über die Prozessqualität vermittelter Zusammenhang angenommen. Struktur- und Orientierungsqualität stellen somit Bedingungsfaktoren der Prozessqualität dar, die als Schaltstelle für die Wirkung anderer Dimensionen fungiert (Kuger & Kluczniok, 2008). Um Aussagen zur Auswirkung der Qualität treffen zu können wird im TP1 des Promotionsvorhabens daher auf die Prozesskomponenten fokussiert.

Während Komponenten der Struktur- und Orientierungsqualität zumeist über vergleichsweise wenig aufwändige Verfahren wie Befragung, Interview und Dokumentenanalyse zugänglich sind, basieren die Verfahren zur Erfassung der Prozessqualität in der Regel auf direkten Beobachtungen (Sylva et al., 2006; Tietze, 2008; siehe Tabelle 2 in Anhang B; vgl. Halle et al., 2010; Kluczniok & Roßbach, 2014). Weit verbreitete *Beobachtungsverfahren der globalen Prozessqualität* sind die Caregiver Interaction Scale (CIS; Arnett, 1989), das Classroom Assessment Scoring System (CLASS; Pianta, La Paro & Hamre, 2008) oder die Early Childhood Environment Rating Scale-Revised Edition (ECERS-R; Harms, Clifford & Cryer, 1998) bzw. übersetzte Versionen der Skala wie die Kindergarten-Skala Revidierte Fassung (KES-R; Tietze et al., 2007). Ausgewählte Verfahren enthalten Prozesskomponenten, welche die kognitive Förderung (z. B. die Dimension Instructional Support der CLASS-Pre-K, insbesondere die Subskala Concept Development), teilweise sogar spezifisch im mathematischen Bereich abbilden (z. B. die Aktivitätenskala der ECERS-R bzw. KES-R, insbesondere das Item zum mathematischen Verständnis; Kluczniok & Roßbach, 2014).

Beispiele für Beobachtungsverfahren der bereichsspezifischen Prozessqualität sind die Early Childhood Environment Rating Scale-Extension (ECERS-E; Sylva, Siraj-Blatchford & Taggart, 2003) und übersetzte Versionen der Skala wie die deutsche Kindergarten-Skala-Erweiterung (KES-E; Tietze, Schlecht & Wellner, 2005) oder die Dortmunder Ratingskala zur Erfassung sprach-, mathematisch- und naturwissenschaftsförderrelevanter Interaktionen – Erweiterung (DO-RESI-E-Ki; Isele, 2014). Die Verfahren bilden die Förderung in verschiedenen Bereichen mit Subskalen ab, so auch in Mathematik. Darüber hinaus existieren Skalen, die spezifisch für einen Bildungsbereich konzipiert wurden, z. B. die Classroom Observation of Early Mathematics Environment and Teaching (COEMET; Sarama & Clements, 2009) für Mathematik, oder die Early Language and Literacy Classroom Observation (ELLCO; Smith, Dickinson, Sangeorge & Anastasopoulos, 2002) für den Bereich Sprache/Literacy oder die für die Studie Utrecht Mixed Preschool Groups entwickelten Skalen zu fachkraftinitiierten Aktivitäten in Literacy und Mathematik (Teacher-manged literacy and math activities; de Haan et al., 2013). Obschon im Allgemeinen betont wird, dass eine reliable und valide Erfassung der Prozessqualität Beobachtungen voraussetzt (Halle et al., 2010; Mashburn et al., 2008; Pianta & Hamre, 2009; Sylva et al., 2006), wurden in wenigen Studien selbstkonstruierte *Fragebögen bereichsspezifischer Prozessqualität* eingesetzt, mit denen die frühpädagogischen Fachkräfte das Ausmaß der Förderung in verschiedenen Bereichen einschätzten (Bäumer et al., 2013; Silinskas, Pakarinen, Poikkeus, Lerkkanen & Nurmi, 2015; Slot, 2014; Slot, Mulder, Verhaagen & Leseman, 2014) oder Selbstevaluationsskalen, welche für das Monitoring und die Feldsteuerung entwickelt wurden wie die Skalen der schottischen Aufsichtsbehörde Social Care and Social Work Improvement Scotland (SCSWIS) oder die Qualitätsindikatoren des Education Scotland Instituts (QI; Bradshaw, Lewis & Hughes, 2014).

Die vorgestellten Verfahren sind gruppenbezogen konzipiert, d. h. sie beziehen sich auf das interaktive Geschehen in der Gruppe, während andere Verfahren die pädagogischen Prozesse auf Individualebene betrachten, z. B. die Aktivitäten, die frühpädagogische Fachkräfte mit einem spezifischen Kind durchführen (Durda, 2015; vgl. Kluczniok & Roßbach, 2014). Viele Angaben beinhalten Aussagen zu Qualitätsstandards, mit welchen Skalenwerten bestimmte Qualitätsstufen zugeordnet werden. So verwenden die CLASS, ECERS-R/-E bzw. KES-E/-R und DO-RESI-E-Ki eine Skalierung von eins bis sieben, wobei Werte unter drei für unzureichende Prozessqualität stehen, sowie Werte zwischen drei und fünf mittelmäßige und Werte von fünf bis sieben gute bis ausgezeichnete Qualität kennzeichnen.

5.2 Bildungseffektivität als Ergänzung für etablierte Qualitätsmodelle

In der Schulforschung wird seit einiger Zeit national (Bischof, Hochweber, Hartig & Klieme, 2013; Sälzer, 2010) und international (Chapman, Muijs, Reynolds, Sammons & Teddlie, 2015) der Ansatz der Bildungseffektivität verwendet um das durchschnittliche Niveau der Kompetenzen bzw. den Grad an Kompetenzentwicklung in Schulen zu vergleichen (Creemers, Kyriakides & Sammons, 2010; Reynolds et al., 2014; Sylva et al., 2010). Bislang ist EPPSE die einzige Studie, in welcher der Effektivitätsansatz in der FIBB verwendet wurde (vgl. Melhuish, 2001; Sylva et al., 2010). Im Effektivitätsansatz steht die Produktionsfunktion von Bildungseinrichtungen im Fokus, d. h. wie in der FIBB mit einem bestimmten Input (u. a. strukturelle und ideelle Bedingungen) pädagogische Prozesse generiert werden, welche gewünschte Bildungsergebnisse produzieren (z. B. frühe Kompetenzen in Mathematik; ECCRN, 2002a; OECD, 2001; Roux & Tietze, 2007; vgl. Input-Prozess-Output-Modelle zur Schulqualität; Ditton, 2000; Scheerens & Bosker, 1997). Die Effektivität bezieht sich auf die erzielten Entwicklungsfortschritte in Mathematik am Ende der Betreuungszeit (*den Bildungsoutput*), wobei letztlich überdauernde positive Effekte auf die weitere akademische und berufliche Laufbahn sowie den generellen Lebenserfolg anvisiert werden (Outcomes; Anders, 2013; Melhuish et al., 2015; OECD, 2013). Prinzipiell steht eine Beurteilung der Bildungseffektivität von Einrichtungen im Einklang mit dem gängigen Qualitätsverständnis von FIBB, in dem die Förderung der kindlichen Entwicklung im Zentrum steht (Roux & Tietze, 2007). Das Erreichen von Lernzielen bzw. Kompetenzzuwächse wird vereinzelt auch explizit als Qualitätsdimension aufgelistet (vgl. OECD, 2006; Viernickel & Schwarz, 2009).

Die Idee einer Beurteilung FIBB anhand des Bildungsoutputs ist mit einem steuerungsstrategischen Paradigmenwechsel verbunden (vgl. Pant et al., 2013): In Ansätzen ist in einigen Ländern eine Verlagerung von einer Steuerung des Inputs FIBB (z. B. ökonomische Ressourcen, detaillierte Curricula oder Vorschriften struktureller Merkmale) hin zur einer stärkeren Outputsteuerung (z. B. durch ein kontinuierliches Monitoring des Bildungsoutputs in Mathematik und weiteren Entwicklungsbereichen) erkennbar (Network on Early Childhood Education and Care, 2013; OECD, 2015b). Der Outputsteuerung FIBB sowie der Erfassung von Bildungsergebnissen generell werden in Deutschland allerdings große Vorbehalte entgegengebracht (Moser & Studer, 2013), wo nach wie vor eine inputorientierte Steuerungsstrategie dominiert, welche allmählich durch eine zunehmende Prozesssteuerung ergänzt wurde (Larra, 2005; Moser & Studer, 2013; OECD, 2015b; Roux & Tietze, 2007).

Im TP2 des Promotionsvorhabens wird sich daher eingehender mit der mathematischen Bildungseffektivität als zusätzliche Komponente der Qualität mathematischer Bildung auseinandergesetzt um den Bildungsoutput frühkindlicher Einrichtungen zu beurteilen (siehe Abbildung 2 für eine Illustration).

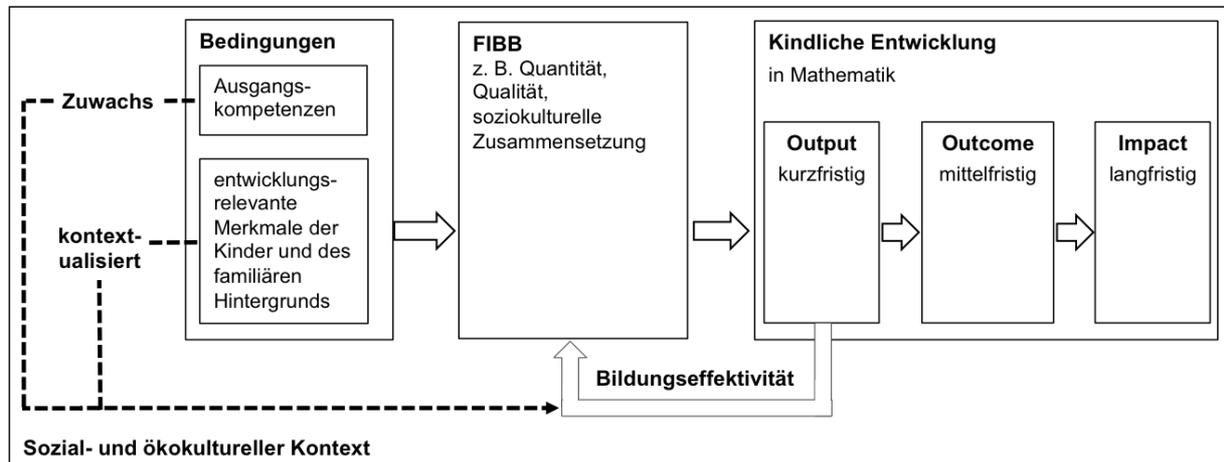


Abbildung 2. Effektivitätsvergleiche als Qualitätskomponente zur Beurteilung des mathematischen Bildungsoutputs FIBB.

Für *Effektivitätsanalysen* werden Angaben über kindliche Kompetenzen in frühkindlichen Einrichtungen benötigt (Creemers et al., 2010; Melhuish, 2001; Racherbäumer, Funke, van Ackeren & Clausen, 2013). In Mathematik sind dies in der Regel Angaben über mathematische Kompetenzen aus Testverfahren wie die Skala Early Number Concepts der British Ability Scales Second Edition, welche in den mathematischen Effektivitätsanalysen in EPPSE eingesetzt wurden (BASII; Elliot, Smith & McCulloch, 1996; vgl. Sammons et al., 2002; Sylva et al., 2010). Effektivitätsurteile ergeben sich nicht durch einen Vergleich des Kompetenzniveaus bzw. -zuwachs mit vorab definierten Bildungsstandards (vgl. Pant et al., 2013) oder mit den Lernzielen in elementarpädagogischen Bildungsplänen (siehe Kapitel 3.2), sondern durch einen Vergleich mit dem Bildungsoutput in einer Referenzstichprobe anderer Einrichtungen. Eine Einrichtung gilt als effektiv, wenn der Output in einer Einrichtung höher ausfällt als in anderen Einrichtungen (Creemers et al., 2010; Racherbäumer et al., 2013; Reynolds et al., 2014).

Effektivitätsanalysen basieren auf Mehrebenenanalysen und somit auf der hierarchischen Struktur der FIBB (siehe Kapitel 3.4). Es gibt unterschiedliche Varianten für Analysen (Creemers et al., 2010; Sammons et al., 2002). Statt eines einfachen (a) *Niveauvergleichs* der durchschnittlichen Kompetenz in Einrichtungen ist es für die „Fairness von Effektivitätsver-

gleichen“ (Scheerens, 1992) ggfs. wichtig, in (b) *Zuwachsvergleichen* oder sogenannten Value Added Modellen (VAM; Lenkeit, 2013; Racherbäumer et al., 2013) das Ausgangsniveau der Kompetenzen zu berücksichtigen. Zudem fließen ggfs. in sogenannten (c) *kontextualisierten Niveauvergleichen* und (d) *Zuwachsvergleichen* entwicklungsrelevante Hintergrundmerkmale ein, also verschiedene Merkmale der Kinder und des familialen Hintergrunds, welche sich als bedeutsam für die frühe mathematische Entwicklung erwiesen haben (CAM für Contextualized Attainment Models bzw. CVA für Contextualized Value-Added Models; Sammons et al., 2013; Sylva et al., 2014). Ausgangskompetenzen und entwicklungsrelevante Hintergrundmerkmale werden berücksichtigt, um den Anteil an der Kompetenz bzw. am Kompetenzzuwachs zu vergleichen, der Effekte der FIBB widerspiegelt und nicht ausschließlich durch individuelle oder familiale Faktoren erklärbar ist und so auf eine effektive Bildungsarbeit oder weitere förderrelevante Eigenschaften von Einrichtungen schließen lässt.

5.3 Ausdifferenzierung fachkraftbezogener Qualitätskomponenten in Mathematik

Frühpädagogische Fachkräfte sind Schlüsselakteure der frühkindlichen Bildung und an die mathematische Bildungsarbeit frühpädagogischer Fachkräfte werden hohe Erwartungen gestellt (siehe Kapitel 3.3). Umfang und Qualität der Bildung und damit die Bildungsergebnisse FIBB hängen maßgeblich von der Umsetzung des Bildungsauftrags durch die Fachkräfte in der täglichen pädagogischen Arbeit ab. Allerdings deuten die Ergebnisse verschiedener Studien an, dass Fachkräfte generell einen geringen Anteil des pädagogischen Alltags für die mathematische Förderung nutzen und z. B. Sprachförderaktivitäten den Vorrang geben (vgl. Cabell, DeCoster, LoCasale-Crouch, Hamre & Pianta, 2013; Ginsburg et al., 2008). Zudem gab es große Unterschiede in der mathematischen Bildungsarbeit von Fachkräften, wobei insgesamt eine Beschränkung auf bestimmte Teilbereiche der Mathematik erkennbar war (Boonen, Kolkman & Kroesbergen, 2011; Ginsburg et al., 2008; Klibanoff, Levine, Huttenlocher, Vasilyeva & Hedges, 2006; McCray & Chen, 2012). Forscher und Forscherinnen kritisieren zudem, dass Fachkräfte in den USA mathematische Bildungsgelegenheiten bzw. Teachable Moments nur unzureichend nutzen (Ginsburg et al., 2008; Hyun & Marshall, 2003). Darüber hinaus deuten Prozessmaße, welche auf das pädagogische Handeln und die Interaktionen der Fachkräfte fokussieren (z. B. CLASS) eine unzureichende kognitive Förderung durch die Fachkräfte an (siehe Tabelle 2 im Anhang B). Lehl et al. (2014) fanden für die BiKS-Studie in Deutschland, dass Fachkräfte von geringer Förderaktivität in Mathematik berichteten, wenn auch häufiger als in Literacy/Sprache. Zudem zeichneten sich ebenfalls eine Konzentration auf

gewisse mathematische Teilbereiche und Unterschiede im Förderausmaß und den Förderaktivitäten der Fachkräfte ab.

Eine ausreichende Qualifizierung und Professionalisierung der frühpädagogischen Fachkräfte in Aus- und Fortbildung ist eine essentielle Voraussetzung dafür, dass Fachkräfte die notwendigen professionellen Kompetenzen erwerben, um eine qualitativ hochwertige und effektive Bildungsarbeit in Mathematik umzusetzen (vgl. Anders, 2013). Die Voraussetzungen der frühpädagogischen Fachkräfte gelten als zentrale Ansatzpunkte zur Optimierung der Qualität und des Bildungoutputs in Mathematik (siehe Abbildung 3). Daher zählen Komponenten der Struktur- und Orientierungsqualität, welche sich auf die Fachkräfte beziehen, zu den proximalen Bedingungsfaktoren von Bildungsprozessen in der FIBB (Cryer et al., 1999; Pianta et al., 2005; vgl. Bronfenbrenner & Morris, 2006).

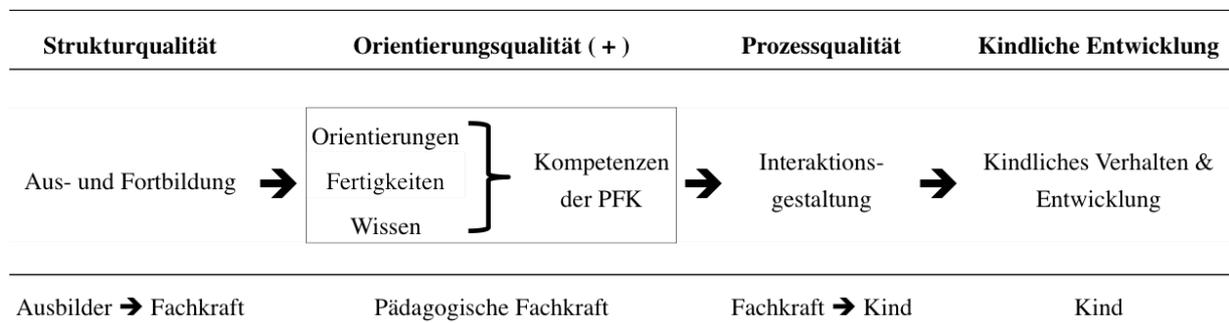


Abbildung 3. Fachkraftbezogene Komponenten der Qualität und ihre Auswirkung (angepasst nach Fukkink & Lont, 2007).

In den meisten pädagogischen Qualitätsmodellen sind *fachkraftbezogene Komponenten der Qualität* jedoch unzureichend repräsentiert. So gehören zwar Aspekte der Qualifikation und Professionalisierung von frühpädagogischen Fachkräften zur Strukturqualität und die Überzeugungen, Wertvorstellungen sowie Einstellungen der Fachkräfte zu den Komponenten der Orientierungsqualität (Kluczniok & Roßbach, 2014; OECD, 2006; Pianta et al., 2005; Roux & Tietze, 2007), es fehlen jedoch insbesondere Vorschläge für eine Differenzierung und Operationalisierung mathematikspezifischer Komponenten.

Mittlerweile liegen differenzierte Vorschläge und wertvolle Erkenntnisse der Professions-, Professionalisierungs- und Interaktionsforschung vor, wobei die unzureichende spezifische Professionalisierung als eine Hauptursache für eine mangelhafte Umsetzung des Bildungsauftrags durch frühpädagogische Fachkräfte diskutiert wird (Anders, 2012; Aktionsrat Bildung, 2012; Chen & Chang, 2006; Chen & McCray, 2012; Cross et al., 2009; Gasteiger, 2017). Zudem liegen wichtige Erkenntnisse und Vorschläge zur Operationalisierung mathema-

tikspezifischer Orientierungen von Fachkräften (z. B. Anders & Roßbach, 2015; Benz, 2012a, 2012b; Chen, McCray, Adams & Leow, 2014; Thiel, 2010) aus der wissenschaftlichen Auseinandersetzung mit den professionellen Kompetenzen frühpädagogischer Fachkräfte vor (z. B. Anders, 2012; Aktionsrat Bildung, 2012; Dunekacke et al., 2013; Gasteiger, 2010, 2014; Morawiak, Schulz, Jungmann & Koch, 2017; Schuler, Streit & Wittmann, 2017).

Die Beiträge anderer Forschungsrichtungen finden allerdings in der Forschung zur Qualität FIBB zu wenig Berücksichtigung. Lange Zeit beschränkte sich die öffentliche und wissenschaftliche Debatte zur Qualitätsentwicklung in der FIBB auf ein Anheben des *allgemeinen Qualifikationsniveaus* der frühpädagogischen Fachkräfte (z. B. durch einen akademischen Abschluss ggfs. mit frühkindlicher Spezialisierung). Zunehmend findet der Ausbau des Angebots *bereichsspezifischer Professionalisierung in der Aus- und Fortbildung* der Fachkräfte und weiteren Bereichen Einzug in die Debatten (Anders, 2012; Aktionsrat Bildung, 2012; Cross et al., 2009; EACEA/Eurydice, 2009; Early et al., 2006; Early et al., 2007; Fukkink & Lont, 2007; Hardy & Steffensky, 2014; OECD, 2006, 2011, 2013, 2015b; Pianta et al., 2005; Whitebook, 2003). In Ansätzen werden auch bereichsspezifische Ausdifferenzierungen ebenfalls für die fachkraftbezogenen Komponenten der Orientierungsqualität erkennbar, z. B. werden frühpädagogische Fachkräfte vermehrt zur Bedeutung verschiedener Entwicklungsbereiche befragt (Anders et al., 2014). Dennoch werden die *mathematikspezifischen Fachkraftorientierungen* in der Forschung zur Qualität FIBB kaum einbezogen.

Insgesamt erscheinen die Beiträge anderer Forschungsrichtungen bislang unzureichend in Qualitätsmodellen berücksichtigt, womit die Konzeptualisierung und Operationalisierung mathematikspezifischer Komponenten in der Qualitätsforschung noch vergleichsweise unscharf ausfällt. Im TP3 des Promotionsvorhabens wird sich eingehender mit zentralen Ansatzpunkten zur Ausdifferenzierung fachkraftbezogener Struktur- und Orientierungskomponenten beschäftigt. Insbesondere werden die mathematikspezifische Professionalisierung sowie Orientierungen von Fachkräften in der FIBB betrachtet, wofür auf Modelle, Erkenntnisse und Messverfahren anderer Forschungsrichtungen zurückgegriffen wird (u. a. der Professions- und Professionalisierungsforschung sowie der Motivations- und Schulforschung).

6 Evidenz zu Qualitätskomponenten und Auswirkungen in der FIBB

Mit dem Promotionsvorhaben wird eine evidenzbasierte Diskussion und Ergänzung der etablierten Qualitätsmodelle angestrebt. Ausgangspunkt bilden Überlegungen zu wichtigen Forschungsdesiderata, die sich aus einer Sichtung des Forschungsstands zur Qualität mathematischer Bildung in der FIBB-Regelversorgung ableiten lassen. Kapitel 6 gibt daher einen Überblick über die vorliegende Evidenz zu den im Kapitel 5 thematisierten Komponenten der Qualität mathematischer Bildung und deren Auswirkungen. Bei der Schilderung werden entsprechende Schwerpunkte gesetzt und es wird ggfs. exemplarisch auf ausgewählte Befunde eingegangen, um unzureichend geklärte Forschungsfragen zu veranschaulichen, die leitend für die TP des Promotionsvorhabens waren. Im Kapitel 6.1 werden bestehende Inkonsistenzen in der Befundlage zur Auswirkung der Prozessqualität auf kindliche Bildungsergebnisse in Mathematik am Beispiel skizziert. Zudem wird auf die damit verbundenen Fragen verwiesen, die Anlass zur Metaanalyse des TP1 gaben. Wichtige Ergebnisse zur mathematischen Bildungseffektivität der EPPSE-Studie werden in Kapitel 6.2 zusammengetragen, um den potentiellen Erkenntnisgewinn aufzuzeigen, der zu den mathematischen Effektivitätsanalysen der FIBB in Deutschland des TP2 motivierte. In Kapitel 6.3 werden die Schwerpunkte der Forschung zu fachkraftbezogenen Qualitätskomponenten aufgezeigt um auf wichtige Forschungslücken hinzuweisen, welche Anlass zur theoretischen und empirischen Auseinandersetzung mit mathematikspezifischen Komponenten im TP3 gaben.

6.1 Befundlage zur Prozessqualität und ihrer Auswirkungen auf Bildungsergebnisse

Mittlerweile existieren international viele Untersuchungen zu den Zusammenhängen von pädagogischer Qualität und kindlicher Entwicklung (vgl. Anders, 2013; Dalli et al., 2011; Melhuish et al., 2015), in denen, entsprechend strukturell-prozessuaalem Qualitätsverständnis (siehe Kapitel 5.1), Effekte der Prozessqualität untersucht und zunehmend sowohl globale als auch bereichsspezifische Prozesskomponenten betrachtet werden.

In Anbetracht der besonderen Entwicklungsrelevanz erscheinen die erreichten Durchschnittswerte der Prozessqualität in der FIBB problematisch: International ergaben sich in den meisten Ländern mittlere Werte globaler Prozessqualität und niedrige Werte bereichsspezifischer Prozessqualität, wobei die Werte in Mathematik besonders niedrig ausfielen (Clifford, Reszka & Roßbach, 2010; Kuger & Kluczniok, 2008; Slot, Lerkkanen & Leseman, 2015; Vermeer, van IJendoorn, Cárcamo & Harrison, 2016). Dies zeigte sich so auch für die 103 deutschen Kindergärten der Studie European Child Care and Education (ECCE; Krumm et al.,

1997; Tietze, Hundertmark-Mayser & Roßbach, 1999), in der Nationalen Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit (NUBBEK; Tietze et al., 2012) und für die 97 hessischen und bayerischen Einrichtungen der BiKS-Studie (siehe Tabelle 2 im Anhang B für Details; Anders et al., 2012; Kuger & Kluczniok, 2008).

Die Mehrheit der Befunde zur Auswirkung der Prozessqualität in der FIBB deutet auf positive Effekte der Prozessqualität auf Bildungsergebnisse in Mathematik hin (vgl. Anders, 2013; Burger, 2010; Dalli et al., 2011; Melhuish et al., 2015). Dennoch fallen relevante Inkonsistenzen in der Befundlage auf, welche im Folgenden exemplarisch beschrieben werden (siehe Tabelle 4 im Anhang B für weitere Einzelbefunde):

In EPPSE wurde berichtet, dass eine höhere bereichsspezifische Prozessqualität (gemessen mit der ECERS-E) die mathematischen Leistungen von fast 3 000 englischen Kindern über eine Altersspanne von fünf bis 16 Jahren vorhersagt (Sammons et al., 2014; Sammons et al., 2004; Sammons et al., 2007; Sylva et al., 2006), während sich keine signifikanten Effekte globaler Prozessqualität erhoben mit der ECERS-R ergaben. Für BiKS wurden signifikante Zusammenhänge für beide Prozesskomponenten zur mathematischen Entwicklung von über 450 deutschen Kindern berichtet, die mit KES-R, KES-E oder Subskalen erfasst wurden, allerdings ausschließlich im Zeitraum von drei bis fünf Jahren (Anders et al., 2012). Für den weiteren Entwicklungsverlauf in der Grundschule berichteten Anders et al. (2013) und Lehl et al. (2016) nur bereichsspezifische Prozesseffekte. Abreu-Lima et al. (2013) fanden, dass weder die globale noch die bereichsspezifische Prozessqualität beobachtet mit ECERS-R und ELLCO mit mathematischen Bildungsergebnissen von 215 portugiesischen Kindern im Alter von vier bis fünf Jahren im Zusammenhang standen. Für die fast 700 sechsjährigen Kinder der EPPNI-Studie aus Nordirland zeigten sich unterschiedliche Befunde für die globale Prozessqualität in Abhängigkeit vom verwendeten Prozessmaß (verschiedene Subskalen der CIS).

Insgesamt unterschieden sich die Befunde einzelner Studien teilweise und es ergeben sich auch abweichende Befunde innerhalb von Studien. So zeichnen sich Unterschiede für globale und bereichsspezifische Prozesskomponenten ab, aber auch für ähnliche Prozesskomponenten wurden in Studien teils unterschiedliche Resultate erzielt. Unklar ist demnach, von welchen durchschnittlichen Prozesseffekten ausgegangen werden kann und ob bereichsspezifische und globale Prozessqualität gleichermaßen eine Rolle für die mathematische Entwicklung von Kindern spielen. Zudem stellt sich die Frage, wodurch abweichende Befunde innerhalb und zwischen Studien erklärt werden können. Obschon sich die Studien in ihren Fragestellungen und zugrundeliegenden Forschungsdesigns ähneln, gibt es zusätzlich zu regionalen Unterschieden und den damit einhergehenden Unterschieden im System und der pädagogischen

Traditionen FIBB weitere Abweichungen im Design der Studien. Welche Rolle z. B. die Wahl des Prozessmaßes, das Alter der Kinder oder eine methodisch-statistische unterschiedliche Umsetzung der Fragestellung für Befunde zu Prozesseffekten spielt, ist noch nicht hinreichend geklärt.

Vor dem Hintergrund der Evidenz zur Bereichsspezifität von Entwicklungs- und Anreizungsprozessen sind zudem Unterschiede prozessualer Effekte zwischen Entwicklungsbereichen anzunehmen. In den meisten Längsschnittstudien wurde die Auswirkung FIBB in diversen Entwicklungsbereichen wie Sprache/Literacy, sozio-emotionale Bereiche und die Entwicklung allgemein kognitiver Fähigkeiten erforscht. Gerade der Bildungsbereich Sprache/Literacy hat in den letzten Jahren enormes öffentliches und wissenschaftliches Interesse gebündelt und die frühen sprachlichen Kompetenzen gelten wie in Mathematik prädiktiv für spätere Schulleistungen (vgl. Roßbach & Weinert, 2008).

Einige Beispiele machen deutlich, dass es auch innerhalb der kognitiv-leistungsbezogenen Domäne fraglich scheint, dass Befunde auf verschiedene Entwicklungsbereiche generalisierbar sind, weshalb eine bereichsspezifische Erforschung der Prozesseffekte als wichtig erachtet wird. So deuteten sich in EPPSE wichtige Unterschiede prozessualer Effekte in den Bereichen Mathematik und Sprache/Literacy an, aber die Befunde weisen über die Entwicklungszeiträume hinweg keine klaren Muster auf (Sammons et al., 2014; Sammons et al., 2011; Sylva et al., 2006). In BiKS resultierten für die Wortschatzentwicklung anders als für die mathematische Entwicklung bereits im Zeitraum von drei bis fünf Jahren keine bereichsspezifischen Prozesseffekte für die ECERS-E Subskala Lesen (Ebert et al., 2013; Weinert, Ebert, Lockl & Kuger, 2012). Abreu-Lima et al. (2013), die in ihrer Studie keinen Nachweis für mathematische Bildungsergebnisse fanden, berichteten von Prozesseffekten im Bereich Sprache/Literacy (insbesondere für das globale Prozessmaß ECERS-R und teilweise für das bereichsspezifische Prozessmaß ELLCO). Insgesamt wirft die Befundlage zur Rolle der Prozessqualität für die mathematische Entwicklung demnach mehrere Fragen auf, denen im TP1 des Promotionsvorhabens nachgegangen wird.

6.2 Evidenz zur mathematischen Bildungseffektivität frühkindlicher Einrichtungen

In EPPSE wurde nicht nur die Prozessqualität sondern ebenfalls der Bildungsausgang in Mathematik von 141 Fremdbetreuungsangeboten (v. a. FIBB, aber auch Spielgruppen) aus fünf Regionen Englands mit Effektivitätsanalysen untersucht (vgl. Melhuish, 2001; Sammons et al., 2014; Sammons et al., 2002; Siraj-Blatchford, Sammons, Taggart, Sylva & Melhuish, 2006; Sylva et al., 2010). Die Effektivitätsanalysen basierten auf Angaben der BAS-II (Elliot et al., 1996) zur mathematischen Kompetenz von fast 3 000 Kindern im Alter von fünf Jahren. Als Ausgangsniveau dienten die Angaben der Skala zur kognitiven und verbalen Kompetenz im Alter von drei Jahren. Die Ergebnisse wiesen in allen Analysen auf Unterschiede in der mathematischen Bildungseffektivität der untersuchten Einrichtungen hin. Die ermittelte Varianz zwischen Einrichtungen sank jedoch von 17 % in den Vergleichen zum Kompetenzniveau (ohne Vergleich der Ausgangskompetenzen oder entwicklungsrelevanter Merkmale) auf 6 % in den kontextualisierten Vergleichen zum Zuwachs, was zeigt, dass es in der FIBB ebenfalls für die Fairness von Effektivitätsvergleichen wichtig ist (vgl. Scheerens, 1992), das Ausgangsniveau der Kompetenz und entwicklungsrelevante Merkmale zu berücksichtigen (Sammons et al., 2002).

Der Besuch einer effektiveren Einrichtung sagte in EPPSE die mathematische Entwicklung über die Grundschulzeit und ein besseres Abschneiden bei den Abschlussprüfungen in Mathematik der weiterführenden Schule vorher (Sammons et al., 2011; Sylva et al., 2014; Sylva et al., 2010). Die Ergebnisse der Effektivitätsanalysen boten den Ausgangspunkt für vertiefende, qualitative Analysen in einer kleinen Substichprobe aus sehr und wenig effektiven Einrichtungen, für die sich Unterschiede im Interaktionsstil und im pädagogischen Handeln der frühpädagogischen Fachkräfte andeuteten (Siraj-Blatchford et al., 2006). Damit erbrachten die Analysen Erkenntnisse über pädagogische Prozesse in der FIBB, welche über Befunde aus den etablierten Prozessmaßen hinausgingen. In der Zusammenschau zeigen die Ergebnisse in EPPSE den Erkenntniswert von Effektivitätsanalysen für die Forschung zur Qualität mathematischer Bildung in der FIBB bei einem komplementären Einsatz von Effektivitäts- und Prozessqualitätsanalysen sowie ggfs. zusätzlichen vertiefenden, qualitativen Verfahren (Melhuish, 2001; Sammons et al., 2014; Sylva et al., 2010). In TP2 des Promotionsvorhabens soll daher der Frage nachgegangen werden, ob Effektivitätsanalysen für die Untersuchung mathematischer Bildung in der deutschen FIBB nützlich sind.

6.3 Evidenz zu fachkraftbezogenen Qualitätskomponenten

Bei einer Sichtung der Evidenz zu den fachkraftbezogenen Qualitätskomponenten werden wichtige Forschungslücken ersichtlich, insbesondere zu den spezifischen Komponenten in Mathematik. Erkenntnisse zum allgemeinen Qualifikationsniveau und zur Berufserfahrung von frühpädagogischen Fachkräften und ihrer Rolle für die Prozessqualität und die mathematischen Bildungsergebnisse in der FIBB liegen auch für die Regelversorgung vor (z. B. Brown, Molfese & Molfese, 2008; Burchinal, Cryer, Clifford & Howes, 2002; Early et al., 2006; Fukkink & Lont, 2007; Ghazvini & Mullis, 2002; Phillipsen, Burchinal, Howes & Cryer, 1997; Saracho & Spodek, 2007; Scarr, Eisenberg & Deater-Deckard, 1994; Slot, Lerkkanen, et al., 2015; Whitebook, 2003). Ebenfalls scheinen diese globalen Strukturkomponenten mit den Orientierungen von Fachkräften im Zusammenhang zu stehen, weil Fachkräfte mit verschiedenen Abschlüssen und verschiedener Berufserfahrung sich in Studien hinsichtlich ihrer Orientierungen unterschieden (z. B. Cobanoglu, 2011; McMullen, 1997; Mischo, Wahl, Hendler & Strohmer, 2012; Mischo, Wahl, Strohmer & Hendler, 2012; Snider & Fu, 1990; Tirosh et al., 2011; Wilkins, 2008; siehe aber Guo, Justice, Sawyer & Tompkins, 2011; Vartuli, 1999). Kritisch bezüglich der Aussagekraft für die deutsche FIBB ist anzumerken, dass viele Belege aus Ländern mit anderen FIBB-Systemen stammen, in denen die Fachkräfte einen anderen Professionalisierungshintergrund und mehr Fachkräfte akademische Abschlüsse besitzen.

Die Mehrheit der Belege für die Wirkung mathematikspezifischer Aus- und Fortbildungsangebote auf die Qualität pädagogischer Arbeit, Bildungsergebnisse und Fachkraftorientierungen stammen aus Evaluationsstudien zu Fachkräftetrainings sowie Qualitätsoffensiven und –programmen (z. B. Bercovici, 2015; Clements & Sarama, 2011; Evans, Leonard, Krier & Ryan, 2013; Fried, et al., 2012; Fukkink & Lont, 2007; Gasteiger, 2010, 2012, 2014; Ginsburg et al., 2008; Guarino, Hamilton, Lockwood & Rathbun, 2006; Linder, 2012; Sarama & Clements, 2009). Über die Situation, Implementation und Wirkung mathematikspezifischer Professionalisierungsangebote in der regulären FIBB in Deutschland gibt diese Evidenz nur unzureichend Aufschluss (vgl. Baumeister & Grieser, 2011; Beher & Walter, 2012; Cross et al., 2009; Deppe, 2011; Gasteiger, 2017; Ginsburg et al., 2008; Janssen, 2010; Ledig, 2011; Sarama & Clements, 2009).

Für Fachkraftorientierungen und ihre Bedeutung für die Prozessqualität und Bildungsergebnisse lassen sich zwar Befunde anführen (z. B. Aslan, Ogul & Tas, 2013; z. B. Brown, 2005; Chen et al., 2014; Kluczniok et al., 2011; Kuger & Kluczniok, 2008; Lee & Ginsburg, 2007; Mischo, Wahl, Hendler, et al., 2012; Mischo, Wahl, Strohmer, et al., 2012; Pianta et al., 2005). Zusätzlich bieten einzelne Erhebungen von Fachkraftorientierungen ebenfalls wertvolle

Einsichten in die Situation der Orientierungsqualität FIBB für den mathematischen Bereich (Chen et al., 2014; Copley & Padrón, 1999; Gasteiger, 2010; Ginsburg et al., 2008; Sparr, Chen & McCray, 2011), jedoch gibt es nur wenige Untersuchungen zu den mathematischen Orientierungen deutscher frühpädagogischer Fachkräfte (Anders & Roßbach, 2015; Benz, 2012a, 2012b; Fried, et al., 2012; Oppermann, Anders & Hachfeld, 2016; Thiel, 2009). Insgesamt ergibt sich ein Nachholbedarf für die wissenschaftliche Auseinandersetzung mit den fachkraftbezogenen Qualitätskomponenten der FIBB-Regelversorgung in Deutschland, weil die Befundlage zu den spezifischen Professionalisierungsangeboten und Orientierungen in Mathematik unzureichend scheint. Die fachkraftbezogenen Qualitätskomponenten stehen daher im Zentrum des TP3.

7 Anliegen des Promotionsvorhabens

Anliegen des Promotionsvorhabens ist es, etablierte Modelle und die empirische Evidenz zu den Komponenten und Auswirkungen der Qualität mathematischer Bildung in der FIBB zu diskutieren und zu ergänzen. In verschiedenen Forschungssträngen wurde das mathematische Bildungspotenzial frühkindlicher Einrichtungen dokumentiert und wertvolle Förderansätze wurden unterbreitet (siehe Kapitel 4.2 bis 4.4). Mit Hilfe von Qualitätsmodellen und -forschung gelang es, Wege für die flächendeckende Überprüfung und Optimierung mathematischer Bildung in der FIBB-Regelversorgung aufzuzeigen und so erheblich zum Verständnis und zur Weiterentwicklung mathematischer Lern- und Bildungsprozesse in der FIBB beizutragen. Dennoch ergeben sich in der Gesamtschau zentrale Desiderata, welche Anlass zur theoretischen und empirischen Auseinandersetzung mit den Komponenten und Auswirkungen der Qualität mathematischer Bildung in frühkindlichen Einrichtungen in drei TP gaben. Um die Desiderata zu adressieren, ist es erforderlich, die Evidenz verschiedener Studien zur FIBB in Europa zusammenzuführen. Wie die Evidenz unterschiedlicher Studien gebündelt werden kann, ist eine methodische Frage mit genereller Bedeutung für die frühkindliche Bildungsforschung. Im Folgenden geht Kapitel 7.1 auf die Desiderata ein, die ausschlaggebend für die TP und das methodische Anliegen des Promotionsvorhabens waren. Anschließend werden im Kapitel 7.2 die sich daraus ergebenden Ziele der TP und des Promotionsvorhabens erläutert.

7.1 Die zugrundeliegenden Desiderata

Erstens ergeben sich Desiderata im Zusammenhang mit der empirischen Evidenz zu den Auswirkungen der Prozessqualität auf die mathematische Entwicklung. Generell sprechen die Ergebnisse von Längsschnittstudien der FIBB in Europa für die angenommene Entwicklungsrelevanz der Prozessqualität (siehe Kapitel 5.1 und 6.1). Dennoch schwanken die Befunde und prozessuale Effekte lassen sich nicht für jedes Alter und für alle Prozessmaße nachweisen. Somit ist nach wie vor unklar, von welcher Stärke der Prozesseffekte durchschnittlich ausgegangen werden kann und ob globale und bereichsspezifische Komponenten der Prozessqualität gleichermaßen relevant für Bildungsergebnisse sind. Inwiefern die Befunde spezifisch für den mathematischen Bereich oder auf weitere schulleistungsbezogene Bereiche (z. B. Sprache/Literacy) generalisierbar sind, ist noch nicht abschließend zu beantworten. Mittlerweile stehen Forschern und Forscherinnen verschiedene Prozessmaße (siehe Kapitel 5.1) und Herangehensweisen zur methodisch-statistischen Umsetzung der gleichen Fragestellungen zur Verfügung. Welche Rolle die Wahl des Prozessmaßes und der methodisch-analytischen Her-

angehensweise für die Befunde zu prozessualen Effekten spielt, ist ungeklärt. Existierende Reviews oder Metaanalysen zur Auswirkung FIBB (z. B. Anders, 2013; Blok et al., 2005; Burger, 2010; Camilli et al., 2010; Dalli et al., 2011; Karoly, Kilburn & Cannon, 2005; Manning, Homel & Smith, 2010; Melhuish et al., 2015; Nores & Barnett, 2010) liefern weder Angaben über die durchschnittliche Stärke prozessualer Effekte noch ausreichende Erklärungen für die Variation der Befunde (siehe Kapitel 8.1).

Zweitens bestehen Desiderata zur mathematischen Bildungseffektivität FIBB. In Anbetracht der Verbreitung in der nationalen und internationalen Schulforschung sowie der vielversprechenden Ergebnisse der EPPSE-Studie für den frühkindlichen Bereich in England (siehe Kapitel 5.2 und 6.2) stellt sich die Frage, warum bislang keine Effektivitätsanalysen für die deutsche FIBB durchgeführt wurden. Zu klären ist daher, ob der Effektivitätsansatz prinzipiell in Deutschland anwendbar ist oder ob überzeugende Argumente gegen eine Verwendung sprechen. Zusätzlich ist unklar, ob sich die mathematische Effektivität deutscher Einrichtungen unterscheidet. Bislang fehlen eine Diskussion sowie empirische Überprüfung des Ansatzes für FIBB in Deutschland.

Drittens bestehen Desiderata zu den fachkraftbezogenen Komponenten der Struktur- und Orientierungsqualität mathematischer Bildung (siehe Kapitel 5.3 und 6.3). Während in der Professionalisierungsforschung sowie in Evaluationsstudien zu Fachkrafttrainings und Bildungsoffensiven auf die Bedeutung mathematikspezifischer Professionalisierungsangebote für die Qualität pädagogischer Bildungsarbeit und -prozesse hinweisen wird, ist wenig zum mathematikspezifischen Professionalisierungshintergrund von Fachkräften der FIBB-Regelversorgung in Deutschland bekannt. Die Dimension der Orientierungsqualität ist generell unterforscht und die Konzeptualisierung und Operationalisierung vergleichsweise unscharf. Im Besonderen fehlen Kenntnisse über Fachkraftorientierungen im mathematischen Bereich.

Viertens resultiert die methodische Frage, wie die zusammenhängenden Fragestellungen der Desiderata mit geeigneter Evidenz zu adressieren sind. Obschon verschiedene Studien relevante Evidenz beinhalten, erscheint deren Aussagekraft nicht ausreichend, um alle Desiderata zu erforschen. In der Bildungsforschung ist eine allgemeine Tendenz erkennbar, Daten einer breiteren (wissenschaftlichen) Öffentlichkeit zur Verfügung zu stellen und Re- bzw. Sekundäranalysen voranzutreiben. Einerseits ist das Ziel, die Potenziale von Studien möglichst vollständig auszuschöpfen sowie Befunde und deren Generalisierbarkeit überprüfbar zu machen. Andererseits besteht die Möglichkeit, durch die Bündelung der Evidenz unterschiedlicher Studien die Limitierung einzelner Studien zu überwinden (z. B. durch Querschnittsdes-

igns oder die Begrenzung auf bestimmte Altersspannen, Bildungsbereiche sowie Länder und FIBB-Systeme) sowie neuartige und komplexe Analysen durchzuführen. Im Zuge dieser Bemühungen haben Forschungsvorhaben einen großen Mehrwert, in denen sowohl der Erkenntniswert der Evidenzbündelung als auch die Herausforderungen und Grenzen am empirischen Beispiel aufgezeigt werden.

7.2 Die Ziele des Dissertationsvorhabens und der Teilprojekte (TP)

Inhaltliches Ziel der Arbeit ist es, in drei TP zentrale Forschungsfragen zu den Komponenten und Auswirkungen der Qualität mathematischer Bildung in frühkindlichen Einrichtungen zu adressieren, um etablierte Modelle und die vorliegende Evidenz zu diskutieren und zu ergänzen (siehe Abbildung 4 für eine Übersicht). Ziel ist es ebenfalls aufzuzeigen, wie ein vielfältiges, methodisch-statistisches Vorgehen die Evidenz unterschiedlicher Studien bündeln kann, um zusammenhängende Fragestellungen der frühkindlichen Bildungsforschung zu beantworten.

Im *TP1* wird eine Bündelung der Evidenz europäischer Längsschnittstudien zu den durchschnittlichen Effekten globaler und bereichsspezifischer Prozessqualität anvisiert, um Aufschlüsse über die Auswirkungen der FIBB-Regelversorgung auf Bildungsergebnisse in Mathematik zu erlangen, die existierende Forschungssynthesen nicht bieten. In diesem Zusammenhang soll ebenfalls geklärt werden, ob die Effekte spezifisch für den Bereich Mathematik ausfallen (im Vergleich zum Bereich Sprache/Literacy) und welche Rolle das gewählte Prozessmaß und methodisch-analytische Vorgehen der Studien für die Befunde zu Prozesseffekten spielen.

Im *TP2* soll die Anwendbarkeit des Effektivitätsansatzes für den Vergleich des mathematischen Bildungsausgangs frühkindlicher Einrichtungen in Deutschland diskutiert und überprüft werden. Hierfür soll geklärt werden, ob sich Herausforderungen für den Qualitätsansatz ergeben, die eine ergänzende Beurteilung der mathematischen Bildung in der FIBB notwendig erscheinen lassen. Demgegenüber sollen die Argumente abgewogen werden, die gegen eine ergänzende Beurteilung der mathematischen Effektivität in der deutschen FIBB angeführt werden können. Weiterhin sollen die Voraussetzungen an Studiendesigns für mathematische Effektivitätsanalysen dargestellt werden, um zu überprüfen, ob Studien zur deutschen FIBB hierfür in Frage kommen. Zusätzlich ist das Ziel, die mathematische Bildungseffektivität frühkindlicher Einrichtungen aus der BiKS-Studie zu vergleichen und die Ergebnisse vor dem Hintergrund der existierenden BiKS-Befunde zur Prozessqualität und weiterer, relevanter Befunde zur FIBB zu diskutieren.

Im *TP3* wird angestrebt, wichtige Ansatzpunkte zur Ausdifferenzierung fachkraftbezogener Komponenten der Struktur- und Orientierungsqualität in Mathematik durch Beiträge anderer Forschungsrichtungen aufzuzeigen (u. a. frühpädagogische und schulische Professions- und Professionalisierungsforschung sowie Motivationspsychologie). Auf Grundlage der Beiträge sollen in einer explorativen Studie zusätzlich zur generellen Qualifikation die Professionalisierung und Schulerfahrung in Mathematik deutscher Fachkräfte der FIBB-Regelversorgung als Strukturkomponenten untersucht werden. Darüber hinaus werden die Wertorientierung, die Selbstwirksamkeit und das Selbstkonzept in Mathematik als Komponenten der Orientierungsqualität betrachtet. Zusätzlich zu einem deskriptiven Bild der fachkraftbezogenen Qualitätskomponenten für die deutsche FIBB-Regelversorgung sollen die Zusammenhänge zwischen den Struktur- und Orientierungskomponenten erforscht werden. Die Befunde sollen mit der Evidenz für Fachkräfte in Deutschland und den USA und aus Evaluationsstudien von Fachkräftetrainings und Bildungsoffensiven verglichen werden, wobei die Spezifika der Arbeit und Professionalisierung in der FIBB-Regelversorgung beleuchtet werden sollen.

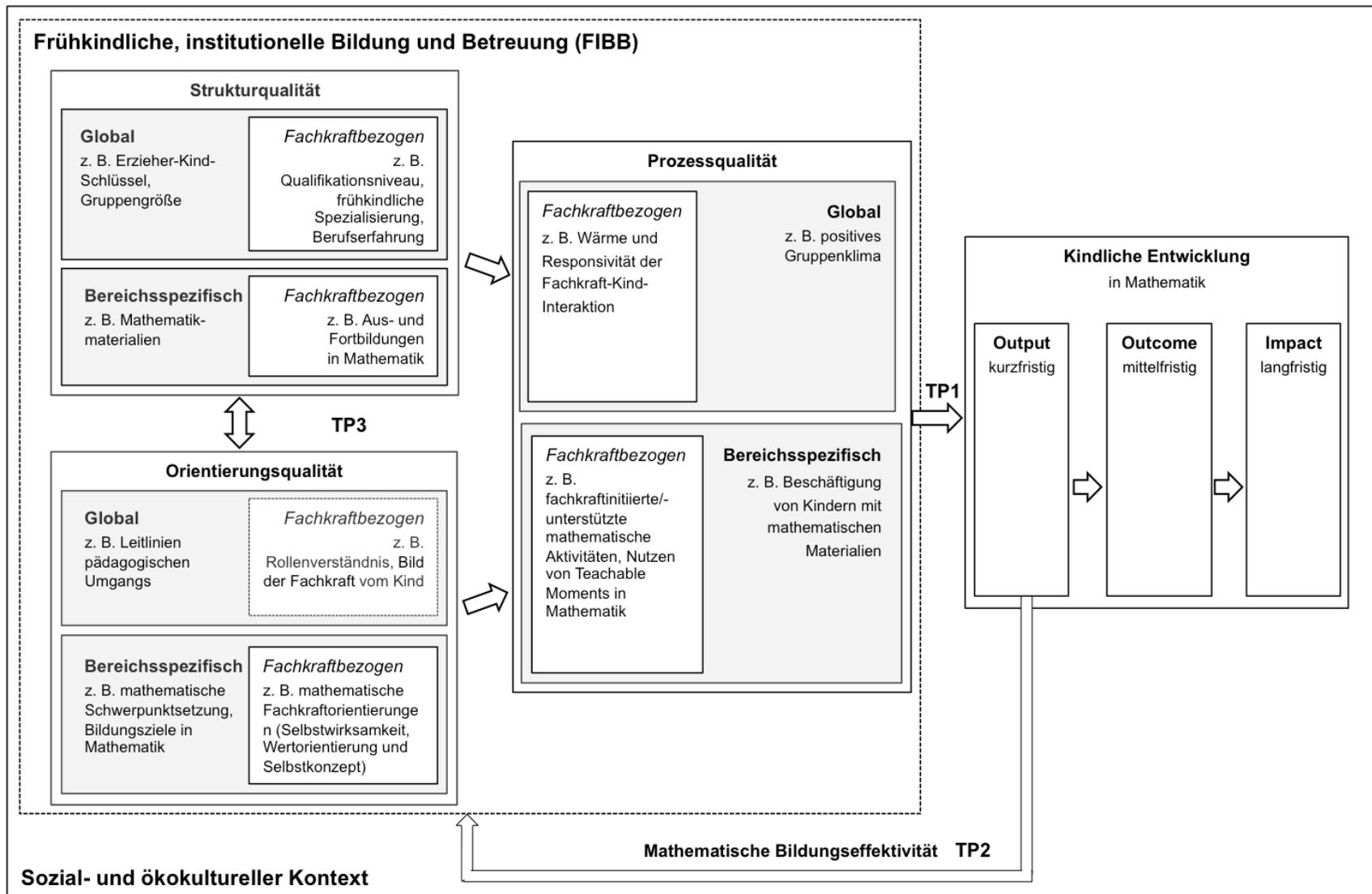


Abbildung 4. In der Arbeit aufgezeigte, ausgewählte Ansatzpunkte zur Differenzierung und Erweiterung des strukturell-prozessualen Qualitätsmodell (angepasst nach Kluczniok & Roßbach, 2014; Roux & Tietze, 2007; Tietze et al., 1998).

Als *methodisches Anliegen* soll demonstriert werden, wie mit einer Vielfalt an methodisch-statistischem Vorgehen die empirische Evidenz unterschiedlicher Studien zusammengeführt werden kann, um zusammenhängende Fragestellungen zu adressieren. Tabelle 1 bietet einen Überblick über den Forschungsgegenstand und die empirische Grundlage jedes TP. In der Metaanalyse in *TP1* wurde die Evidenz aus 17 Längsschnittstudien, neun europäischen Ländern, 2 887 Einrichtungen bzw. Gruppen und 16 461 Kindern zusammengefasst. Grundlage boten die Analysen zu den Effekten von Prozess- und Strukturqualität sowie Quantität FIBB, welche im EU-Projekt Curriculum and Quality Analysis and Impact Review of European Early Childhood Education and Care (CARE) in Form eines Berichts an die Europäische Kommission angefertigt wurden (Ulferts & Anders, 2015). Die mathematischen Effektivitätsanalysen im *TP2* basieren auf den Angaben der Längsschnittstudie BiKS zu den mathematischen Kompetenzen und entwicklungsrelevanten Merkmalen von 544 Kindern aus 97 Einrichtungen in Bayern und Hessen. Im *TP3* werden die Angaben zu fachkraftbezogenen Qualitätskomponenten von 221 frühpädagogischen Fachkräften aus 29 Einrichtungen in Berlin und Bayern der Querschnittstudie Muster, Formen, Raum und Zahlen im Kindergarten entdecken – Pädagogische Einstellungen von frühpädagogischen Fachkräften zu mathematischer Frühförderung analysiert. Damit werden in den TP unterschiedliche Ebenen betrachtet: die Studienebene (TP1), die Einrichtungsebene (TP2) und die Fachkräfteebene (TP3). Zur Anwendung kommen eine Metaanalyse auf Basis von Längsschnittmehrebenenmetaanalyse und Mixed-Effects-Moderatoranalysen (TP1), Mehrebenenregressionen (TP2) und blockweise Regressionen (TP3). Die folgenden drei Kapitel bieten für die drei TP jeweils eine Zusammenfassung des spezifischen theoretischen Hintergrunds sowie eine kurze Beschreibung der Studien einschließlich Fragestellungen, Methode sowie der für das Promotionsthema relevanten Ergebnisse und deren Diskussion.

Tabelle 1

Übersicht über die Teilprojekte (TP), das methodische Vorgehen und die Datengrundlagen des Promotionsvorhabens

	TP1	TP2	TP3
Forschungsgegenstand			
Qualitätsdimension	Prozessqualität	Outputqualität	Fachkraftbezogene Struktur- und Orientierungsqualität
Qualitätskomponenten	Globale und bereichsspezifische Prozesskomponenten	Mathematische Bildungseffektivität	Allgemeine Qualifikation sowie Professionalisierung, Schulerfahrungen und motivationale Orientierungen in Mathematik
Fokus der Fragestellung	Auswirkungen von Qualität	Ergänzung von Qualitätskomponenten	Ergänzung von Qualitätskomponenten
Empirische Grundlage			
Datengrundlage	Europäische Studien (u. a. BiKS, EPPSE, pre-COOL)	BiKS-3-10	Muster, Formen, Raum und Zahlen im Kindergarten entdecken –Pädagogische Einstellungen von frühpädagogischen Fachkräfte zu mathematischer Frühförderung
Ebene	Studien	Einrichtungen	Fachkräfte
Stichprobengröße	17 Studien 2 887 Einrichtungen/Gruppen 16 461 Kinder	1 Studie 97 Einrichtungen 544 Kinder	1 Studie 29 Einrichtungen 221 Fachkräfte
Region	9 Europäische Länder	Bayern und Hessen	Berlin und Bayern
Studiendesign	Längsschnitt	Längsschnitt	Querschnitt
Analyse	Längsschnittmehrebenenmetaanalyse Mixed-Effects-Moderatoranalysen	Mehrebenenregressionen Random-Intercepts-Modelle	Blockweise Regressionen mit robusten Standardfehlern

8 Auswirkungen der Prozessqualität FIBB auf Bildungsergebnisse in Mathematik und Sprache/Literacy – Eine Metaanalyse europäischer Längsschnittstudien (TP1)

Im TP1 sollen die Auswirkungen globaler und bereichsspezifischer Prozesskomponenten der Qualität auf kindliche Bildungsergebnisse untersucht werden (siehe Abbildung 1 und Kapitel 5.1). Hauptmotivation für das TP1 lieferten die vielversprechende, aber heterogene Befundlage europäischer Längsschnittstudien zu prozessualen Effekten auf die mathematische Entwicklung von Kindern der FIBB-Regelversorgung und Abweichungen zu Befunden in Sprache/Literacy als weiteren leistungsbezogenen Bereich sowie die bislang unzureichenden Erklärungen für variierende Befunde (siehe Kapitel 6.1). Kapitel 8 führt zunächst den Hintergrund des TP1 weiter aus. Im Kapitel 8.1 wird erörtert, warum existierende Forschungssynthesen nicht ausreichen, um Fragen im Zusammenhang mit den prozessualen Effekten zu beantworten. Im Anschluss wird in Kapitel 8.2 darauf eingegangen, welche Unterschiede in den untersuchten Bildungsergebnissen und Prozessmaßen variierende Prozesseffekte erklären könnten und daher als potentielle Moderatoren in der Metaanalyse untersucht werden. Kapitel 8.3 und 8.4 beschreiben die Metaanalyse des TP1 und die methodische Umsetzung. In Kapitel 8.5 und 8.6 werden die zentralen Ergebnisse der Studie zusammengefasst und diskutiert. Die Beschreibung des TP1 beschränkt sich auf die für das Promotionsthema relevanten Teile der Metaanalyse. Eine detaillierte Darstellung des TP1 kann dem Anhang A entnommen werden und ist ebenfalls in Ulferts und Anders (2015) nachzulesen.

8.1 Erkenntnisse und Limitierungen bisheriger Forschungssynthesen

Ergebnisse existierende Metaanalysen deuteten zwar auf positive Effekte der FIBB für kognitiv-leistungsbezogene Entwicklungsbereiche hin, haben bislang jedoch keine Prozesseffekte untersucht und sind durch weitere Limitierungen in der Aussagekraft für das Promotionsthema begrenzt (z. B. Blok et al., 2005; Camilli et al., 2010; Nores & Barnett, 2010). Zuerst wurden in den Metaanalysen die hauptsächlich aus dem US-amerikanischen Raum stammende Evidenz zu Effekten des Besuchs frühkindlicher Einrichtungen per se oder der Teilnahme an Offensiven und Initiativen wie Head Start, High/Scope Perry Preschool oder Abecedarian ausgewertet. Die Evidenz basiert auf Gruppenvergleichen der Leistungen von Kindern mit und ohne Fremdbetreuungserfahrung bzw. verschiedener Bildungsmaßnahmen. Üblicherweise dienen standardisierte Mittelwertdifferenzen als Effektstärkemaß wie Cohens *d* oder Hedges *g* (Card, 2012; Cooper, 2009). Die aggregierten Effekte beinhalten hauptsächlich

Informationen über die Effektivität spezifischer Maßnahmen oder von Multikomponenten-Interventionen, die zusätzliche Unterstützung für Familien anboten (siehe Kapitel 3.5). Zahlreiche Befunde aus Studien der FIBB-Regelversorgung, in welchen keine Vergleiche von Maßnahmen erfolgten oder keine Stichproben von Kindern ohne Fremdbetreuungserfahrungen herangezogen wurden, blieben unberücksichtigt. Auch wegen der hohen Betreuungsrate betrifft dies insbesondere europäische Studien. Die Generalisierbarkeit bisheriger metaanalytischer Befunde auf die Wirkung der FIBB-Regelversorgung in Europa ist somit fraglich und der reine Effekt FIBB bleibt unbekannt.

Wegen der Bereichsspezifität von Entwicklungs- und Anregungsprozessen erscheint zudem das Aggregieren von Befunden aus sehr unterschiedlichen Entwicklungsbereichen problematisch. Die gebündelte Evidenz der Metaanalysen reichte bisweilen von Auswirkungen auf die Intelligenz und allgemein kognitiven Fähigkeiten von Kindern bis zur Wirkung auf verschiedene Bildungsergebnisse. Ein Teil der Varianz in den kognitiven Effekten bisheriger Metaanalysen könnte daher auf die Bündelung bereichsübergreifender Evidenz zurückgehen. Die Konzentration auf Entwicklungsbereiche, welche explizit zum Bildungsauftrag der meisten Ländern zählen (OECD, 2006; Kapitel 3.3 für Mathematik; Sylva et al., 2015) und einen zentralen Stellenwert für spätere akademische Leistung wie Mathematik und Sprache/Literacy haben (siehe Kapitel 2.1; Hellmich, 2007; Kaufmann, 2011; Roßbach & Weinert, 2008), erscheint sinnvoll (Burger, 2010).

Internationale Reviews fassen zwar die Evidenz zur FIBB-Regelversorgung zusammen und betonen, dass gerade für kognitiv-leistungsbezogene Entwicklungsbereiche überzeugende Evidenz für die Auswirkung globaler und bereichsspezifischer Prozessqualität vorliegt (Anders, 2012; Dalli et al., 2011; Melhuish et al., 2015). Die Reviews verweisen jedoch ebenfalls auf die Inkonsistenzen in der Befundlage (siehe Kapitel 6.1). Metaanalysen, in welchen einzelne Befunde nicht nur inhaltlich zusammengefasst werden, sondern quantitativ zu prozessualen Gesamteffekten aggregiert und potentielle Erklärungen für variierende Befunde statistisch getestet werden, fehlen bislang.

8.2 Herleitung der Moderatoren der Prozesseffekte

Aus den Erkenntnissen bisheriger Synthesen sowie der frühkindlichen Bildungsforschung im Allgemeinen lassen sich potentielle Moderatoren der Prozesseffekte mit Bezug zu den untersuchten Bildungsergebnissen und Prozessmaßen ableiten.

8.2.1 Moderatoren mit Bezug zu den Bildungsergebnissen

Zum einen gilt es die *Persistenz prozessualer Effekte* zu klären. Obschon sich besonders in kognitiv-leistungsbezogenen Bereichen überdauernde Entwicklungsvorteile FIBB in Forschungssynthesen andeuten (Anders, 2013; Barnett, 2011; Blok et al., 2005; Camilli et al., 2010; Nores & Barnett, 2010), wird in Metaanalysen von einer Abnahme prozessualer Effekte mit der Zeit berichtet. Metaanalytische Erkenntnisse zur Persistenz prozessualer Effekte existieren bislang nicht. Denkbar ist einerseits, dass sich Effekte mit zunehmendem Alter und zeitlichem Abstand zur FIBB graduell auswaschen (ein sogenannter fade-out Effekt), andererseits, dass die Effekte deutlich abnehmen, nachdem Kinder die frühkindlichen Einrichtungen verlassen (vgl. Camilli et al., 2010) und neuen institutionellen Einflüssen in den Grundschulen und weiterführenden Schulen ausgesetzt sind (z. B. Anders et al., 2013; Lehrl et al., 2016; Sylva et al., 2010).

Darüber hinaus ist es wichtig zu untersuchen, ob es sich um *Effekte auf die kindliche Entwicklung* handelt und Alternativerklärungen auszuschließen. Alternativ können soziale Segregationsprozesse, welche für viele Länder Europas auch bei universellem Zugang zu FIBB dokumentiert sind (siehe Kapitel 3.5; vgl. EACEA/Eurydice, 2009; Leseman & Slot, 2014; OECD, 2006), dazu führen, dass in Studien Zusammenhänge zwischen Prozessqualität und Bildungsergebnissen beobachtet werden. So wäre ein Zusammenhang auch beobachtbar im Falle selektiver herkunftsbedingter Zuweisung zur FIBB-Qualität, welche Kinder mit Migrationshintergrund und niedrigem SES benachteiligt, die ebenfalls in der Regel häufiger Kompetenzrückstände aufzeigen (Anders, 2013; Anders et al., 2012; Arnold & Doctoroff, 2003; EACEA/Eurydice, 2009; Ebert et al., 2013; Melhuish et al., 2015; siehe Kapitel 2.3). Eine weitere Erklärung läge darin, dass es in Gruppen mit vielen benachteiligten Kindern schwieriger zu sein scheint, Prozesse hoher Qualität zu gewährleisten. Zum Beispiel hing der Anteil von Kindern mit Migrationshintergrund negativ mit der globalen und bereichsspezifischen Prozessqualität zusammen (Kuger & Kluczniok, 2008; Kuger, Kluczniok, Kaplan & Roßbach, 2016; Tietze et al., 2012; siehe allerdings Pakarinen, Lerkkanen, Poikkeus, Siekkinen & Nurmi, 2011 für gegenläufige Ergebnisse). Längsschnittstudien bieten die Möglichkeit Zusammenhänge zum Kompetenzzuwachs bzw. zur Kompetenzentwicklung von Kindern zu untersuchen, z. B. in autoregressiven Analysen (z. B. Anders et al., 2013; Sylva et al., 2010) oder Wachstumsmodellen (z. B. Anders et al., 2013; Anders et al., 2012; de Haan et al., 2013) und dadurch teilweise für Selektions- bzw. Segregationseffekte zu kontrollieren.

Darüber hinaus stellt sich die Frage nach der *Bereichsspezifität von Prozesseffekten* aufgrund der abweichenden Befunde für die Bereiche Mathematik und Sprache/Literacy (siehe

Kapitel 6.1) und der Bereichsspezifität von Entwicklungs- und Anregungsprozessen (siehe Kapitel 2.2).

8.2.2 Moderatoren mit Bezug zu den Prozessmaßen

Trotz großer Ähnlichkeiten gibt es wichtige Unterschiede zwischen Prozessmaßen, z. B. welche *Bedeutung der materiell-räumlichen Umgebung* für das pädagogische Geschehen in Settings FIBB beigemessen wird. So beinhalten die Qualitätsskalen ECERS-R und ECERS-E bzw. KES-R und KES-E Einschätzungen des Materials, des Mobiliars sowie zum Einhalten von Sicherheits- und Hygienestandards. Höhere Qualitätswerte sind oft nur erreichbar, wenn die materiell-räumlichen Voraussetzungen erfüllt sind. Hofer (2010) zeigte empirisch, wie sehr diese Ratingregel Qualitätsurteile beeinflusst. Andere Skalen wie die CLASS oder CIS fokussieren beim Rating auf die Interaktionen, insbesondere die Fachkraft-Kind-Interaktionen. Pianta et al. (2005) berichten von lediglich niedrigen bis mittleren Zusammenhängen zwischen CLASS und ECERS-R, wobei die Zusammenhänge zu räumlich-materiellen Items der ECERS-R geringer als zu interaktionsbezogenen Items ausfielen. Von der Äquivalenz der Messverfahren ist daher nicht auszugehen.

Derzeit gibt es verschiedene Forschungsprojekte, in welchen ökonomischere Alternativen zu Beobachtungsmaßen erprobt werden, z. B. Fragebögen für Fachkräfte zur bereichsspezifischen Förderung (Bäumer et al., 2013; Bradshaw et al., 2014; Slot, 2014). Die *Vergleichbarkeit von Fragebogen- und Beobachtungsmaßen der bereichsspezifischen Prozessqualität* ist jedoch nicht ausreichend geklärt.

8.3 Beschreibung der Metaanalyse zu Prozesseffekten des TP1

In Anbetracht der Limitierungen bisheriger Synthesen, ist das Ziel der Metaanalyse des TP1 eine quantitative Synthese der Befunde europäischer Längsschnittstudien für zwei wichtige schulleistungsbezogene Bildungsbereiche: Mathematik und Sprache/Literacy. Angestrebt werden Aussagen über prozessuale Effekte, weil diese auf unterschiedliche Länder, Systeme FIBB und pädagogische Ansätze übertragbar sind (Kluczniok & Roßbach, 2014; Pianta et al., 2005), im Gegensatz zu Aussagen zur Effektivität spezifischer Offensiven und Initiativen. Zum einen soll die durchschnittliche (a) *Höhe der Gesamteffekte globaler und bereichsspezifischer Prozessqualität* der FIBB-Regelversorgung ermittelt werden, wobei von positiven prozessualen Effekten ausgegangen werden kann (Anders, 2013; Dalli et al., 2011; Melhuish et al., 2015). Zum anderen soll das Ausmaß der (b) *Variation der Prozesseffekte zwischen und innerhalb Studien* quantifiziert werden (z. B. für verschiedene Messzeitpunkte oder Teilkompe-

tenzen der Bereiche). In Anbetracht der heterogenen Befundlage ist eine substantielle Varianz anzunehmen.

Darüber hinaus wird untersucht, ob ausgewählte (c) *Moderatoren mit Bezug zu den untersuchten Bildungsergebnissen und Prozessmaßen* die Varianz der Prozesseffekte erklären können. Mit Bezug zu Bildungsergebnissen als Moderatoren wird getestet, ob die Effekte mit zunehmendem Alter der Kinder oder nach dem Verlassen FIBB abnehmen, ob es sich um Effekte auf den Kompetenzzuwachs bzw. die Entwicklung handelt sowie ob sich die Prozesseffekte für Mathematik und Sprache/Literacy unterscheiden. Bezüglich der Prozessmaße wird den Fragen nachgegangen, ob sich die globalen Prozesseffekte für interaktionsfokussierte Maße zu denen mit zusätzlichem Umgebungsrating unterscheiden und ob vergleichbare bereichsspezifische Prozesseffekte mit Fragebogen- und Beobachtungsmaßen erzielt werden.

8.4 Methode der Metaanalyse des TP1

8.4.1 Suche, Auswahl und Kodierung der Primärstudien und Referenzen

Die systematische Suche nach Studien und Referenzen umfasste mehrere Schritte (siehe Anhang A für Details). Insgesamt ergab der Suchprozess 207 potentiell relevante Referenzen, von denen 27 Referenzen die Auswahlkriterien erfüllten. Zu den Einschlusskriterien gehörte, dass in den Studien primär FIBB in Europa mit einem längsschnittlichen Design untersucht wurde. Voraussetzung war ebenfalls, dass in den Studien mindestens ein Maß für den Zusammenhang von Prozessqualität (globales und/oder bereichsspezifisches Prozessmaß) und Bildungsergebnis (in Mathematik und/oder Sprache/Literacy) beinhaltet sein sollten. Nicht berücksichtigt wurden Evaluationsstudien spezifischer Offensiven, Initiativen oder Multikomponenten-Interventionen, welche nicht die Auswirkungen regulärer FIBB widerspiegeln. Eher unspezifische Offensiven bzw. Initiativen, welche lediglich auf ein Anheben der Qualität oder Quantität der FIBB-Regelversorgung abzielten, gingen jedoch in die Metaanalyse ein. Zudem wurde sich in der Metaanalyse auf Studien mit einem Beginn nach 1990 beschränkt. Studien zu spezifischen Risikogruppen, z. B. hyperaktive Kinder oder Kinder mit Lernstörungen, wurden ausgeschlossen. Kodiererinnen extrahierten die Informationen aus den Referenzen unabhängig voneinander nach einem strukturierten Kodierschema und fügten sie in SPSS ein, wobei sie im Vorfeld intensive Schulungen, ein Kodiermanual und Beispielkodierungen erhielten sowie durch fortwährende Feedbackgespräche unterstützt wurden (für Details zum Kodierprozess siehe Anhang A). Insgesamt basiert die Metaanalyse auf 143 einzelnen Effektgrößen (*ES* für Effect Sizes) aus 17 Studien (siehe Tabelle 4 im Anhang B).

8.4.2 Statistisches Vorgehen für die Berechnung der Gesamteffekte

In die Berechnung der Gesamteffekte flossen standardisierte Koeffizienten für den Zusammenhang von Prozessmaß und Bildungsergebnis aus verschiedenen Auswertungsverfahren ein (u. a. multiple Regression, Korrelation und Wachstumsmodelle; siehe Anhang A für Details zum Moderatoreffekt des statistisch-analytischen Vorgehens; Bowman, 2012; Card, 2012; Kim, 2011). Alle *ES* gehörten der Korrelationsfamilie an (Döring & Bortz, 2014) und lagen in standardisierter Form vor mit einem Wertebereich von -1 bis 1, wobei eine *ES* von 0 gegen einen Zusammenhang von Prozessqualität und Bildungsergebnis sprach. Wenn gemäß Theorie ein negativer Effekt für ein Prozessmaß zu erwarten war, z. B. für die CIS-Subskalen Detachment und Punitiveness, wurde post-hoc die Inverse der *ES* gebildet (Arnett, 1989). Die Längsschnittmehrebenenmetaanalyse (Maas, Hox & Lensvelt-Mulders, 2004) wurde mit dem Package metafor im Statistikprogramm R durchgeführt (Viechtbauer, 2010), um die hierarchische Struktur der Metadaten zu berücksichtigen: *ES* für den Zusammenhang verschiedener Prozess- und Bildungsergebnismaße (Ebene 1) sind innerhalb Messzeitpunkten bzw. in Bezug auf das Alter der Kinder (Ebene 2) und Studien genestet.

Die *ES* wurden jeweils mit den Inversen ihrer Varianz gewichtet (Card, 2012; Cooper, 2009), so dass Befunde größerer Studien stärkeres Gewicht erhielten. Die Analysen basieren auf einem Restricted-Maximum-Likelihood (REML)-Schätzer (vgl. Viechtbauer, 2005). Zwei separate Gesamteffekte für den mittleren Zusammenhang der beiden Prozesskomponenten wurden gebildet (siehe Abbildung 5 und Kapitel 5.1 für Details zu Prozessmaßen).

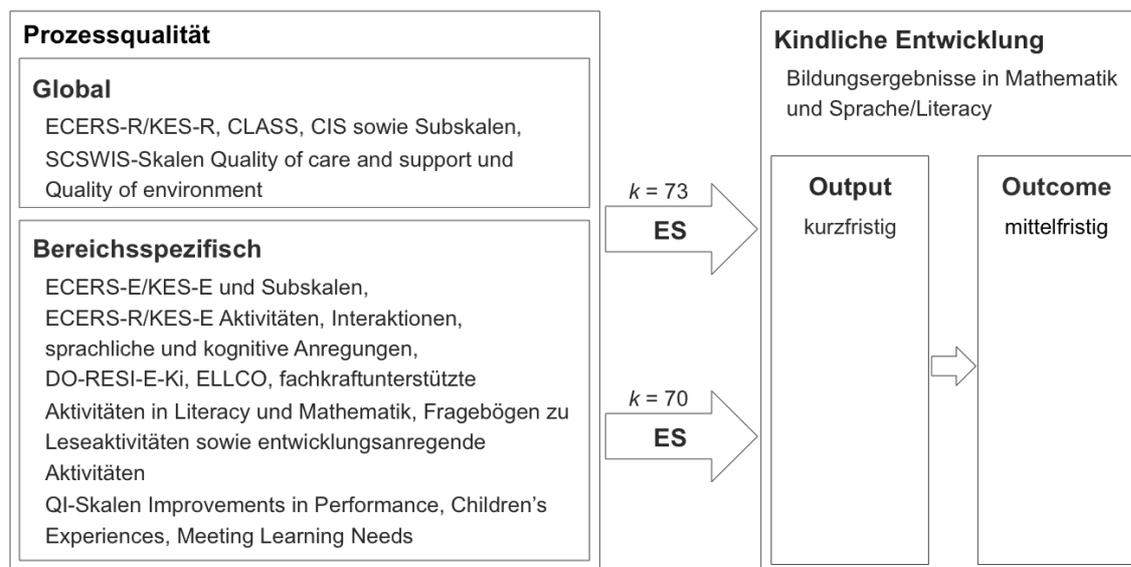


Abbildung 5. Übersicht über die berücksichtigten Befunde bzw. Prozessmaße in den Effekten globaler und bereichsspezifischer Prozessqualität.

8.4.3 Moderatoranalysen und Evaluation möglicher Verzerrungen

Die Moderatoranalysen erfolgten mit Mixed-Effects-Modellen (Viechtbauer, 2010) basierend auf den zuvor beschriebenen Längsschnittsmehrebenenmodellen. Das Alter der Kinder wurde als kontinuierlicher Moderator getestet. Für die restlichen Moderatoren wurden kategoriale Moderatoren mit Bezug zu untersuchten Bildungsergebnissen gebildet (akademische Phase der Erhebung in der FIBB- vs. Schulphase, Zusammenhänge zum Zuwachs bzw. zur Entwicklung und Bildungsbereich Mathematik vs. Sprache/Literacy) sowie mit Bezug zu Prozessmaßen (Interaktionsfokus vs. zusätzliches Umgebungsrating globaler Prozessmaße und Beobachtung vs. Fragebogen bereichsspezifischer Prozessmaße). Die Untersuchung der einzelnen Moderatoren erfolgte separat, um dem Verlust an statistischer Power und dem Problem der Multikollinearität vorzubeugen (Cooper, 2009). Das Ausmaß und die Konsequenzen möglicher Verzerrungen wurden mit gängigen Verfahren eingeschätzt (siehe Anhang A für Details; vgl. Higgins & Green, 2008).

8.5 Ergebnisse der Metaanalyse zu den Prozesseffekten des TP1

8.5.1 Deskriptiver Überblick über die Primärstudien

Die Metaanalyse basiert auf 17 Längsschnittstudien aus neun europäischen Ländern. In fünf Studien (KiDZ, FC Eva, TransKiGs, pre-COOL und Utrecht Mixed Preschool Groups) wurden Offensiven bzw. Initiativen evaluiert, welche insbesondere auf eine Anhebung der Prozessqualität in der Regelversorgung abzielten. Insgesamt lagen Befunde für 16 461 Kinder in 2 887 Einrichtungen für eine Altersspanne von drei bis 16 Jahren vor, jedoch gehen die meisten Befunde nicht über die Grundschulphase hinaus. Die berücksichtigten Studien unterschieden sich in Stichprobencharakteristika, Prozess- und Bildungsergebnismaßen und methodisch-statistischem Vorgehen (siehe Anhang A für weitere Beschreibungen der Primärstudien). Die durchschnittliche Prozessqualität der berücksichtigten Studien befand sich zumeist auf mittlerem Qualitätsniveau, wobei das bereichsspezifische Qualitätsniveau niedriger ausfiel (siehe Tabelle 2 im Anhang B).

8.5.2 Höhe und Varianz der Prozesseffekte

Tabelle 2 stellt die Ergebnisse für die Gesamteffekte globaler und bereichsspezifischer Prozessqualität sowie Angaben zur Varianz dar. Gemittelt über die 73 bzw. 70 *ES* aus 13 bzw. 11 Studien resultierten signifikante, positive Effekte in geringer Höhe für globale ($ES = .11$, $se = .04$, $p < .01$, Cohens $d = .22$) und bereichsspezifische Prozessqualität ($ES = .10$, $se = .03$, $p < .001$, Cohens $d = .20$). Für beide Effekte ergaben sich Varianzen innerhalb ($\sigma_{Alter} = .023$ und $.01$) sowie zwischen Studien ($\sigma_{Studie} = .005$ und $.002$; siehe Abbildung 1 und 2 im Anhang B).

für Forest Plots). Für beide Effekte deutete der Cochran's Q-Test auf Homogenität die Signifikanz der Varianz an: $Q_{ES}(69)= 264.55$ und $Q_{ES}(72)= 2482.15$ ($p < .001$; Higgins & Green, 2008).

Tabelle 2

Ergebnisse der Längsschnittsmehrebenenmetamodelle für globale und bereichsspezifische Prozesseffekte

Effekt	Stichprobe		Effekt			Varianz			
	k_{ES}	$k_{Studien}$	ES	SE	$CI_{95\%}$	Q_{ES}	df_{ES}	σ^2_{Alter}	$\sigma^2_{Studien}$
Global	73	13	.11**	.04	.04 - .18	2482.15***	72	.02	.00
Bereichs.	70	11	.10***	.03	.05 - .15	264.55***	69	.01	.00

Anmerkung. *** $p < .001$, ** $p < .01$, * $p < .05$.

8.5.3 Moderatoreffekte mit Bezug zu den Bildungsergebnissen und Prozessmaßen

Tabelle 3 gibt einen Überblick über die Ergebnisse der Moderatoranalysen, z. B. ob ein Moderatoreffekt vorlag ($df = 1$ für alle Q_M) und ob nach Berücksichtigung des Moderatoreffekts noch substantielle Varianz verblieb (Q_E). Keine der untersuchten Moderatoren erklärte die Varianz der Effekte vollständig: $Q_E(71) = 2139.65$ bis 2476.44 für globale Prozessqualität und $Q_E(68) = 226.07$ bis 264.53 für bereichsspezifische Prozessqualität. Die Prozesseffekte verloren nicht an Stärke, weder graduell mit zunehmendem Alter der Kinder ($Q_M = 0.33$, $p = .56$ und $Q_M = 0.004$, $p = .95$) noch nach dem Übergang der Kinder in die Schule. Beide Effekte spiegeln auch Entwicklungseffekte wider. Für die globale Prozessqualität ergaben sich ausschließlich Effekte zum Kompetenzzuwachs bzw. zur kindlichen Entwicklung ($Q_M = 3.93$, $p < .05$, $ES = .14$, $p < .001$ vs. $ES = .05$, $p = .33$). Obschon sich die Effekte in beiden Bildungsbereichen zeigten, wurden Bereichsunterschiede deutlich: $Q_M = 36.04$, $p < .001$ bzw. $Q_M = 6.32$, $p < .05$. Globale Prozessqualität hing stärker mit Bildungsergebnissen im Bereich Sprache/Literacy zusammen ($ES = .13$, $p < .001$ vs. $ES = .07$, $p < .05$), während die bereichsspezifischen Prozesseffekte stärker im Bereich Mathematik ausfielen ($ES = .11$, $p < .01$ vs. $ES = .08$, $p < .001$). Die Art des Prozessmaßes war ein signifikanter Moderator für beide Prozesskomponenten ($Q_M = 9.35$, $p < .01$ und $Q_M = 4.8$, $p < .05$). Für interaktionsfokussierte Maße ergaben sich stärkere Effekte als für Ratings globaler Prozessqualität, die auch die räumlich-materielle Umgebung umfassten ($ES = .19$, $p < .001$ vs. $ES = .09$, $p < .05$). Ein signifikanter bereichsspezifischer Effekt resultierte nur für Beobachtungs- und nicht für Fragebogenmaße ($ES = .12$, p

< .001 vs. $ES = .01$, $p = .78$). Zusätzliche Mixed-Effects-Modelle testeten, ob Unterschiede im methodisch-statistischen Vorgehen die Prozesseffekte moderierten (siehe Anhang A für Details). Insgesamt zeigten sich lediglich geringe Anzeichen für Verzerrungen, wobei sich die Gesamteffekte als robust erwiesen (siehe Anhang A für Details).

Table 3

Ergebnisse der Mixed-Effects-Modelle für Moderatoren mit Bezug zu Bildungsergebnissen und Prozessmaßen

Subeffekt	Moderator-	Residualvarianz		Stichprobe		Effekt			
	test	Q_M^a	Q_E	df_E	k_{ES}	$K_{Studien}$	ES	SE	$CI_{95\%}$
Akademische Phase ^b									
Global - FIBB	1.78	2332.6***	71	48	11	.09#	0.04	.00 - .17	
Global - Schule				25	6	.17**	0.05	.06 - .28	
Bereichs. - FIBB	1.12	263.75***	68	45	10	.12***	0.03	.05 - .18	
Bereichs. - Schule				25	6	.07#	0.04	.00 - .14	
Niveau vs. Entwicklung									
Global - Niveau	3.93*	2195.97***	71	32	6	.05	0.05	-.05 - .15	
Global - Entwicklung				41	13	.14***	0.04	.07 - .22	
Bereichs. - Niveau	1.17	263.22***	68	39	6	.12***	0.03	.05 - .19	
Bereichs. - Entwicklung				31	10	.08*	0.03	.02 - .14	
Bildungsbereich									
Global – Sprache/Literacy	36.04***	2346.47***	71	46	12	.13***	0.04	.06 - .21	
Global - Mathematik				27	7	.07*	0.04	.00 - .15	
Bereichs. - Sprache/Literacy	6.32*	235.83***	68	43	10	.08**	0.03	.03 - .14	
Bereichs. - Mathematik				27	9	.11***	0.03	.06 - .17	
Prozessmaß									
Global – Interaktionsfokus	9.35**	2139.65***	71	24	5	.19***	0.05	.09 - .29	
Global - Umgebungsrating				49	10	.09*	0.04	.01 - .18	
Bereichs. - Beobachtung	4.80*	226.07***	68	55	8	.12***	0.03	.07 - .18	
Bereichs. - Fragebogen				15	3	.01	0.04	-.07 - .10	

Anmerkung. *** $p < .001$, ** $p < .01$, * $p < .05$, # $p < .10$. ^a $df = 1$ für alle Q_M . ^bAlter als kontinuierlicher Prädiktor moderierte die Effekte nicht.

8.6 Diskussion zu Auswirkungen der Prozessqualität des TP1

Die Metaanalyse fand signifikante und robuste, positive Zusammenhänge der Prozessqualität zu kindlichen Bildungsergebnissen trotz Varianz der 143 Einzelbefunde. Die Effekte spiegeln die gebündelte Evidenz aus 17 europäischen Längsschnittstudien aus neun Ländern mit verschiedenen FIBB-Systemen und pädagogischen Traditionen wider. Die gefundenen Gesamteffekte ($ES = .11$ bzw. $d = .22$ für globale und $ES = .10$ bzw. $d = .20$ für bereichsspezifische Prozessqualität) sind gering, fallen im Vergleich mit Ergebnissen aktuellerer Metaanalysen aber nur etwas niedriger aus ($d = .23 - .34$; Blok et al., 2005; Camilli et al., 2010; Karoly et al., 2005; Manning et al., 2010; Nores & Barnett, 2010). Das ist insofern erstaunlich, als dass die evaluierten Offensiven und Initiativen sich teilweise durch eigene Curricula auszeichneten, welche speziell auf die Förderung in kognitiven Bereichen ausgerichtet waren und sich hinsichtlich qualitativer und quantitativer Merkmale von regulärer FIBB abhoben sowie teils in Kombination mit weiteren Unterstützungsangeboten erfolgten. Zudem zogen die bisherigen Metaanalysen zwar teils andere Fördermaßnahmen, vielmals aber Stichproben von Kindern regulärer FIBB oder ohne Betreuungserfahrung zum Vergleich heran. In dem Vergleich beider Studienarten errechneten Camilli et al. (2010) nur für Studien ohne vergleichbare Fördermaßnahmen als Kontrollgruppe einen signifikanten Effekt für die kognitive Domäne ($d = .23$ vs. $d = .07$). Die vorliegende Metaanalyse hingegen fand signifikante Effekte für den Vergleich von FIBB-Angeboten unterschiedlicher prozessualer Qualität.

Die Ergebnisse der Moderatoranalysen sprechen für die Persistenz der Prozesseffekte über die untersuchte Altersspanne hinweg auch nach dem Übergang in die Schulphase, in der Kinder neuen institutionellen Einflüssen ausgesetzt sind (z. B. Anders et al., 2013; Lehl et al., 2016; Sylva et al., 2010). Dies entspricht Befunden zu Langzeiteffekten bestimmter Förderprogramme, allerdings fällt der Follow-up-Zeitraum in den Evaluationsstudien vergleichsweise kurz aus. In Metaanalysen zur Effektivität von Offensiven bzw. Initiativen wird generell für die kognitive Domäne von einer signifikanten Abnahme der Effekte über die Zeit berichtet, allerdings wird Evidenz einbezogen, die bis ins junge oder sogar mittlere Erwachsenenalter reicht (Blok et al., 2005; Camilli et al., 2010; Karoly et al., 2005; Manning et al., 2010; Nores & Barnett, 2010). Die Befunde scheinen Zusammenhänge von Prozessqualität und kindlicher Entwicklung und nicht Alternativerklärungen widerzuspiegeln, z. B. herkunftsbedingte Zuweisung zu Qualität oder leistungsstarken Gruppen als günstige Rahmenbedingungen für die pädagogische Arbeit und eine hohe Prozessqualität (vgl. EACEA/Eurydice, 2009; Lehl et al., 2014; Leseman & Slot, 2014; OECD, 2006).

Die Ergebnisse sprechen zwar dafür, dass globale und bereichsspezifische Prozesskomponenten für Bildungsergebnisse in beiden untersuchten schulleistungsbezogenen Bereichen relevant sind, jedoch deutet sich eine Bereichsspezifität prozessualer Effekte an. Ein besonderer Stellenwert für Bildungsergebnisse in Mathematik scheint der bereichsspezifischen Anregung und Stimulation des Lernens zuzukommen (vgl. Burger, 2014; Hamre, Hatfield, Pianta & Jamil, 2014; Kluczniok & Roßbach, 2014; Kuger & Kluczniok, 2008; OECD, 2006, 2011). Bildungsergebnisse im Bereich Sprache/Literacy scheinen stärker von der globalen Prozessqualität abzuhängen. Eine mögliche Erklärung hierfür ist der inhärente Sprachbezug der Prozessmaße. Verbale Interaktionen stellen ein wesentliches Fundament des pädagogischen Geschehens in FIBB dar (Halle et al., 2010; Harms et al., 1998; Pianta et al., 2008). Ein FIBB-Setting hoher globaler Prozessqualität offeriert somit stets ein gewisses Niveau sprachlicher Anregungen und Stimulation. Nicht notwendigerweise haben die Interaktionen und Prozesse allerdings hohen mathematischen Anregungsgehalt. Gute Werte der globalen Prozessqualität sind also auch bei einem niedrigen Niveau mathematischer Förderung erreichbar.

Die Ergebnisse der Moderatoranalysen verwiesen auf wichtige Unterschiede zwischen existierenden Prozessmaßen. Interaktionsfokussierte Maße resultierten in stärkere globale Prozesseffekte, als Maße, die zusätzlich Ratings der räumlich-materiellen Umwelt einschlossen. Generell, aber im Speziellen auch zur Weiterentwicklung von Prozessratings, bedarf es weiterer Forschung darüber, welche Funktion die räumlich-materielle Umgebung im Bildungs- und Lernprozess einnimmt (z. B. Schuler, 2009). Um das Bildungspotenzial frühkindlicher Einrichtungen aufzuzeigen, scheinen Beobachtungsmaße empfehlenswert, denn die untersuchten Fragebogenmaße bildeten die entwicklungsrelevanten Komponenten bereichsspezifischer Prozessqualität nur unzureichend ab. Zudem verweist der Befund auf wichtigen Forschungs- und Entwicklungsbedarf für Fragebogenverfahren zur Erfassung der Prozessqualität.

Insgesamt weisen die Ergebnisse des TP1 darauf hin, dass das Bildungspotenzial frühkindlicher Einrichtungen durch eine Anhebung der Prozessqualität in der FIBB-Regelversorgung optimierbar ist. In Anbetracht der niedrigen Durchschnittswerte der Prozessqualität bei geringer Varianz (Kuger & Kluczniok, 2008; Slot, Lerkkanen, et al., 2015; Tietze et al., 2012; Vermeer et al., 2016), insbesondere der bereichsspezifischen Förderung im mathematischen Bereich, ist anzunehmen, dass frühkindliche Einrichtungen in Europa dieses Bildungspotenzial noch unzureichend ausschöpfen. Wertvolle Ideen zur Qualitätsentwicklung in der FIBB sowie Vorschläge, wie mathematische Bildung in frühkindlichen Einrichtungen umsetzbar ist, liegen vor (Burger, 2014; Dooley et al., 2012; Gasteiger, 2012; Ginsburg et al., 2008; Kaufmann, 2011; Sommerlatte et al., 2009).

9 Effektivität FIBB in Deutschland – Diskussion und empirische Untersuchung des Ansatzes im Bildungsbereich Mathematik (TP2)

Im TP2 wird die Diskussion fortgeführt, ob sich die mathematische Bildungseffektivität als wichtige Ergänzung zu Komponenten etablierter Modelle eignet, um den Bildungoutput FIBB zu vergleichen (siehe Abbildung 4 und Kapitel 5.2). In Anbetracht der Verbreitung in der Schulforschung und vielversprechenden EPPSE-Befunde (siehe Kapitel 6.2) stellte sich die Frage, warum in der FIBB-Forschung in Deutschland der Effektivitätsansatz bislang nicht verwendet wurde. Im Kapitel 9.1 werden zuerst die konzeptuellen und methodologischen Herausforderungen des Qualitätsansatzes insbesondere im Zusammenhang mit der Prozessqualität zusammengetragen, welche als Argumente für eine Ergänzung durch den Effektivitätsansatz angeführt werden können. In den Kapiteln 9.2 und 9.3 werden schließlich mögliche Gründe gegen Effektivitätsanalysen in der FIBB erwogen und die Voraussetzungen an Studiendesigns beschrieben. Anschließend werden in den Kapiteln 9.4 und 9.5 die Untersuchung der mathematischen Effektivität FIBB in Deutschland mit Daten der Studie BiKS und die methodische Umsetzung beschrieben. Eine Beschreibung und Diskussion der Ergebnisse vor dem Hintergrund der Prozessqualitätsbefunde und relevanter Befunde anderer Studien der FIBB folgen in den Kapiteln 9.6 und 9.7. Bei der Beschreibung werden dem Promotionsthema entsprechend Schwerpunkte gesetzt. Details zur Studie sowie weitere Ergebnisse und ihre Interpretation können dem Anhang C entnommen werden.

9.1 Herausforderungen des Qualitätsansatzes

9.1.1 Konzeptuelle Herausforderungen

Eine erste Herausforderung ergibt sich aus der Notwendigkeit (a) *verschiedene Perspektiven im Qualitätskonzept zu berücksichtigen* (Mehrperspektiven-Ansatz; Katz, 1996), denn generell ermisst sich Qualität als Grad, indem das Betreuungs- und Bildungsangebot den Anforderungen unterschiedlicher Interessengruppen entsprechen (z. B. Eltern, Kinder, Fachkräfte oder Gesellschaft; DIN, 2005). Die definierten Qualitätskomponenten sollten zumindest die zentralen Anforderungen wichtiger Interessengruppen widerspiegeln. Zudem sind jeweils Standards festzulegen, welches Niveau als befriedigende Qualität gilt. Außerdem ist die Art des Verrechnens zu einem Gesamturteil zu bestimmen, z. B. ob alle, die meisten oder ganz spezielle Kriterien erfüllt sein müssen, damit das Gesamturteil gut bzw. befriedigend ausfällt. Jedoch können die Anforderungen verschiedener Interessengruppen teilweise sehr unterschiedlich ausfallen oder sogar im Widerspruch stehen. Beispielsweise erleichtern ausgedehnt-

te Öffnungszeiten Eltern die Vereinbarkeit von Familie und Beruf, erfordern aber vom Personal und Träger einen hohen Aufwand. Außerdem weisen Befunde darauf hin, dass eine zu lange Betreuungsdauer in den ersten zwei Lebensjahren mit Problemverhalten bei den Kindern einhergehen kann (Anders, 2013). Im Allgemeinen stellen die Bildungsförderung und das Wohlbefinden des Kindes die zentralen Bezugspunkte für die Auswahl und Gewichtung von Qualitätskomponenten dar (Roux & Tietze, 2007).

Weil Qualität einen Konsens unterschiedlicher Vorstellungen und Werte repräsentiert, handelt es sich um ein (b) *normatives Konstrukt* (Roux & Tietze, 2007). Problematisch aus wissenschaftlicher Perspektive wird die Wertverankerung, wenn sich dominante Vorstellungen und Werte nicht mit den Erkenntnissen der Forschung decken. So gibt es zahlreiche Befunde über wichtige Schritte mathematischer Entwicklung in den ersten drei Lebensjahren (Kaufmann, 2011; Pauen & Herber, 2009; Roßbach & Weinert, 2008). Broekhuizen, Leseman, Moser und van Trijp (2015) berichten allerdings, dass fast 2 500 Eltern Europas dem Erlernen früher akademischer Kompetenzen (u. a. frühe mathematische Kompetenzen) für Kinder unter drei Jahren nur eine geringe Bedeutung beimessen.

Aufgrund der Wertverankerung kann es zur Anpassung von Modellen an den jeweiligen Kontext und mit dem Wandel der Zeit kommen, weshalb (c) *Qualitätsvorstellungen ggfs. kultur-, system- und kontextspezifisch unterschiedlich ausfallen* (Moss & Pence, 1994; Siraj-Blatchford et al., 2006). Beispielsweise deuten sich zwischen Ländern Unterschiede in den elterlichen Vorstellungen zum frühen akademischen Lernen an (Broekhuizen et al., 2015). Auch kommt es zu veränderten Qualitätsauffassungen über die Zeit mit dem Wandel von Ansprüchen, FIBB-Systemen und Rahmenbedingungen sowie neuen Erkenntnissen über die kindliche Entwicklung und institutionelle Einflüsse. Fokussierten z. B. klassische Qualitätsauffassungen auf die strukturellen Qualitätskomponenten, rücken aktuell die pädagogischen Prozesse in den Mittelpunkt. Das Fundament etablierter Modelle bilden Expertenmeinungen, langjährige Praxiserfahrungen und empirische Evidenz, weshalb sich insgesamt eine hohe Überschneidung in Qualitätsauffassungen ergibt. Dennoch gibt es unterschiedliche Vorstellungen z. B. zur Rolle der räumlich-materiellen Umgebung für die Qualität pädagogischer Prozesse (siehe Kapitel 8).

Eine weitere Herausforderung resultiert aus dem gestiegenen Bewusstsein für die Bereichsspezifität von Anregungsprozessen (Hamre et al., 2014; Kuger & Kluczniok, 2008; Roßbach et al., 2008), weshalb in (d) *Qualitätsmodellen stärker bereichsspezifische Komponenten berücksichtigt werden müssen*, wie etwa die mathematikspezifische Professionalisie-

rung der frühpädagogischen Fachkräfte und anregende Erzieher-Kind-Interaktionen über mathematische Erfahrungen.

9.1.2 Methodologische Herausforderungen des Qualitätsansatzes

Eine erste methodologische Herausforderung betrifft die (a) *aufwändigen Erhebungsverfahren*. Ein umfassendes Bild der pädagogischen Qualität von Einrichtungen erfordert nach Katz (1996) die Einschätzung zahlreicher Qualitätskomponenten aus verschiedenen Perspektiven, z. B. Fachkräfte, Eltern, Kinder und Leitung, wobei sich in der Regel in etablierten Modellen auf ausgewählte Komponenten beschränkt wird. Während einige Komponenten sich vergleichsweise einfach erfassen lassen wie Personal- und Gruppenstruktur, erfordert die Erfassung der Prozessqualität im Allgemeinen gut geschulte Beobachterinnen und Beobachter, mehrstündige Begehungen und zusätzliche Interviews mit Fachkräften (Halle et al., 2010; Sylva et al., 2006; Tietze et al., 2007).

Trotz des großen Aufwands stellt sich die Frage, ob der (b) *begrenzte Einblick in das pädagogische Geschehen für eine zuverlässige Einschätzung der Prozessqualität ausreicht*. Hofer (2010) beobachtete für fast 200 frühkindliche Einrichtungen, dass eine Verlängerung der Zeitstichprobe zu niedrigeren Qualitätswerten in der ECERS-R führte. Praetorius, Pauli, Reusser, Rakoczy und Klieme (2014) schlussfolgern aus einer Untersuchung zur mathematischen Instruktionsqualität in Schulen, dass es mindestens neun Unterrichtsstunden bedarf, um das Ausmaß kognitiver Aktivierung verlässlich einzuschätzen. Kane und Staiger (2012) gaben auf Grundlage einer Untersuchung mit 1 333 Lehrkräften die Empfehlung, mehr als vier beobachtete Unterrichtsstunden auszuwerten und zusätzlich Maße der Bildungseffektivität hinzuzuziehen, um die Unterrichtsqualität einzuschätzen. Die geringere Strukturierung frühkindlicher Bildungsangebote im Vergleich zu schulischen Angeboten erschwert Prozesseinschätzungen vermutlich zusätzlich (z. B. keine festgelegten Zeiten, Orientierung an den Interessen und Bedürfnissen der Kinder; siehe Kapitel 3.3).

Die (c) *eingeschränkte Differenzierungsfähigkeit vieler Qualitätskomponenten* stellt Forschungsvorhaben vor eine Herausforderung, denn politische Regulierungen und Vorschriften zu Qualitätsstandards, z. B. zum Erzieher-Kind-Schlüssel (Viernickel & Schwarz, 2009), zielen im Allgemeinen darauf ab die Qualität FIBB zu vereinheitlichen und in den Studien zur Qualität wurden mehrheitlich die in der Regelversorgung innerhalb der Reglementierungen vorliegenden Qualitätsunterschiede untersucht. Die im Durchschnitt mittelmäßigen bis niedrigen Werte der Prozessqualität bei niedriger Varianz stellen nicht nur die Praxis vor eine Herausforderung sondern erschweren ebenfalls den empirischen Nachweis von Zusammenhängen (siehe Kapitel 8; Clifford et al., 2010; Kuger & Kluczniok, 2008; Tietze et al., 2012). Aus Per-

spektive der empirischen Bildungsforschung sind allerdings Nachweise zentraler Annahmen von Modellen erforderlich, insbesondere zu den Auswirkungen auf die kindliche Entwicklung.

Fraglich ist ebenso, ob die (d) *an einer spezifischen Gruppe gemessene Prozessqualität repräsentativ für die gesamte Einrichtung* ist, denn die meisten Untersuchungen wählen lediglich eine Gruppe der Einrichtung für Messungen aus, um Aussagen über die Prozessqualität der Einrichtung zu treffen (z. B. EPPSE; Melhuish et al., 2013; Sylva et al., 2006). Ergebnisse der Schulforschung verweisen allerdings darauf, dass die Unterrichtsqualität innerhalb von Schulen stark variiert und dass zumeist größere Unterschiede innerhalb als zwischen Schulen existieren (Goodwin, 2011; Konstantopoulos, 2006).

Zudem stellt sich die Frage, ob die (e) *einmalig erhobene Prozessqualität für das pädagogische Geschehen über einen längeren Zeitraum repräsentativ* ist. In vielen Längsschnittstudien wird von einer einmaligen Messung der Prozessqualität auf den gesamten Betreuungszeitraum der Kinder geschlossen (z. B. pre-COOL oder EPPSE; Slot et al., 2014; Sylva et al., 2006). Die Prozessqualität in frühkindlichen Einrichtungen zeigte jedoch lediglich über kurze Zeiträume eine mittlere bis hohe Stabilität und über längere Zeiträume innerhalb eines Jahres nur eine befriedigende Stabilität (Clifford, 2005; Clifford et al., 2010; Curby, Grimm & Pianta, 2010; Hofer, 2010; Kuger et al., 2016; Pianta & Hamre, 2009; Tietze, Roßbach & Grenner, 2005) und sie scheint für die bereichsspezifische Prozessqualität noch niedriger auszufallen (Kuger et al., 2016).

Bei einer aus empirischer Sicht notwendigen Überprüfung der dimensional Struktur mit Faktorenanalysen ergaben sich mehrheitlich (f) *Abweichungen der empirisch ermittelten von der angenommenen dimensional Struktur* für etablierte Prozessmaße (Clifford et al., 2010; Gordon, Fujimoto, Kaestner, Korenman & Abner, 2013; Hamre et al., 2014; Pianta & Hamre, 2009; Pianta et al., 2005). Beispielsweise resultierte für die ECERS-R zumeist eine zweidimensionale Struktur aus materiell-räumlichen und interaktionsbezogenen Ratings.

Die ermittelten niedrigen Prozesseffekte auf Bildungsergebnisse im TP1 könnten auch auf unzureichend gelöste konzeptuelle und methodologische Herausforderungen etablierter Qualitätsmodelle hinweisen (siehe Kapitel 8). Daher bedarf es dem Ansatz einer kontinuierlichen Weiterentwicklung und Ergänzung z. B. durch den Effektivitätsansatz.

9.2 Herausforderungen des Effektivitätsansatzes

Der Effektivitätsansatz stellt den Bildungoutput und somit mathematische Bildungsergebnisse in den Fokus (Creemers et al., 2010). Als erstes Argument gegen eine Verwendung in der FIBB in Deutschland kann die verbreitete (a) *Skepsis gegenüber einem stark lernziel-*

orientierten Qualitätsindikator angeführt werden. Wenn auch die Bildungsfunktion frühkindlicher Einrichtungen allgemein anerkannt wird, geben die bildungspolitische und pädagogische Umsetzung und die Implikationen für die Forschung noch immer Anlass zu Debatten, besonders wenn es um die Definition von Lernzielen und die Erfassung kindlicher Kompetenzen geht (Sylva et al., 2015). Zudem steht einer (b) *bereichsspezifische, ausschließlich auf Mathematik bezogene Beurteilung einem dominierend holistischem Verständnis kindlicher Entwicklung in der FIBB gegenüber* (siehe Kapitel 3.2 und 3.3). Erkennbar ist zudem eine starke Priorisierung sozio-emotionaler gegenüber kognitiv-leistungsbezogener Bereiche, mit der eine Angst vor der Verschulung der FIBB verbunden ist, welche dem Spiel, den individuellen kindlichen Bedürfnissen und der Persönlichkeitsentfaltung zu wenig Raum lässt. Dabei liegen überzeugende Vorschläge zur Umsetzung mathematischer Bildung vor, welche die spielbasierte, kindorientierte und in den FIBB-Alltag eingebettete Förderung in Mathematik und weiteren Bereichen vorsehen (siehe Kapitel 3.3; Burger, 2014; Dooley et al., 2012; Fthenakis et al., 2009; Gasteiger, 2012; Ginsburg et al., 2008; Kaufmann, 2011; NAEYC & NCTM, 2010; Rechsteiner et al., 2012; Siraj-Blatchford, 2009; Sommerlatte et al., 2009; van Oers, 2010).

Weil ausschließlich der Bildungsoutput FIBB betrachtet wird, (c) *werden keine Aussagen zur Art der Lernerfahrungen und Umstände getroffen*. Damit bleiben ggfs. Merkmale unberücksichtigt, die für das Wohlbefinden und die kindliche Entwicklung generell förderlich sind wie die Wärme und Responsivität der Interaktionen. Prinzipiell lassen sich Lernfortschritte auch unter wenig kindgerechten Bedingungen erreichen wie mit Druck sowie monotonem und repetitivem Einüben. Der verbreiteten Kritik, dass eine frühe Lernförderung sich negativ auf die sozio-emotionale Entwicklung und das Wohlbefinden auswirke, widersprechen Befunde europäischer Längsschnittstudien allerdings, die überwiegend positive Zusammenhänge berichteten (z. B. Abreu-Lima et al., 2013; Bradshaw et al., 2014; Kluczniok, Anders, Sechtig & Roßbach, 2016; Leal, Abreu-Lima & Cadima, 2013; Tsakiri, 2012).

Die (d) *Effektivitätsergebnisse geben keine Auskunft über mögliche Gründe für gefundene Unterschiede*, weshalb sich weniger direkt Implikationen ableiten lassen z. B. Ansatzpunkte zur Optimierung der pädagogischen Arbeit oder zu den Rahmenbedingungen FIBB. Des Weiteren ist die (e) *Aussagekraft des Effektivitätsindikators auf die erfassten Kompetenzen beschränkt* und muss daher vor dem Hintergrund des verwendeten Messverfahrens bewertet werden (z. B. erfasste und nicht erfasste Kompetenzen, Gütekriterien des Testverfahrens). Effektivitätsanalysen bieten ebenfalls (f) *keine Antwort auf die Frage, ob das Durchschnittsniveau zufriedenstellend ist*, d. h. für eine ausreichend mathematische Bildungsarbeit in der Ein-

richtungsstichprobe spricht, weil normative Standards wie im Qualitätsansatz fehlen. Das Effektivitätsurteil ermittelt sich relational durch den Vergleich des Bildungsoutputs einer Einrichtung mit dem durchschnittlichen Output in der Referenzstichprobe anderer Einrichtungen (Chapman et al., 2015; Creemers et al., 2010). Die Qualitätsurteile und die Differenzierungsfähigkeit des Effektivitätsindikators sind somit abhängig von den Verteilungseigenschaften der Referenzstichprobe. Insgesamt scheinen die Argumente ideeller Natur ohne überzeugendes empirisches Fundament zu sein oder sogar im Widerspruch mit der Befundlage und vorgeschlagenen Bildungsansätzen zu stehen. Der Beitrag der Bildungseffektivität zur Forschung und Praxis FIBB scheint eher von geeigneten Erhebungsdesigns und einem sensiblen und gut durchdachten Umgang mit Effektivitätsergebnissen abzuhängen.

9.3 Anforderungen von Effektivitätsanalysen an Studiendesigns

Die Voraussetzungen, die Effektivitätsanalysen an Studiendesigns stellen, könnten ebenfalls erklären, warum Effektivitätsanalysen bislang in der deutschen FIBB nicht zur Anwendung kamen. Weil Effektivitätsurteile auf Mehrebenenanalysen basieren, sollten Studiendesigns die Mehrebenenstruktur FIBB (Kinder befinden sich in Gruppen, welche in Einrichtungen organisiert sind) in ausreichendem Maße widerspiegeln (Creemers et al., 2010; siehe Kapitel 3.4). Hierbei ist neben der Repräsentativität die Größe der Stichproben der Ebenen entscheidend, z. B. empfiehlt Melhuish (2011) mindestens 20 Kinder pro Einrichtung zu testen. Liegen wenig Angaben zu Kindern aus einer Einrichtung vor, wird der für diese Einrichtung ermittelte Effektivitätswert fehleranfälliger (Ballou & Springer, 2015). Für Effektivitätsanalysen scheint im Besonderen die Größe der Einrichtungstichprobe entscheidend (Maas & Hox, 2005). Die Grundlage für mathematische Effektivitätsanalysen bilden die Daten der Kinder zu Bildungsergebnissen in Mathematik (siehe Kapitel 6.2). Längsschnittdaten ermöglichen den Zuwachs oder die Entwicklung der Kompetenz in Einrichtungen zu vergleichen, was die Fairness von Effektivitätsvergleichen erhöht (VAM-Modelle; Creemers et al., 2010; Reynolds et al., 2014). Liegen Angaben über individuelle und familiäre Merkmale vor, können diese in kontextualisierten Vergleichen genutzt werden (CAM- oder CVA-Modelle), um ungleiche Ausgangsbedingungen für die Kompetenzentwicklung in Einrichtungen zumindest teilweise zu berücksichtigen. International sowie in Deutschland erfüllen die Designs einiger Studien grundsätzlich die Voraussetzungen für Effektivitätsanalysen (z. B. BiKS, NEPS oder NUB-BEK).

9.4 Beschreibung der Effektivitätsanalysen für die BiKS-Einrichtungen des TP2

Im TP2 sollen mathematische Effektivitätsanalysen für die deutsche FIBB mit Daten der Studie BiKS-3-10 umgesetzt werden (von Maurice et al., 2007), in welcher jährlich zu drei Messzeitpunkten (MZP) mathematische Kompetenzen in Bereichen erhoben wurden, die sich als besonders prädiktiv für die weitere mathematische Entwicklung erwiesen haben (Jordan et al., 2009; Schneider et al., 2013). Wie in EPPSE eignet sich das Design der BiKS-Studie daher, um unterschiedliche Effektivitätsanalysen durchzuführen und die Ergebnisse mit bereits publizierten Befunden zur Prozessqualität zu vergleichen. Für eine maximale Vergleichbarkeit richtet sich die Auswahl entwicklungsrelevanter Merkmale nach vorhergehenden BiKS-Analysen zur Prozessqualität (Anders et al., 2012; vgl. Anders et al., 2013; Lehl et al., 2016; Lehl et al., 2014), welche sich in anderen Studien ebenfalls als relevant für die mathematische Entwicklung gezeigt haben (Abreu-Lima et al., 2013; Arnold & Doctoroff, 2003; z. B. Sammons et al., 2002; Schneider et al., 2013; Sylva et al., 2010). Der Fokus der Analysen liegt jedoch erstmalig auf der Mehrebenenstruktur von BiKS und bezieht sich nicht auf die individuellen sondern institutionellen Kompetenzniveaus und -verläufe. Anders als in EPPSE liegen in BiKS Kompetenzangaben zu drei MZP vor, weshalb ein Vergleich der mathematischen Bildungseffektivität zu den MZP möglich ist (Chapman et al., 2015).

Im TP2 soll untersucht werden, ob sich die mathematische Bildungseffektivität der BiKS-Einrichtungen für die MZP unterschiedet. Es stellen sich die Fragen, ob sich für die Einrichtungen Unterschiede (a) *im mittleren Kompetenzniveau* oder (b) *im mittleren Kompetenzzuwachs* in den untersuchten mathematischen Kompetenzbereichen ergeben und ob sich Unterschiede zwischen Einrichtungen in *kontextualisierten Vergleichen* (c) *des Niveaus* oder (d) *des Zuwachses* zeigen, in welchen entwicklungsrelevante Merkmale berücksichtigt werden. Der Anhang C beinhaltet zudem die Ergebnisse und Interpretation der Rolle einzelner entwicklungsrelevanter Merkmale für die Effektivitätsvergleiche, welche Anhaltspunkte für die Auswahl von Merkmalen für zukünftige Effektivitätsanalysen bieten.

9.5 Methode der mathematischen Effektivitätsanalysen des TP2

9.5.1 Stichprobe

In BiKS wurden zu drei MZP kindliche Kompetenzen in 97 Kindergärten in Bayern und Hessen erhoben (60 bzw. 37 Einrichtungen). In der Studie wurde ein gruppen- und altersbasiertes Sampling verwendet: Nur eine Einrichtung pro Gruppe nahm teil (gruppenbasiert), wobei die Wahl auf die Gruppe mit den meisten Kindern im Zielalter fiel (altersbasiert; von Maurice et al., 2007). Das Zielalter ermittelte sich am regulären Einschulungstermin im Schuljahr

2008/2009 der beiden Bundesländer (Geburtsfenster in Bayern: 01.10.2001 - 31.10.2002 und Hessen: 01.07.2001 - 30.06.2002). In jeder Gruppe wurde lediglich ein Teil der Kinder getestet (durchschnittliche Anzahl pro Gruppe: 5.71; Spannweite: 4 bis 10). Die insgesamt 554 Kinder (48 % Mädchen) waren zum ersten MZP durchschnittlich drei Jahre ($M = 44.55$, $SD = 5.00$), zum zweiten MZP vier Jahre ($M = 55.73$, $SD = 4.47$) und zum dritten MZP fünf Jahre ($M = 67.43$, $SD = 4.25$) alt. Die Kinder waren ungefähr drei Jahre alt bei Eintritt in die Einrichtungen ($M = 37.54$, $SD = 5.21$). Für die Mehrheit der Kinder fanden die Testungen somit im ersten bis dritten Kindergartenjahr statt. Die Eltern der Kinder gaben mehrheitlich Deutsch als Muttersprache an. 10% bzw. 12 % der Kinder entstammten Elternhäusern mit einem Elternteil bzw. beiden Eltern anderer Muttersprache. 34 % der Mütter besaßen die mittlere Hochschulreife, weitere 34 % Abitur und 25 % hatten die Schule ohne Abschluss oder mit Hauptschulabschluss verlassen. 7 % der Mütter gaben einen sonstigen, meist im Ausland erlangten Abschluss an.

9.5.2 Instrumente

In die Effektivitätsanalysen flossen die Summenwerte der Kinder zu den drei MZP aus dem Untertest Rechnen der deutschen Version der Kaufman-Assessment Battery for Children (K-ABC; Melchers & Preuß, 2003). Der Test erfasst elementare numerische und geometrische Kompetenzen wie Zählfertigkeiten, Ziffern- und Formkenntnis, das grundlegende Mengenverständnis sowie basale Rechenfertigkeiten. Der Mittelwert der Stichprobe stieg über die drei MZP mit zunehmendem Alter der Kinder an: Mit etwa drei Jahren erreichten die Kinder durchschnittlich 4.90 Punkte ($SD = 3.38$), mit etwa vier Jahren 10.26 Punkte ($SD = 3.98$) und mit etwa fünf Jahren 15.03 Punkte ($SD = 3.70$). In die kontextualisierten Vergleiche flossen zusätzlich entwicklungsrelevante, individuelle Kind- und Familienmerkmale ein (Anders et al., 2013; Anders et al., 2012; Lehl et al., 2014; vgl. Schneider et al., 2013): Geschlecht, Alter zum jeweiligen MZP und bei Eintritt in die Einrichtung, die Muttersprache der Eltern, der mütterliche Bildungsabschluss, der Highest International Socio-Economic Index (HISEI) als SES-Indikator (vgl. Ganzeboom, De Graaf & Treiman, 1992) sowie die häusliche Lernumwelt (HLE; siehe Anhang C für Details).

9.5.3 Statistische Auswertung

Die Effektivitätsvergleiche erfolgten mit unterschiedlichen Mehrebenenanalysen (Ebene 1: Kindebene, Ebene 2: Einrichtungsebene). Die Vergleiche wurden mit Random-Intercept-Modellen in *Mplus* durchgeführt (type = twolevel, cluster = Einrichtung). Für alle Analysen wurde ein Maximum-Likelihood-Schätzer mit robusten Standardfehlern (MLR) verwendet,

um für die Clusterung und fehlende Normalverteilung von Daten zu korrigieren, sowie die Full Information Maximum Likelihood-(FIML)-Methode, um fehlende Werte zu schätzen (Lüdtke, Robitzsch, Trautwein & Köller, 2007). Die vier in der Fragestellung skizzierten Vergleichsarten wurden mit unterschiedlichen Modellen umgesetzt. Die (a) Niveauvergleiche der mathematischen Kompetenz zu drei MZP basierten auf Intercept-Only- bzw. Null-Modellen (Geiser, 2011). Mittels Random-Intercepts-Modellen erfolgten die (b) Zuwachsvergleiche der mathematischen Kompetenz zum zweiten und dritten MZP unter Berücksichtigung früherer Kompetenz auf Kindebene (VAM-Modelle) sowie die kontextualisierten Vergleiche des Niveaus (c) und des Zuwachses (d), in die zusätzlich entwicklungsrelevante Merkmale auf Kindebene einfließen (CAM- bzw. CVA-Modelle). Signifikante ICCs (für Intra-Class-Correlation) deuten Varianzen auf Einrichtungsebene und somit Effektivitätsunterschiede an. Liegen signifikante Unterschiede vor, bildet das Residuum dieser Ebene für die einzelnen Einrichtungen den Effektivitätswert (Sammons et al., 2002; Sylva et al., 2010). Das Residuum zeigt an, wie stark der erzielte Bildungsausgang der jeweiligen Einrichtung vom Durchschnittsausgang der Einrichtungsstichprobe abweicht. Der Determinationskoeffizient $R^2_{innerhalb}$ gibt den Anteil der Varianz der mathematischen Kompetenz an, den die entwicklungsrelevanten Merkmale insgesamt erklären.

9.6 Ergebnisse zur mathematischen Effektivität FIBB des TP2

Tabelle 4 fasst die Ergebnisse für die vier Vergleiche zusammen. Die (a) Niveauvergleiche zeigen, dass sich das mittlere Kompetenzniveau in den Einrichtungen unterscheidet, d. h. es deuten sich signifikante Unterschiede in der mathematischen Bildungseffektivität an. Obschon für alle MZP signifikant, nimmt der Anteil erklärter Varianz über die Zeit von 21% zum ersten auf 13 % zum zweiten auf 8 % zum dritten MZP ab.

In BiKS wurden keine Kompetenzen vor dem Kindergarten erhoben, allerdings wurden die Kinder mehrheitlich unmittelbar nach Betreuungsbeginn getestet. Zusätzlich berechnete Nullmodelle für entwicklungsrelevante Merkmale deuten auf Unterschiede in der Zusammensetzung hinsichtlich der meisten entwicklungsrelevanten Merkmale (Alter zum MZP, Eintrittsalter, Migrationshintergrund, mütterlicher Bildungsabschluss und SES). Vermutlich unterschieden sich die Kinder in den Einrichtungen von Beginn an hinsichtlich mathematischer Kompetenz und entwicklungsrelevanter Merkmale.

Tabelle 4

Effektivitätsunterschiede (*ICC*) in den vier Vergleichen und durch entwicklungsrelevante Merkmale erklärte Varianzanteile in den mathematischen Kompetenzen ($R^2_{innerhalb}$) zu den drei Messzeitpunkten (MZP)

Effektivitätsvergleich	<i>ICC</i>	$R^2_{innerhalb}$	<i>N</i>	Mittlere Clustergröße
a Niveau				
1. MZP	.21***		554	5.71
2. MZP	.13**		554	5.71
3. MZP	.08*		554	5.71
b Zuwachs				
1. auf 2. MZP	.03	.43***	551	5.68
2. auf 3. MZP	.01	.52***	523	5.39
c Niveau kontextualisiert				
1. MZP	.08**	.33***	547	5.64
2. MZP	.02	.24***	547	5.64
3. MZP	.01	.18***	547	5.64
d Zuwachs kontextualisiert				
1. auf 2. MZP	.01	.40***	547	5.64
2. auf 3. MZP	.01	.53***	547	5.64

Anmerkungen. $N = 97$ Einrichtungen; * $p < .05$; ** $p < .01$; *** $p < .001$. Die Kinder waren zu den drei MZP etwa 3, 4 und 5 Jahre alt.

Die Befunde deuten insgesamt allerdings keine Effektivitätsunterschiede für „faire“ Vergleiche an. Die in (b) Zuwachsvergleichen ermittelten *ICCs* (weniger als 4 % für beide MZP) sprechen dagegen, dass der durchschnittliche Zuwachs mathematischer Kompetenz in den Einrichtungen signifikant voneinander abweicht. Bei Berücksichtigung entwicklungsrelevanter Merkmale in den (c) kontextualisierten Niveauvergleichen ergeben sich lediglich zum ersten MZP abweichende Niveaus mathematischer Kompetenz der Einrichtungen ($ICC = .08$). Zu den anderen MZP und in (d) kontextualisierten Zuwachsvergleichen ergeben sich keine signifikanten Effektivitätsunterschiede ($ICC = .02$ bzw. $.01$).

Die (b) Zuwachsvergleiche zeigen, dass fast die Hälfte der Varianz in den Kompetenzen durch die Kompetenz zum vorherigen MZP erklärt wird, wobei der Anteil zum zweiten MZP sogar ansteigt (43 % zum zweiten und 52 % zum dritten MZP). Wie aus den (c) kontextuali-

sierten Niveauvergleichen deutlich wird, klären entwicklungsrelevante Merkmale einen signifikanten Anteil der Kompetenz auf, aber insgesamt weniger als die Kompetenz zum vorherigen MZP und der Zusammenhang nimmt über die MZP von 33 % zum ersten auf 24 % zum zweiten MZP auf 18 % zum dritten MZP ab.

9.7 Diskussion zur mathematischen Effektivität der BiKS-Einrichtungen des TP2

Keine Unterschiede in der mathematischen Effektivität ergab sich für die BiKS-Einrichtungen in „fairen Vergleichen“ (Racherbäumer et al., 2013; Scheerens, 1992). Die divergierenden Kompetenzniveaus sind vermutlich nicht durch eine unterschiedliche Bildungsarbeit sondern durch unterschiedliche Zusammensetzungen in den Einrichtungen hinsichtlich entwicklungsrelevanter Merkmale erklärbar. Disparitäten in der frühen mathematischen Kompetenz entwickeln sich höchstwahrscheinlich bereits vor Eintritt in die Einrichtungen und sind durch Merkmale des familialen Hintergrunds und der Anregung begründet (Anders, 2012) und führen dazu, dass sich die Kompetenzniveaus in den Einrichtungen von Beginn an unterscheiden. In weiteren deutschen Studien wurden ebenfalls soziokulturelle Segregationseffekte in der FIBB beobachtet, beispielsweise bezüglich Migrationshintergrund und elterlichem Bildungsabschluss (siehe Kapitel 3.5; Becker, 2010; Becker & Schober, 2015). Problematisch scheinen Segregationseffekte auch, weil die soziokulturelle Zusammensetzung von Gruppen und Einrichtungen ein Bedingungsfaktor der Lern- und Bildungsprozesse in der FIBB zu sein scheint. So wurde in der BiKS-Studie gefunden, dass der Anteil von Kindern mit Migrationshintergrund negativ mit der Prozessqualität zusammenhing (Kuger & Kluczniok, 2008; Kuger et al., 2016; vgl. Tietze et al., 2012) und in EPPSE wurde ein Zusammenhang zwischen mütterlichem Bildungsabschluss und mathematischer Kompetenzentwicklung in Einrichtungen beobachtet (Sammons et al., 2002).

Effektivitätsanalysen sind für weitere deutsche Studien durchführbar, z. B. NEPS (Bäumer et al., 2013; Durda, 2015) oder NUBBEK (Tietze et al., 2012). Bei den von Leyendecker, Agache und Madsen (2014) berichteten signifikanten *ICCs* in NUBBEK für Kompetenzen in anderen Entwicklungsbereichen handelt es sich um einfache Niveauvergleiche und damit in Anbetracht der TP2 Ergebnisse vermutlich um keine fairen Effektivitätsvergleiche. Auch in NUBBEK ist von soziokulturellen Segregationseffekten auszugehen, welche unterschiedliche Kompetenzniveaus in Einrichtungen auch für andere Entwicklungsbereiche erklären können (Roßbach & Weinert, 2008). Im Falle von Querschnittserhebungen wie NUBBEK sollten daher zumindest entwicklungsrelevante Merkmale berücksichtigt werden. Für die Auswahl kann sich an den zusätzlichen Ergebnissen des TP2 (siehe Anhang C) sowie den EPPSE-Befunden

(Sammons et al., 2002) und der allgemeinen Evidenz zur Entwicklungsrelevanz von Merkmalen orientiert werden (z. B. Anders et al., 2013; Anders et al., 2012; Kaufmann, 2011; Schneider et al., 2013; Sylva et al., 2006).

Die BiKS-Effektivitätsbefunde weichen von den EPPSE-Befunden ab, wo mathematische Effektivitätsunterschiede in fairen Vergleichen beobachtet wurden (kontextualisierte, Zuwachsvergleiche; Sammons et al., 2002; Sylva et al., 2010). Ein Grund hierfür könnte in FIBB-Spezifika beider Länder liegen: Beispielsweise gilt das englische System als stärker lernzielorientiert und die offizielle Implementierung mathematischer Bildung im nationalen Curriculum erfolgte früher (Kwon, 2002; OECD, 2006). Weitere mögliche Gründe für divergierende Befunde liegen in unterschiedlichen Forschungsdesigns beider Studien (siehe Anhang C für Limitierungen der BiKS-Effektivitätsanalysen).

Wichtige Diskrepanzen deuten sich ebenfalls in Bezug auf die Befunde zur Prozessqualität in BiKS an, denn es fanden sich geringe Unterschiede der Prozessqualität von Einrichtungen (Kuger & Kluczniok, 2008), welche mit der mathematischen Entwicklung zusammenhängen (siehe Kapitel 6.2; Anders et al., 2013; Anders et al., 2012; Lehl et al., 2016). Lehl et al. (2014) fanden auch keinen Nachweis für eine herkunftsbedingte Zuweisung zur bereichsspezifischen Prozessqualität: Bildungsbenachteiligte Kinder befanden sich in Einrichtungen vergleichbarer Qualität wie andere Kinder. Während in bisherigen Publikationen Einschätzungen zum prozessualen Geschehen in der Gesamtgruppe vorgenommen und Zusammenhänge zu individuellen Entwicklungsverläufen der Kinder analysiert wurden, werden mit Hilfe der Effektivitätsanalysen Aussagen zum Bildungsergebnis FIBB anhand der durchschnittlichen Bildungsergebnisse der getesteten Kinder einer Gruppe getroffen, welche mit Mehrebenenanalysen und nicht Wachstumsmodellen umgesetzt wurden. In der Gesamtschau wird daher deutlich, dass es sich bei Prozessqualität und Bildungseffektivität um unterschiedliche, sich ergänzende Komponenten zur Beurteilung der mathematischen Bildung handelt. Implikationen für die Forschung und Praxis FIBB lassen sich daher besser in einer Gesamtdiskussion der Evidenz zur mathematischen Effektivität und Prozessqualität FIBB treffen (siehe Kapitel 11).

10 Zentrale fachkraftbezogene Komponenten der Struktur- und Orientierungsqualität in der FIBB in Deutschland und ihr Zusammenhang (TP3)

Das TP3 widmet sich in einer theoretischen und empirischen Auseinandersetzung zentralen fachkraftbezogenen Qualitätskomponenten der FIBB-Regelversorgung in Deutschland (siehe Abbildung 4), über die bislang wenig bekannt ist. Obschon fachkraftbezogene Komponenten der Orientierungs- und Strukturqualität mathematischer Bildung mittlerweile als Fundament für die Prozessqualität und den Bildungoutput FIBB gelten, sind sie vergleichsweise unscharf konzeptualisiert und unterforscht (siehe Kapitel 5.3 und 6.3). Dabei wird im TP3 an mehreren Stellen auf Beiträge anderer Forschungsrichtungen zurückgegriffen, insbesondere der Professions- und Professionalisierungsforschung sowie der Motivations- und Schulforschung, um fachkraftbezogene Qualitätskomponenten auszdifferenzieren und die Rolle der Komponenten für die pädagogische Arbeit abzuleiten. Zu Beginn werden im Kapitel 10.1 die allgemeine Qualifikation sowie der Professionalisierungshintergrund und die Schulerfahrungen in Mathematik als Strukturkomponenten thematisiert. Anschließend geht es im Kapitel 10.2 um die Fachkraftorientierungen als Facette professioneller Kompetenz. Zudem werden die Wertorientierung, Selbstwirksamkeit und das Selbstkonzept in Mathematik als besonders handlungsrelevante Orientierungen behandelt. Nachfolgend wird im Kapitel 10.3 auf den Zusammenhang zwischen den untersuchten Struktur- und Orientierungskomponenten eingegangen. Ein Überblick über die explorative Studie sowie das methodische Vorgehen wird in den Kapiteln 10.4 und 10.5 geboten. Die Beschreibung und Diskussion relevanter Befunde folgen in Kapitel 10.6 und 10.7. Weitere Details zur Studie und Ergebnisse befinden sich im Anhang D.

10.1 Strukturkomponenten – Qualifikation und Professionalisierung in Mathematik

Obschon als zentrale strukturelle Qualitätskomponente seit einiger Zeit diskutiert (siehe Kapitel 5.3 und 6.3; Anders, 2012; Aktionsrat Bildung, 2012; Roux & Tietze, 2007), verfügen die meisten frühpädagogischen Fachkräfte hierzulande nicht über einen Hochschulabschluss, anders als Lehrkräfte des Schulwesens oder frühpädagogische Fachkräfte anderer Länder (Cross et al., 2009; EACEA/Eurydice, 2009; OECD, 2015b). Die Mehrzahl der Fachkräfte hat eine dreijährige Ausbildung zur Erzieherin/zum Erzieher an Fachschulen absolviert. Im Zeitraum von 2006 bis 2013 stieg allerdings die Anzahl der Fachkräfte mit Hochschulabschluss von 3.2 % auf 4.9 % (Statistisches Bundesamt, 2013). Aufgrund der relativ späten offiziellen

Implementierung der Mathematik als Bildungsbereich und der unzureichenden Einbettung mathematikspezifischer Angebote in die Aus- und Fortbildung (Baumeister & Grieser, 2011; Behr & Walter, 2012; Deppe, 2011; Gasteiger, 2017; Janssen, 2010; Ledig, 2011; OECD, 2011; Sylva et al., 2015) ist davon auszugehen, dass die mathematikspezifische Professionalisierung der Fachkräfte begrenzt ist, wenngleich Evidenz zum mathematikspezifischen Aus- und Fortbildungshintergrund der Fachkräfte in Deutschland weitgehend fehlt (siehe Kapitel 6.3). Fried (2011) berichtete, dass rund 64 % von 367 niedersächsischen Fachkräften angaben, an Fortbildungen zur mathematischen Grundbildung teilgenommen zu haben, weitaus weniger als zum Thema Sprache und Sprechen (rund 85%). Vermutlich fußen mathematische Erfahrungen vieler Fachkräfte daher größtenteils auf Schulerfahrungen und werden ggfs., durch Erfahrungen während der täglichen pädagogischen Arbeit in den Einrichtungen ergänzt. Um den enormen Anforderungen an die professionelle Kompetenz gerecht zu werden, benötigen Fachkräfte allerdings eine ausreichende mathematikspezifische Qualifizierung und Professionalisierung, welche wirksam und umfassend die für die mathematische Bildungsarbeit notwendigen Kompetenzen adressieren (Anders, 2012; Chen & Chang, 2006; Chen & McCray, 2012).

10.2 Orientierungskomponenten – Motivationale Orientierungen

Modelle professioneller Kompetenz für den frühpädagogischen Bereich (z. B. Anders, 2012; Fröhlich-Gildhoff, Nentwig-Gesemann & Pietsch, 2011) zeigen auf, welches breites Portfolio professioneller Kompetenz für eine erfolgreiche pädagogische Arbeit erforderlich ist wie verschiedene Wissenskomponenten, professionelle Haltungen und selbstregulatorische Fähigkeiten. Hierzu gehören ebenso die Orientierungen sowie motivationale und emotionale Aspekte. Fachkräfte benötigen neben bereichsübergreifenden Kompetenzen spezifische Facetten in den verschiedenen Bildungsbereichen (Aktionsrat Bildung, 2012; Chen & Chang, 2006), z. B. mathematikspezifische Orientierungen (Dunekacke et al., 2013; Schuler et al., 2017).

Die Schulforschung bietet differenzierte Taxonomien von Orientierungen und schlüsselt die verschiedenen Funktionen einzelner Orientierungen für das Fachkräftehandeln auf (siehe Abbildung 6; Fives & Buehl, 2012; Pajares, 1992). Als unmittelbar handlungsrelevant gelten die Selbstwirksamkeit und die Wertorientierung von Fachkräften wegen ihrer motivierenden Funktion, weshalb sie teilweise auch als motivationale Orientierungen bezeichnet werden (Fives & Buehl, 2012; Wigfield & Eccles, 2002). In Erwartungs-Wert-Modellen der Motivationspsychologie wird drauf hingewiesen, dass die Kombination beider Orientierungen über die individuelle Motivation entscheidet (z. B. Eccles et al., 1983; Krapp & Ryan, 2002; Wigfield,

1994; Wigfield & Eccles, 2000) und beide Orientierungen in direkter Verbindung zum Selbstkonzept eigener Fähigkeiten stehen (Wigfield & Eccles, 2000).

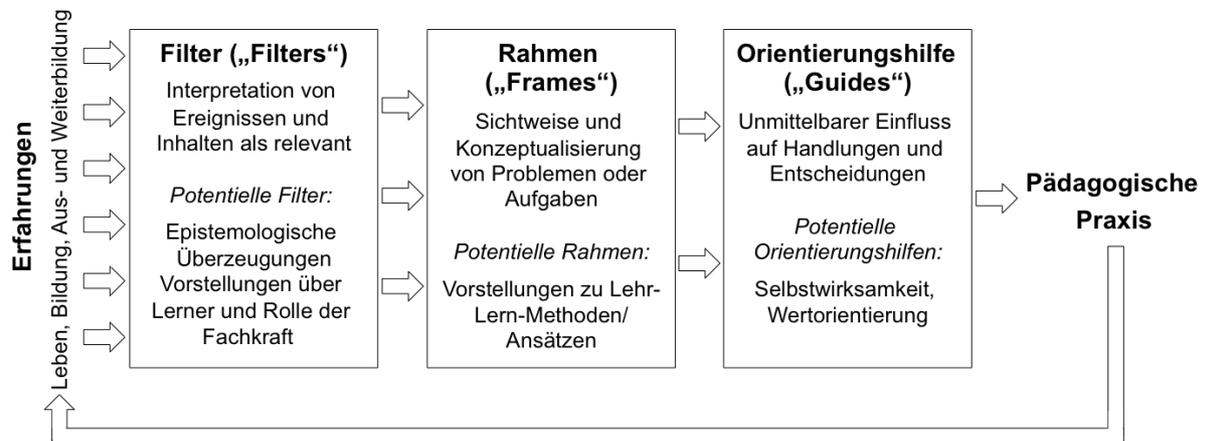


Abbildung 6. Die Funktionen von Orientierungen als Filter, Rahmen und Orientierungshilfen für die pädagogische Praxis von Fachkräften (nach Fives & Buehl, 2012).

Die *mathematische Wertorientierung* bezieht sich auf den Nutzen und Wert, den der Mathematik zugeschrieben wird (Eccles et al., 1983). Mathematische Selbstwirksamkeit und mathematisches Selbstkonzept stellen beides Einschätzungen eigener mathematischer Fähigkeiten dar (Bandura, 1986; Krapp & Ryan, 2002), jedoch gibt es wichtige Unterschiede (Bong & Skaalvik, 2003; Ferla, Valcke & Cai, 2009; Pajares, 1997). *Das mathematische Selbstkonzept* beschreibt die Selbstzuschreibungen eigener mathematischer Kompetenzen einer Person, also die Überzeugung in Mathematik gut oder schlecht zu sein. Es speist sich aus zurückliegenden Erfahrungen mit der Mathematik, repräsentiert das gesammelte Wissen über die eigenen Fähigkeiten und ist vergleichsweise stabil. Die *mathematische Selbstwirksamkeit* bezeichnet die Überzeugung von Personen spezifische mathematische Aufgaben oder Situationen erfolgreich lösen oder bewältigen zu können. Statt einer generellen Einschätzung mathematischer Kompetenz auf Grundlage vergangener Erfahrungen richten sich Einschätzungen der Selbstwirksamkeit auf das Meistern konkreter, zukünftiger Herausforderungen, z. B. das Lösen mathematischer Aufgaben. Durch die stärkere Veränderbarkeit ist die Selbstwirksamkeit für die Professionalisierungsforschung besonders interessant.

Motivationale Orientierungen von Fachkräften können sich sowohl auf die Mathematik selbst als auch auf die mathematische Bildungsaufgabe beziehen, beispielsweise die Selbstwirksamkeit von Fachkräften, welche sich auf mathematische Lehr- bzw. Bildungskompetenz bezieht oder auf die Wertorientierung über Mathematik als wichtigen Bereich früher Bildung

(z. B. Anders & Roßbach, 2015; Chen et al., 2014; Cobanoglu, 2011; Guo et al., 2011; Tsamir, Tirosh, Levenson, Tabach & Barkai, 2014; Vural & Gunhan, 2011).

Bezüglich der *Evidenz zu motivationalen Orientierungen* frühpädagogischer Fachkräfte werden Diskrepanzen zwischen dem Forschungsstand in den USA und in Deutschland deutlich. In den USA ist die Annahme verbreitet, dass frühpädagogische Fachkräfte allgemein negative mathematische Orientierungen besitzen und dass sie der frühen mathematischen Förderung große Vorbehalte entgegenbringen (Baroody et al., 2006; Bates, Latham & Kim, 2013; Copley, 2004; Copley & Padrón, 1999; Cross et al., 2009; Sweeting, 2011). Außerdem wird von einem geringen Vertrauen in die mathematischen Fähigkeiten und die eigene mathematische Bildungsarbeit sowie einer hohen Mathematikangst der Fachkräfte berichtet. Die zumeist ältere Evidenz basiert jedoch mehrheitlich auf kleinen qualitativen Studien und die Befragungen fokussieren vornehmlich auf negative Orientierungen und Ängste, wohingegen neuere, quantitative Studien zu positiveren Ergebnissen der motivationalen Überzeugungen frühpädagogischer Fachkräfte kommen (Chen et al., 2014; Sweeting, 2011).

In aktuelleren deutschen Studien wurde ebenfalls gefunden, dass Fachkräfte generell offen gegenüber Mathematik waren und allenfalls von gemischten, aber nicht überwiegend negativen mathematischen Orientierungen berichteten (Benz, 2012a, 2012b; Fried, 2011, 2012; Thiel, 2010). Fachkräfte nahmen Mathematik als wichtigen Bereich früher Bildung wahr (Benz, 2012b; Thiel, 2010), allerdings hauptsächlich für Kinder ab drei Jahren (Anders & Roßbach, 2015) und ordneten anderen Bildungsbereichen, z. B. Sprache/Literacy und sozio-emotionalen Bereichen einen höheren Stellenwert zu (Broekhuizen et al., 2015; Kowalski, Pretti-Frontczak & Johnson, 2001; Mischo, Wahl, Hendler, et al., 2012; Tietze et al., 1998; Tietze, Roßbach, et al., 2005; Wilkins, 2009).

Für den angenommenen Zusammenhang zwischen motivationalen und weiteren Fachkraftorientierungen und pädagogischem Handeln (Bandura, 1986; Fives & Buehl, 2012; Pajares, 1992; Wigfield & Eccles, 2000) gibt es bislang wenige Studien, in denen gemischte Befunde berichtet werden. Thiel (2010) berichtete beispielsweise, dass deutsche Fachkräfte mit hoher mathematischer Wertorientierung eher ein konstruktivistisches Verständnis pädagogischer Arbeit besitzen, welches aktives Explorieren und gemeinsames Problemlösen in mathematischen Lernprozessen betont (vgl. Benz, 2012b). Ähnliches berichtete Cobanoglu (2011) für eine hohe Selbstwirksamkeit bezüglich mathematischer Bildungsarbeit, während Brown (2003, 2005) in zwei Studien, in denen pädagogisches Handeln von US-Fachkräften beobachtet wurde, keinen Zusammenhang mit Wertorientierung oder Selbstwirksamkeit bezüglich der eigenen Bildungsarbeit fand.

10.3 Zusammenhang zwischen den Struktur- und Orientierungskomponenten

Trotz der vergleichsweisen hohen Stabilität (Pajares, 1992) weisen Befunde darauf hin, dass sich mathematische Orientierungen von Fachkräften in Abhängigkeit von Erfahrungen mit Mathematik und mathematischer Bildung verändern können (Fives & Buehl, 2012) und dass die behandelten fachkraftbezogenen Strukturkomponenten Bedingungsfaktoren der Orientierungen darstellen können.

10.3.1 Mathematische Schulerfahrungen als Bedingungsfaktoren

Grundlegende Erfahrungen mit Mathematik sammeln Fachkräfte bereits in ihrer Schulzeit. Insbesondere affektiv-emotionale Erfahrungen sind für motivationale Orientierungen in der Schule von Relevanz (Wigfield & Eccles, 2000, 2002). Zudem besteht ein enger Bezug zur Mathematikmotivation, v. a. dem Interesse. Es wird angenommen, dass Schulerfahrungen eine bleibende Wirkung auch über die Schulzeit hinaus auf Fachkraftorientierungen entfalten (Pajares, 1992; Relich, 1996; van Oers, 2002; Wilkins, 2008). Überwiegend negative Schulerfahrungen werden als Hauptursache der negativen mathematischen Orientierungen und Mathematikangst von US-Fachkräften diskutiert (Baroody, 2004; Bates et al., 2013; Brady & Bowd, 2005; Copley, 2004; Copley & Padrón, 1999; Kelley & Tomhave, 1985; Sweeting, 2011; Tobias, 1987).

In vergleichsweise wenigen Studien wurde sich der positiven Emotionen und dem Interesse an Mathematik frühpädagogischer Fachkräfte gewidmet, deren Rolle für die pädagogische Arbeit von Lehrerinnen und Lehrern hervorgehoben wird (z. B. Baumert & Kunter, 2006; vgl. Anders, 2012; Aktionsrat Bildung, 2012). In Studien zu deutschen Fachkräften wurde allerdings von gemischten Emotionen bezüglich Mathematik berichtet (Benz, 2012a, 2012b; Thiel, 2010). Thiel (2010) beobachtete zudem für deutsche Fachkräfte, dass die Offenheit gegenüber Mathematik mit einer höheren Wertorientierung verbunden war. Über die zurückliegenden Schulerfahrungen deutscher frühpädagogischer Fachkräfte ist bislang wenig bekannt. Anders und Roßbach (2015) fanden heraus, dass deutsche Fachkräfte mit der Schulmathematik nicht übermäßig negative Emotionen, ebenso wenig jedoch positive Emotionen verbanden und dass Schulemotionen nicht mit der Wertorientierung bezüglich der Bedeutung früher mathematischer Bildung zusammenhängen.

10.3.2 Qualifikation und mathematikspezifische Professionalisierung als Bedingungsfaktoren

Im Anschluss an die Schule sollen Fachkräfte durch Qualifikations- und Professionalisierungsmaßnahmen die nötigen professionellen Kompetenzen für die mathematische Bildung erwerben, zu denen positive Fachkraftorientierungen gehören. Für den Zusammenhang zwi-

schen Qualifikation und Orientierungen lieferte Vartuli (1999) Evidenz: Fachkräfte mit einer frühpädagogischen Spezialisierung besaßen Orientierungen, welche besser zur kindlichen Entwicklung passten (sogenannte *Beliefs of Developmentally Appropriate Practices* bzw. DPA). Wegen des geringen Anteils akademischer Abschlüsse frühpädagogischer Fachkräfte in Deutschland existieren hierzu bislang kaum Studien. Mischo, Wahl, Hendler, et al. (2012) berichteten jedoch, dass deutsche frühpädagogische Hochschulabsolventinnen und -absolventen weniger instruktive Orientierungen aufwiesen und die Bildungsfunktion FIBB stärker betonten als Fachkräfte mit der üblichen Fachschulausbildung.

Evidenz zur besonderen Bedeutung mathematikspezifischer Aus- und Fortbildungsangebote für mathematische Orientierungen (Anders, 2012; Chen & McCray, 2012; Gasteiger, 2010, 2012) stammt vor allem aus Evaluationsstudien von Fachkräftetrainings sowie Qualitätsoffensiven und -initiativen, wobei sich die Befunde mehrheitlich auf eine gesteigerte mathematische Selbstwirksamkeit beziehen (z. B. Akay & Boz, 2010; Albayrak & Unal, 2011; Brand & Wilkins, 2007; Evans et al., 2013; Kalder & Lesik, 2011; Moseley & Utley, 2006; Quinn, 1997; Rosenfeld, 2010; Saçkes, Flevaris, Gonya & Trundle, 2012; Swars, 2005; Tirosh et al., 2011; Wilkins & Brand, 2004). Belege gibt es allerdings auch für die erfolgreiche Steigerung von Wertorientierung (Kalder & Lesik, 2011; Quinn, 1997; Sherman & Christian, 1999; Wilkins, 2008) und das Selbstkonzept in Mathematik (Palmer, 2009). Viele Befunde beziehen sich allerdings auf Fachkräfte der Grundschule. Für den Zusammenhang zwischen mathematikspezifischen Angeboten im Rahmen der Professionalisierung in der FIBB-Regelversorgung auf Fachkraftorientierungen fehlen Befunde weitgehend (vgl. Darling-Hammond, Chung & Frelow, 2002 für schulische Lehrkräfte). Spezifische Studien für die Effekte von Professionalisierungsangeboten sind allerdings notwendig. So kritisieren Cross et al. (2009), dass die Qualität regulärer Professionalisierungsangebote in den USA enorm schwankt. Für Deutschland ist insgesamt von einer geringeren mathematikspezifischen Professionalisierung auszugehen, wobei ebenfalls große Schwankungen im Ausmaß und in der Qualität der mathematikspezifischen Professionalisierungshistorie von Fachkräften der FIBB-Regelversorgung anzunehmen sind (Baumeister & Grieser, 2011; Behr & Walter, 2012; Deppe, 2011; Gasteiger, 2017; Janssen, 2010; Ledig, 2011).

Neben dieser Erfahrung durch formelle Angebote in der Schule bzw. weiterführende Ausbildungsinstitute, sammeln Fachkräfte u. U. informelle Erfahrungen mit Mathematik und mathematischer Bildung im Rahmen der täglichen Arbeit. Im Arbeitskontext sind Fachkräfte kontinuierlich mit Erwartungen und Feedback zu ihrer pädagogischen Arbeit von Kindern, Eltern und Kollegen konfrontiert, was Orientierungen der Fachkräfte prägen kann, z. B. was

sie über das kindliche Lernen, Bildungsinhalte, effektive Bildungsmethoden sowie über sich selbst und ihre eigenen Fähigkeiten denken. Für die Berufserfahrung ergab sich in einer Stichprobe türkischer Fachkräfte ein schwacher Zusammenhang zur Selbstwirksamkeit (Cobanoglu, 2011). Tirosh et al. (2011) konnten zeigen, dass praktizierende Fachkräfte überzeugter waren verschiedene Mathematikaufgaben lösen zu können als Fachkräfte in der Ausbildung. McMullen (1997) hingegen fand keinen Zusammenhang von Berufserfahrung und pädagogischer Selbstwirksamkeit von Fachkräften, aber einen Zusammenhang mit entwicklungsangemessenen Handlungsorientierungen bzw. Developmentally Appropriate Practices (siehe aber Guo et al., 2011; Vartuli, 1999).

10.4 Beschreibung der Studie zu fachkraftbezogenen Qualitätskomponenten des TP3

Im TP3 werden Daten der 2011-2012 durchgeführten Studie Muster, Formen, Raum und Zahlen im Kindergarten entdecken – Pädagogische Einstellungen von frühpädagogischen Fachkräften zu mathematischer Frühförderung verwendet, um neue Erkenntnisse zu fachkraftbezogenen Komponenten der Struktur- und Orientierungsqualität mathematischer Bildung in der deutschen FIBB-Regelversorgung zu erlangen (vgl. Anders & Roßbach, 2015; Oppermann et al., 2016). In der explorativen Studie werden diesbezüglich zwei Ziele verfolgt:

1. Deskriptives Profil der Struktur- und Orientierungsqualität: Als Strukturkomponenten wird ein Profil der (a) mathematischen Schulerfahrungen (positive und negative Emotionen, Interesse und Vermeidung) und zudem der (b) generellen Qualifikation und mathematikspezifischen Professionalisierung der Fachkräfte angestrebt. Zudem werden (c) die motivationalen mathematischen Orientierungen der Fachkräfte (Selbstwirksamkeit, Selbstkonzept und Wertorientierung) als besonders handlungsrelevante Komponenten der Orientierungsqualität betrachtet.

2. Zusammenhänge zwischen Struktur- und Orientierungsqualität: Darüber hinaus sollen in der Studie die Zusammenhänge zwischen den mathematischen Orientierungen und (a) der Schulerfahrung für deutsche Fachkräfte der FIBB-Regelversorgung erforscht werden. Anschließend wird untersucht, ob die (b) Qualifikation und Professionalisierung über die mathematische Schulerfahrung hinaus Zusammenhänge zu den aktuellen Orientierungen aufweisen.

10.5 Methode der Studie zu den fachkraftbezogenen Qualitätskomponenten des TP3

10.5.1 Stichprobe und Durchführung

Die *Stichprobe* besteht aus 221 fröhpädagogischen Fachkräften (91.4 % weiblich) aus 29 Kindertagesstätten aus Berlin (54.8 %) und Bayern (45.2 %). Im Durchschnitt waren die Fachkräfte 40 Jahre alt ($SD = 11.07$) mit einer Spannweite von 18 bis 63 Jahren. Die Teilnahme an der Studie, welche durch geschulte Erheberinnen und Erheber in den Kindertagesstätten durchgeführt wurde, war freiwillig und bestand aus dem Ausfüllen von Fragebögen und Testverfahren zum mathematischen Wissen.

Analytisch wurde für die Zielsetzung 1 über Deskriptiva und Korrelationen ausgewertet. Die Beantwortung der Zielsetzung 2 erfolgt mit blockweisen, multiplen Regressionen in *Mplus* (Version 7, weitere Details im Anhang D). In den Analysen wurde die Full Information Maximum Likelihood-Methode zur Schätzung fehlender Werte (FIML; Enders, 2001; Lüdtke et al., 2007) und einen robusten Standardfehler (type = complex) verwendet, um die hierarchische Struktur der Daten zu berücksichtigen (Fachkräfte gruppieren sich in verschiedenen Einrichtungen).

10.5.2 Instrumente

Für die explorative Studie wurden die Angaben aus Skalen zu vergangenen Schulerfahrungen und aktuellen Orientierungen verwendet, die als Mittelwerte in die Analysen eingingen (von 1 = *gar nicht zutreffend* bis 4 = *voll zutreffend*). Tabelle 5 gibt einen deskriptiven Überblick, ein Beispielitem und beinhaltet Angaben zur internen Konsistenz errechnet über *Crohnbachs α* , welche insgesamt zufriedenstellend ausfiel ($\alpha = .66 - .89$; siehe Anhang D für eine detailliertere Beschreibung der Instrumente).

Der Fragebogen zu den (a) *mathematischen Schulerfahrungen* bat Fachkräfte an ihre Schulzeit zurückzudenken und verschiedene Items in Vergangenheitsform zur Emotion und Motivation gegenüber Mathematik am Ende der Schulzeit zu beantworten. Fachkräfte gaben an, wie sehr sie eine Reihe positiver und negativer Emotionen mit dem Mathematikunterricht verbanden. Zudem enthielt der Fragebogen Items zum Interesse an den Inhalten der Schulmathematik und der Tendenz den Mathematikunterricht zu vermeiden.

Zusätzlich wurden im TP3 Einzelitems mit spezifischem Antwortformat und Dummyvariablen zur (b) *Qualifikation und Professionalisierung* der Fachkräfte ausgewertet. Eine binäre Dummyvariable gab an, ob Fachkräfte die Schule mit oder ohne Abitur verlassen hatten (0 = *ohne* und 1 = *mit Abitur*) und diente als Proxyvariable für eine höhere generelle Qualifikation der Fachkräfte, weil der Anteil an Fachkräften mit einem (Fach-) Hochschulabschluss mit we-

niger als 4 % sehr niedrig war. Zwei weitere Dummyvariablen fassten die Antworten der Fachkräfte auf verschiedene Items zur Teilnahme an mathematikspezifischen Angeboten in der Aus- und Fortbildung zusammen (0 = *kein* und 1 = *mindestens ein Angebot*). Die Berufserfahrung in Jahren diente als Indikator für die Möglichkeit informell Erfahrungen mit Mathematik oder mathematischer Bildung während der pädagogischen Arbeit zu sammeln.

Die Instruktionen und die Formulierung der Items zu den (c) *mathematischen Orientierungen* erfolgten im Präsens und basierten auf verschiedenen Vorarbeiten (Benz, 2012b; Ramm et al., 2006; Thiel, 2010). Die Skala mathematische Wertorientierung erfasste, inwiefern die Fachkräfte Mathematik als wichtig und nützlich für das alltägliche Leben und die Gesellschaft insgesamt betrachteten. Die mathematische Selbstwirksamkeit erfasste, inwiefern sich Fachkräfte sicher waren, verschiedene mathematische Aufgaben lösen zu können. Die mathematische Selbstkonzeptskala gab Auskunft darüber, wie Fachkräfte ihre generellen mathematischen Fähigkeiten und Leistungen im Bereich Mathematik wahrnahmen.

10.6 Ergebnisse zu den fachkraftbezogenen Qualitätskomponenten des TP3

10.6.1 Profil der (a) Schulerfahrungen, (b) der Qualifikation und Professionalisierung sowie (c) der mathematischen Orientierungen

Die Deskriptiva in Tabelle 5 der (a) *Schulerfahrungen* ergeben für Mathematikinteresse und -vermeidung ein ähnliches Ergebnis wie bereits für positive und negative Emotionen berichtet (Anders & Roßbach, 2015). Teilnehmerinnen und Teilnehmer berichteten keine ausgeprägt negativen Emotionen oder starke Mathematikvermeidung ($M = 1.92$, $SD = .81$ und $M = 1.87$, $SD = .71$). Allerdings gaben sie ebenfalls keine übermäßig positiven Emotionen oder starkes Mathematikinteresse an ($M = 1.98$, $SD = .78$ und $M = 2.19$, $SD = .66$). Die Mittelwerte positiver und negativer Emotionen unterschieden sich nicht ($t(196) = 0.76$, $p = \text{n.s.}$), aber das Mathematikinteresse fiel im Durchschnitt höher aus als die Mathematikvermeidung ($t(209) = 4.19$, $p < .001$). Damit sind weder die emotionalen noch die motivationalen Erfahrungen der Fachkräfte so schlecht, wie von vielen Forschern für US-Fachkräfte beschrieben (Baroody, 2004; Baroody et al., 2006; Bates et al., 2013; Copley, 2004; Copley & Padrón, 1999; Hembree, 1999).

Tabelle 5

Deskriptive Statistik der Schulerfahrungen und mathematischen Orientierungen mit Likert-Skalierung von 1 bis 4

Skalen	<i>M</i>	<i>SD</i>	<i>N</i>	α	Items	Beispielitem
Mathematische Schulerfahrungen						
Positive Emotionen	1.98	0.78	203	0.89	3	“Geben Sie bitte an, mit welchen Gefühlen die Erinnerung an Ihren eigenen Mathematikunterricht verbunden ist.“ (z. B. „Freude“, „Spaß“)
Mathematikinteresse	2.19	0.66	215	0.66	3	“Ich wollte in Mathematik nicht nur Sachen auswendig lernen, sondern sie auch wirklich verstehen.”
Negative Emotionen	1.92	0.82	203	0.88	4	“Geben Sie bitte an, mit welchen Gefühlen die Erinnerung an Ihren eigenen Mathematikunterricht verbunden ist.“ (z. B. „Angst“, „Frustration“)
Mathematikvermeidung	1.87	0.71	212	0.76	3	“Es gab Zeiten, da wollte ich aus Angst vor Mathematik nicht in die Schule gehen.“
Mathematische Orientierungen						
Wertorientierung	2.89	0.51	197	0.73	6	“Mathematik hilft, alltägliche Aufgaben und Probleme zu lösen.”
Selbstwirksamkeit	3.1	0.67	211	0.75	5	“Wie sicher glauben Sie, folgende Mathematikaufgaben lösen zu können? Ausrechnen, wie viel billiger ein Fernseher bei 30% Rabatt wäre.”
Selbstkonzept	1.83	0.76	211	0.91	4	“Ich war schon immer überzeugt, dass Mathematik eine meiner besonderen Stärken ist.”

10.6.2 Zusammenhänge der mathematischen Orientierungen mit den (a) Schulerfahrungen sowie der (b) Qualifikation und Professionalisierung

Tabelle 6 zeigt die standardisierten Regressionsgewichte der blockweisen Regression für die drei Orientierungen getrennt, in die Schulerfahrungen zuerst als Prädiktorblock einging. Nicht alle untersuchten (a) *Schulerfahrungen* erwiesen sich als prädiktiv und die Höhe der Zusammenhänge variierte für die verschiedenen Orientierungen. Positive Emotionen hingen signifikant positiv mit der mathematischen Wertorientierung und dem mathematischen Selbstkonzept zusammen ($\beta = .24$ und $\beta = .35$, $p < .001$). Das Mathematikinteresse war ein signifikanter Prädiktor für das Selbstkonzept und ein marginal signifikanter Prädiktor der Wertüberzeugung in Mathematik ($\beta = .25$, $p < .01$ und $\beta = .14$, $p < .10$). Für die negativen Schulerfahrungen ergaben sich lediglich für die Mathematikvermeidung signifikante Zusammenhänge, nicht für negative Emotionen. Die Mathematikvermeidung hing signifikant mit beiden Fähigkeitsüberzeugungen zusammen ($\beta = -.35$, $p < .01$ für mathematische Selbstwirksamkeit und $\beta = -.50$, $p < .001$ für mathematisches Selbstkonzept). Das Hinzufügen des zweiten Prädiktorblocks führte nur zu geringfügigen Änderungen der beschriebenen Koeffizienten.

Der zusätzliche Prädiktorblock zur (b) *Qualifikation und Professionalisierung* der Fachkräfte erklärte insgesamt 6 % bzw. 7 % der Varianz für mathematische Wertorientierung und Selbstwirksamkeit. Formale Qualifikation und Professionalisierung hing ausschließlich mit der Selbstwirksamkeit zusammen: Fachkräfte mit einer höheren generellen Qualifikation (Abitur oder Hochschulzugangsberechtigung) und Fachkräfte, welche an mathematikspezifischen Angeboten während der Ausbildung oder dem Studium teilnahmen, berichteten eine signifikant bzw. marginal signifikant höhere Selbstwirksamkeit ($\beta = .26$, $p < .001$ bzw. $\beta = .11$, $p < .10$). Die Berufserfahrung als informelle Möglichkeit mathematischer Bildungserfahrung war ein positiver Prädiktor mathematischer Wertorientierung ($\beta = .20$, $p < .01$) und mathematischer Selbstwirksamkeit ($\beta = .16$, $p < .05$). Keine Zusammenhänge der einzelnen Qualifikationen und Professionalisierungsaspekte ergaben sich zum mathematischen Selbstkonzept (insgesamt nur 1% zusätzliche Varianzaufklärung).

Tabelle 6

Unstandardisierte Regressionskoeffizienten und dazugehörigen Standardfehler für die Vorhersage von mathematischen Orientierungsfahrungen sowie Qualifikation und Professionalisierung

Prädiktoren	Math. Wertorientierung		Math. Selbstwirksamkeit		Math. Selbstkonzept	
	Schritt a	Schritt b	Schritt a	Schritt b	Schritt a	Schritt b
Schulerfahrungen						
Positive Emotionen	0.15*** (0.04)	0.18*** (0.04)	-0.01 (0.08)	0.03 (0.08)	.29*** (0.06)	.29*** (0.07)
Mathematikinteresse	0.11# (0.05)	0.09# (0.05)	0.09 (0.07)	0.06 (0.07)	.24** (0.07)	.25*** (0.06)
Negative Emotionen	0.05 (0.05)	0.05 (0.06)	-0.10 (0.08)	-0.10 (0.06)	-.03 (0.05)	-.03 (0.06)
Mathematikvermeidung	-0.07 (0.06)	-0.04 (0.07)	-0.32** (0.11)	-0.26** (0.10)	-.46*** (0.06)	-.45*** (0.06)
Qualifikation / Professionalisierung						
Abitur ^a		-0.03 (0.06)		0.34*** (0.09)		-0.02 (0.07)
Math. Kurse in der Ausbildung ^b		0.09 (0.07)		0.16# (0.09)		-0.02 (0.07)
Math. Kurse in der Fortbildung ^b		0.09 (0.08)		-0.06 (0.10)		-0.06 (0.08)
Berufserfahrung in Jahren		0.01** (0.00)		0.01* (0.00)		0.00 (0.00)
R^2	.10**	.16***	.15*	.22**	.44***	.43***
ΔR^2		.06		.07		.01

Anmerkungen. $N = 203 - 218$. # $< .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (zweiseitig). Referenzgruppen der binären Variablen sind ^akein/e Abitur/Hochschulzugangsberechtigung und ^bkeine mathematikspezifischen Kurse.

10.7 Diskussion zu fachkraftbezogenen Qualitätskomponenten des TP3

Das TP3 widmete sich fachkraftbezogenen Komponenten der Struktur- und Orientierungsqualität mathematischer Bildung. Die Ergebnisse zur allgemeinen Qualifikation wie die Berufserfahrung und der geringe Anteil der untersuchten Fachkräfte mit Hochschulabschluss und Abitur decken sich weitgehend mit den Befunden anderer Erhebungen (vgl. Statistisches Bundesamt, 2013; Tietze et al., 2012). Die mathematikspezifische Professionalisierung während der Aus- und Fortbildung, über die für die FIBB in Deutschland bislang wenig bekannt ist, erscheint für die untersuchten Fachkräfte unzureichend. Die Ergebnisse stehen im Einklang mit Forderungen nach einem Ausbau des Angebots. Zusätzlich sollte sichergestellt werden, dass Angebote in Anspruch genommen werden, z. B. durch eine Erweiterung der verpflichtenden Anteile in den Rahmenlehrplänen.

Insgesamt scheint sich der formale Erfahrungshorizont mit Mathematik für viele Fachkräfte weitgehend auf die Schulzeit zu beschränken. Zwar gaben Fachkräfte nicht übermäßig negative Schulerfahrungen an, d. h. negative Emotionen oder Vermeidungstendenzen. Dennoch scheint es ihnen an positiven und motivierenden mathematischen Erfahrungen in der Schulzeit zu mangeln, weil Fachkräfte auch nicht von übermäßigem Interesse und positiven Emotionen berichteten. Damit ähneln die Befunde zu motivationalen und emotionalen Schulerfahrungen dem gemischten Muster der Befunde positiver und negativer Orientierungen in anderen deutschen Studien zu Fachkräften (Benz, 2012a, 2012b; Fried, 2011, 2012; Thiel, 2010), aber es ergeben sich wichtige Unterschiede zu Annahmen und Befunden für US-Fachkräfte (Baroody et al., 2006; Bates et al., 2013; Copley, 2004; Copley & Padrón, 1999; Cross et al., 2009; Sweeting, 2011). Das Profil motivationaler Orientierungen untersuchter Fachkräfte zeichnete ein positives Bild der mathematischen Selbstwirksamkeit und Wertorientierung, wobei lediglich das durchschnittliche mathematische Selbstkonzept vergleichsweise niedrig ausfiel.

Entsprechend der Modellannahmen von Eccles et al. (1983; vgl. Wigfield, 1994; Wigfield & Eccles, 2000) und der US-Forschung standen motivationale und emotionale Schulerfahrungen, mit Ausnahme negativer Emotionen, in Beziehung zu den untersuchten Orientierungen, besonders zum Selbstkonzept. Obschon Befunde auf eine bedeutende Rolle der Schulerfahrungen hinweisen, gab es ebenfalls Anzeichen dafür, dass eine ausreichende Qualifikation und Professionalisierung für günstige motivationale Orientierungen wichtig sind (Anders, 2012; Aktionsrat Bildung, 2012; Chen & McCray, 2012; OECD, 2015b; Pianta et al., 2005). Dass eine höhere Qualifikation und mathematische Professionalisierungsangebote in der Ausbildung für mathematische Selbstwirksamkeit eine Rolle spielen, deckt sich mit den zahlreichen

Befunden zur Wirksamkeit verschiedener mathematischer Trainings und Offensiven und Initiativen (z. B. Akay & Boz, 2010; Albayrak & Unal, 2011; Brand & Wilkins, 2007; Evans et al., 2013; Kalder & Lesik, 2011; Moseley & Utley, 2006; Rosenfeld, 2010; Saçkes et al., 2012; Swars, 2005; Wilkins & Brand, 2004) sowie der Evidenz aus Vergleichen der Orientierungen von Fachkräften verschiedener Abschlüsse (Mischo, Wahl, Hendler, et al., 2012; Mischo, Wahl, Strohmer, et al., 2012; Vartuli, 1999). Unklar ist, warum der Zusammenhang auf die Berufsvorbereitungsphase begrenzt blieb und sich nicht ebenfalls auf besuchte Fortbildungsangebote von praktizierenden Fachkräften der FIBB-Regelversorgung bezog. Erwiesen sich doch in Implementations- und Trainingsstudien verschiedene Maßnahmen als förderlich für die Selbstwirksamkeit praktizierender Fachkräfte. Ebenfalls stellt sich die Frage, warum sich keine Beziehung zur mathematischen Wertorientierung ergab, existieren doch Belege für die Beeinflussbarkeit durch Trainings, Offensiven und Initiativen sowie für Unterschiede in Abhängigkeit vom Abschluss. Der fehlende Zusammenhang von Professionalisierungsangeboten und Berufserfahrung zum Selbstkonzept entspricht den Beschreibungen dieses Konstrukts in der Forschungsliteratur als vergleichsweise stabil und wenig beeinflussbar. Dass die Selbstwirksamkeit und Wertorientierung der Fachkräfte positiv mit einer längeren Berufserfahrung zusammenhängen, könnte auf die Rolle von gesammelten Erfahrungen mit der Mathematik und mathematischen Bildung der Fachkräfte im Rahmen der pädagogischen Praxis im Sinne einer informellen/inzidentellen Professionalisierung hindeuten. Vermutlich kann die tägliche pädagogische Arbeit unter bestimmten Umständen motivierende und positive Erfahrungen mit Mathematik und mathematischer Bildung offerieren. Nachzudenken ist daher über Möglichkeiten den pädagogischen Arbeitsalltag und -kontext in frühkindlichen Einrichtungen dementsprechend zu gestalten, z. B. indem die Salienz und die Bedeutung mathematischer Bildung erhöht wird durch eine Verankerung in der Einrichtungskonzeption als Bildungsschwerpunkt oder der Förderung des fachlichen Austauschs der Fachkräfte über die mathematische Bildungsarbeit in Teammeetings, internen Workshops oder in Elterngesprächen. Einige wenige US-Studien geben Aufschluss darüber, welche Merkmale des pädagogischen Arbeitskontexts für positive Fachkräfteorientierungen Relevanz besitzen (Buchanan, Burts, Bidner, White & Charlesworth, 1998; Guo et al., 2011; Upadyaya, Viljaranta, Lerkkanen, Poikkeus & Nurmi, 2011). Über wirksame Ansätze zur Beeinflussung von Orientierungen und motivational-emotionalen Facetten elementarpädagogischer, professioneller Kompetenzen kann aus den erwähnten Evaluationsstudien gelernt werden (z. B. auch Akay & Boz, 2010; Lake & Kelly, 2014; Tooke & Lindstrom, 1998). Für Anpassungen müssen jedoch erst einmal die Inhalte und Methoden mathematikspezifischer Angebote der Aus- und Fortbildung in der deutschen FIBB-

Regelversorgung und deren Wirkung genauer bekannt sein. Generell erweisen sich die Befunde und Theorien der Schulforschung oder der US-amerikanischen Forschung vor dem Hintergrund der vorliegenden Ergebnisse und Befunde weiterer Studien als nur eingeschränkt übertragbar. Zudem gibt es wichtige Spezifika der pädagogischen Arbeit und der Aus- und Fortbildung in der deutschen FIBB.

11 Gesamtdiskussion

Während Beiträge unterschiedlicher Forschungsrichtungen das Bewusstsein für das Bildungspotenzial FIBB gestärkt haben, ist der große Verdienst des Qualitätsansatzes das gestiegene öffentliche und wissenschaftliche Interesse für die Qualität und Bildungsprozesse der FIBB-Regelversorgung. Der Qualitätsansatz hat nützliche Modelle und Messverfahren hervorgebracht, um das Bildungspotenzial der Regelversorgung FIBB zu erforschen, sicherzustellen und zu optimieren. Inzwischen existieren theoretisch fundierte und empirisch abgesicherte Erkenntnisse über verschiedene Komponenten der Qualität und deren Wirkung auf mathematische Bildungs- und Lernprozesse in der FIBB.

Trotz großer Verdienste und steter Erweiterung wurden im Rahmen des Promotionsvorhabens zentrale Forschungsdesiderata ermittelt, weshalb sich in drei TP theoretisch und empirisch mit den Komponenten und Auswirkungen der Qualität mathematischer Bildung in der FIBB-Regelversorgung auseinandergesetzt wurde. Kapitel 11 wertet die Erkenntnisse zusammenfassend aus. Das Kapitel 11.1 beginnt mit einem Überblick über die Befunde aus den TP und schließt mit einer Reflexion darüber, ob mit der Bündelung von Evidenz unterschiedlicher Studien die Limitationen der einzelnen Studien der TP überwunden werden konnten. Anschließend erörtert Kapitel 11.2 zuerst die Forschungsimplicationen für die weitere Entwicklung von Modellen, Messverfahren und Studien zur Qualität mathematischer Bildung, um dann Trends und Impulse für die zukünftige Evidenzbündelung aufzuzeigen. Zuletzt leitet das Kapitel 11.3 Implikationen für die Bildungspolitik und pädagogische Praxis ab.

11.1 Diskussion der Teilprojekte des Dissertationsvorhabens

11.1.1 Auswirkung der Prozessqualität auf die mathematische Entwicklung (TP1)

Die Ergebnisse des TP1 zeigen, dass trotz der skizzierten Heterogenität der Befundlage und Studiendesigns die Evidenz europäischer Längsschnittstudien für persistente Effekte globaler und bereichsspezifischer Prozessqualität FIBB auf die Entwicklung in Mathematik und Sprache/Literacy spricht. Damit bestätigen die Ergebnisse Annahmen etablierter Modelle zur Entwicklungsrelevanz von Prozesskomponenten (ECCRN, 2002a; Kuger & Kluczniok, 2008; Mashburn et al., 2008; Pianta et al., 2005). In Ergänzung zu den Metaanalysen zur Effektivität von Offensiven und Initiativen (z. B. Barnett, 2011; Camilli et al., 2010; Nores & Barnett, 2010) machen die Prozesseffekte auf das Bildungspotenzial aufmerksam, welches in einer gestärkten Prozessqualität von Einrichtungen der Regelversorgung liegt. In Übereinstimmung mit der Annahme zur Bereichsspezifität von Lern- und Anregungsprozessen auch innerhalb

der kognitiv-leistungsbezogenen Entwicklungsdomäne (Burger, 2014; Hamre et al., 2014; Kluczniok & Roßbach, 2014; Kuger & Kluczniok, 2008; OECD, 2006, 2011; Roßbach & Weinert, 2008; Schneider et al., 2013; Stamm & Edelmann, 2012; Sylva et al., 2015) fielen die bereichsspezifischen Prozesseffekte in Mathematik und die globalen Prozesseffekte im Bereich Sprache/Literacy stärker aus. Trotz feiner Unterschiede in der Operationalisierung und im konzeptuellen Verständnis der Prozessqualität (Kluczniok & Roßbach, 2014; Pianta et al., 2005) scheinen verschiedene Maße entwicklungsrelevante Prozesskomponenten abzubilden. Für das Erfassen des Bildungspotenzials FIBB scheint allerdings ein Fokus auf Interaktionen statt auf die materiell-räumlichen Bedingungen in frühkindlichen Einrichtungen geeigneter (Arnett, 1989; Pianta et al., 2008) und die Verwendung von Beobachtungsverfahren unerlässlich (Halle et al., 2010; Mashburn et al., 2008; Pianta & Hamre, 2009; Sylva et al., 2006). Darüber hinaus unterstrichen zusätzliche Ergebnisse, wie wichtig der methodisch-statistische Fortschritt der Qualitätsforschung für die Evidenz zu prozessualen Effekten ist, damit Analysen die Wirkung der Prozessqualität abbilden können: Wichtig scheint bei der Planung und Auswertung von Studien die multiplen Bedingungsfaktoren kindlicher Entwicklung zu berücksichtigen (Anders, 2013; Burger, 2010; Melhuish et al., 2015) und auf elaborierte statistische Verfahren zurückzugreifen (Bryk & Raudenbush, 1992; Byrne, 2010; Goldstein, 2011; Hox, 2010).

11.1.2 Mathematische Bildungseffektivität als ergänzende Qualitätskomponente (TP2)

Vor dem Hintergrund der Ergebnisse für die Prozesseffekte in TP1, die zwar robust jedoch niedrig und heterogen ausfließen, sowie aufgrund weiterer bislang unzureichend gelöster Herausforderungen des Qualitätsansatzes wurde im TP2 des Promotionsvorhabens die mathematische Bildungseffektivität vorgeschlagen, um Einrichtungen anhand des Bildungsausgangs zu beurteilen. Mathematische Bildungsergebnisse zur Beurteilung der Bildungsförderung FIBB heranzuziehen steht weder im Widerspruch zu gängigen Qualitätsmodellen noch konnten andere überzeugende Argumente angeführt werden, die gegen eine Verwendung in der deutschen FIBB sprechen. Eigene Analysen zur mathematischen Bildungseffektivität der BiKS-Einrichtungen ergaben keine Effektivitätsunterschiede für faire Vergleiche (Racherbäumer et al., 2013; Scheerens, 1992). Unterschiede im Kompetenzniveau der Einrichtungen scheinen von Beginn bis zum Ende des Betreuungszeitraums zu bestehen und sich weitgehend durch andere soziokulturelle Zusammensetzungen in den Einrichtungen zu erklären. Die Befunde unterstrichen erstens die Bedeutung von Längsschnittstudien und der Erfassung entwicklungsrelevanter Merkmale für Effektivitätsvergleiche in der deutschen FIBB. Zweitens ergeben sich wichtige Abweichungen zu den Effektivitätsbefunden in EPPSE (Sammons et al., 2002; 2014;

Sylva et al., 2010) und zu den Ergebnissen der Prozessqualität aus BiKS (Anders et al., 2012; 2013; Lehl et al., 2016), was die ergänzende Funktion des Effektivitätsansatzes unterstreicht. Effektivitäts- und Qualitätsbefunde zusammen deuten wichtigen Optimierungsbedarf der mathematischen Bildung in der deutschen FIBB an. Zudem diskutierte TP2 die leistungsbezogene und soziokulturelle Zusammensetzung der Gruppe bzw. Einrichtung als ein bislang unzureichend erforschtes Merkmal FIBB, aus dem sich wichtige Implikationen für Forschung und Praxis ergeben.

11.1.3 Motivationalen Orientierungen und Professionalisierung in Mathematik als fachkraftbezogene Qualitätskomponenten (TP3)

Trotz weitgehender Anerkennung der zentralen Bedeutung der Fachkräfte (Anders, 2012; Aktionsrat Bildung, 2012; OECD, 2015b; Pianta et al., 2005), erscheinen die Qualitätskomponenten, sich auf die Fachkräfte beziehen, relativ unscharf konzeptualisiert und unterforscht. Das TP3 zeigte auf, wie ein Ausdifferenzieren und Erforschen der fachkraftbezogenen Orientierungs- und Strukturqualität in Mathematik durch den Rückgriff auf Beiträge anderer Forschungsrichtungen möglich ist. Lange Zeit beschränkte sich die Diskussion um fachkraftbezogene Strukturkomponenten auf globale, bereichsübergreifende Komponenten wie den Abschluss oder die Berufserfahrung. Die Ergebnisse des TP3 zu der geringen Anzahl besuchter mathematikspezifischer Aus- und Fortbildungsangebote zeigen einigen Nachholbedarf bei der Erforschung und Intensivierung der mathematikspezifischen Professionalisierung an (Anders, 2012; Chen & McCray, 2012; Gasteiger, 2010, 2012). Scheinbar blieb für die Mehrzahl der untersuchten Fachkräfte die formale Erfahrung mit Mathematik auf die Schulzeit begrenzt. Die emotionalen und motivationalen Schulerfahrungen erwiesen sich allerdings weder als äußerst positiv noch negativ. Zudem deuten Befunde zur mathematischen Selbstwirksamkeit und Wertorientierung der Fachkräfte, welchen eine motivierende Funktion und besondere Handlungsrelevanz zugeschrieben wird (Bandura, 1986; Fives & Buehl, 2012; Wigfield & Eccles, 2002), ein positives Bild der Orientierungsqualität der deutschen FIBB an. Lediglich das mathematische Selbstkonzept fiel vergleichsweise niedrig aus. Die Befunde stehen im Einklang mit vergleichbarer deutscher Evidenz (Benz, 2012a, 2012b; Thiel, 2010), jedoch betonte das TP3 die Unterschiede zu Befunden und dominierenden Annahmen über Schulerfahrungen und Fachkraftorientierungen von US-Fachkräften und unterstrich die Spezifika der Professionalisierung und Arbeit im deutschen FIBB-Kontext (Baroody, 2004; Bates et al., 2013; Copley, 2004; Copley & Padrón, 1999; Sweeting, 2011).

Positive und negative Schulerfahrungen scheinen Bedingungsfaktoren motivationaler Orientierungen zu sein (Copley & Padrón, 1999; Pajares, 1992; Relich, 1996; van Oers, 2002; Wilkins, 2008). Formale Qualifikation und die Professionalisierung vor Berufseinstieg hingen darüber hinaus mit der mathematischen Selbstwirksamkeit von Fachkräften der FIBB-Regelversorgung zusammen, was sich mit Befunden zur Wirksamkeit von Fachkrafttrainings und -initiativen deckt (z. B. Akay & Boz, 2010; Albayrak & Unal, 2011; Brand & Wilkins, 2007; Evans et al., 2013; Kalder & Lesik, 2011; Moseley & Utley, 2006; Rosenfeld, 2010; Saçkes et al., 2012; Swars, 2005; Wilkins & Brand, 2004). Andererseits wirft der fehlende Zusammenhang zu mathematischen Fortbildungen sowie zu anderen Orientierungen Fragen auf (vgl. Kalder & Lesik, 2011; Palmer, 2009; Quinn, 1997; Sherman & Christian, 1999; Wilkins, 2008). Die Befunde zur Berufserfahrung könnten auf die Bedeutung informeller Erfahrungen mit Mathematik und mathematischer Bildung während der pädagogischen Arbeit für Selbstwirksamkeit und Wertüberzeugung hindeuten (Cobanoğlu, 2011; Tirosh et al., 2011).

11.1.4 Möglichkeiten und Grenzen der Überwindung von Limitationen durch Evidenzbündelung

Im Rahmen des Promotionsvorhabens wurden in drei TP erste Antworten auf unzureichend geklärte Fragen zur Qualität mathematischer Bildung gefunden. Damit zeigte das Vorhaben, welcher Erkenntnisgewinn sich aus dem Bündeln der Evidenz verschiedener Studien ergeben kann, sowohl durch die Auswertung hinsichtlich *thematisch zusammenhängender, aber unterschiedlicher Fragestellungen* als auch durch die quantitative Synthese der Evidenz zu *vergleichbaren Fragestellungen*. Die Evidenzbündelung ermöglichte über den „Datenhorizont“ bzw. den „Informationsgehalt, einer Studie hinauszugehen, der auch bei komplexen Designs begrenzt bleibt, z. B. durch regionale oder methodologische Limitierungen. So bietet die BiKS-Studie die Angaben zu mathematischen Kompetenzen und Prozesskomponenten, um sowohl die Auswirkungen (TP1) als auch die mathematische Effektivität FIBB in zwei deutschen Bundesländern zu untersuchen (TP2). Die TP3-Studie ist auf die Fachkräfteebene begrenzt und erlaubte weder Aussagen zur Effektivität noch zur Auswirkungen FIBB, erfasste jedoch fachkraftbezogene Orientierungs- und Strukturkomponenten im mathematischen Bereich, welche in BiKS nicht erhoben wurden.

Die Metaanalyse ermöglichte außerdem regionale und methodische Limitationen der BiKS-Befunde zu Prozesseffekten durch die quantitative Synthese mit der Evidenz weiterer Längsschnittstudien zu überwinden. Die Synthese ermöglichte zum einen die Generalisierbarkeit der Befunde zu überprüfen, z. B. für Länder mit spezifischen FIBB-Systemen und pädagogischen

Tradition. Zum anderen ermöglichte die Metaanalyse Aussagen für Entwicklungszeiträume, die über die Dauer von Einzelstudien hinausgehen, sowie zu verschiedenen Prozessmaßen und methodisch-statistischen Vorgehensweisen. In einzelnen Studien wird selten mehr als eines der aufwändigeren Maße zu vergleichbaren Prozesskomponenten verwenden, z. B. CLASS und ECERS-R (siehe aber Pianta et al., 2005).

Eine große Herausforderung für die Evidenzbündelung in der Metaanalyse ergab sich aus der großen Bandbreite verschiedener Auswertungsverfahren, u. a. Korrelation, multiple Regression und latente Wachstumsmodelle, mit welchen in Primärstudien Prozesseffekte berechnet wurden. Cooper (2009) empfiehlt auf das Aggregieren von Effektgrößen verschiedenartiger Analysen z. B. Regression mit unterschiedlichen Prädiktoren zu verzichten und auf Korrelationen zurückzugreifen, was allerdings in mehrfacher Hinsicht kritisch zu bewerten ist (Bowman, 2012; Card, 2012). Erstens ist ein Zugriff auf Korrelationen komplexer Längsschnittstudien oder die notwendigen Informationen zur Berechnung dieser nur selten möglich. Zweitens ist anzuzweifeln, dass sich Prozesseffekte bivariat abbilden lassen, weil Lernprozesse durch ein Zusammenwirken verschiedener Faktoren bedingt werden (siehe Kapitel 2.3). Das TP1 aggregierte die vorliegenden Effektgrößen und explorierte stattdessen, angelehnt am Vorgehen für die Auswertung sonstiger Designunterschiede, Charakteristika des statistisch-methodischen Vorgehens als Moderatoren prozessualer Effekte (Ahn & Becker, 2011; Herbison, Hay-Smith & Gillespie, 2006; Juni, Witschi, Bloch & Egger, 1999).

Insgesamt konnte das Bündeln von Evidenz in den drei TP die Limitierungen einzelner Studien nur teilweise überwinden. Hieraus ergibt sich weiterer Forschungsbedarf. Auf die Limitierungen der einzelnen TP-Studien wird in den Anhängen A, B und D ausführlich eingegangen.

11.2 Implikationen für die Forschung zur Qualität mathematischer Bildung in der FIBB

11.2.1 Umfassenderes Verständnis mathematischer Lern- und Bildungsprozesse

Viele Messverfahren mathematischer Kompetenz (so auch mehrheitlich die Verfahren in TP1 sowie der K-ABC im TP2), beschränken sich auf die Bereiche Zahlen und Operationen und ggfs. Raum und Form, die als vergleichsweise gut erforscht gelten (vgl. Baroody et al., 2006; Burger, 2014; Kaufmann, 2011; Pauen & Herber, 2009; Schneider et al., 2013). Wenige Testverfahren erfassen weitere Inhaltsbereiche (z. B. Clements & Sarama, 2011; Fried, et al., 2012; Neumann et al., 2013) oder Prozesskompetenzen der Mathematik (Neumann et al., 2013). Ähnliches gilt für motivationale oder volitional-emotionale Kompetenzaspekte, wie die Lernfreude, das Involvement oder die Mathematikmotivation (vgl. aber Fried, et al., 2012;

Kluczniok et al., 2016; Lerkkanen et al., 2012; Richter, Lehl & Weinert, 2016). Ein ebenso verengtes Verständnis mathematischer Bildungsprozesse wird aus vielen Förderprogrammen und Prozessmaßen FIBB ersichtlich: Die ECERS-E bzw. KES-E als verbreitetes bereichsspezifisches Maß erfasst die Anregung in den prädiktiven Förderbereichen Zahlen und Operationen, Raum und Form, Muster und Strukturen, während andere Maße weitere Förderbereiche abdecken (z. B. Isele, 2014; vgl. Halle et al., 2010). Aus fachdidaktischer und entwicklungspsychologischer Perspektive sollte in Zukunft ein umfassenderes Verständnis früher mathematischer Lern- und Bildungsprozesse angestrebt werden (Kaufmann, 2011; Weinert, 2001).

11.2.2 Wahl und Weiterentwicklung von Prozessmaßen

Dass die Prozesseffekte in TP1 niedrig ausfielen und das gewählte Prozessmaß variierende Befunde erklären konnte, lässt für die Wahl und Weiterentwicklung von Prozessmaßen folgende Schlussfolgerungen zu:

Wenn es darum geht, bildungsrelevante Prozesskomponenten zu erfassen, sollte, laut der Ergebnisse aus TP1, *interaktionsfokussierten Maßen* globaler Prozessqualität den Vorzug gegeben und auf ein zusätzliches Rating der räumlich-materiellen Umgebung verzichtet werden. Bestreben, Prozessmaße in diese Richtungen zu verfeinern, sind somit voranzutreiben (z. B. Harms, Clifford & Cryer, 2014). Zusätzlich bedarf es Studien, in denen sich intensiv mit der *Bedeutung des Materials und der räumlichen Umgebung* für die mathematische Bildungsarbeit auseinandergesetzt wird, wie die Studien von Schuler (2013) oder Jörns, Schuchardt, Mähler und Grube (2013; vgl. Ginsburg & Golbeck, 2004). Nützlich erscheinen die Erkenntnisse für die (Weiter-)Entwicklung von Prozessmaßen und mathematischen Bildungsangeboten gleichermaßen, weil Angebote, die mit speziell entwickelten Materialien und räumlicher Umgebung (Bradley & Garrity, 2013; Friedrich & Munz, 2004, 2011; Pauen & Herber, 2009) arbeiten, Angeboten gegenüberstehen, die Alltagsgegenstände und -materialien nutzen, die nicht direkt für die mathematische Förderung ausgelegt sind (Stiftung Bildungspakt Bayern, 2007; Henschen & Teschner, 2013; Kaufmann, 2011; Rathgeb-Schnierer, 2013; Sommerlatte et al., 2009).

In Anbetracht der stärkeren bereichsspezifischen Prozesseffekte in TP1 ist die *Verwendung und Weiterentwicklung bereichsspezifischer Prozessmaße* gerade für die Forschung im mathematischen Bereich wichtig; dabei scheint jedoch, wie bereits erwähnt, das Verständnis mathematischer Bildungsprozesse gängiger Maße erweiterungsbedürftig. Überprüft werden sollte in diesem Zusammenhang auch, ob und wann das Zusammenführen zu einem Globalwert trotz Unterschiede im Ausmaß und in der Qualität der Anregung in verschiedenen Bereichen ratsam ist (TP1; Cabell et al., 2013). Monitorings und Large-Scale-Assessments benöti-

gen ökonomische Verfahren, welche die Prozessqualität in der FIBB reliabel und valide erfassen. Die TP1-Befunde signalisieren, dass *Fragebogenmaße (weiter-) entwickelt* werden müssen und sowohl auf ihren Zusammenhang mit Beobachtungsmaßen (Anders et al., 2014; Camehl, Schober & Spieß, 2012) als auch mit Bildungsergebnissen überprüft werden sollten.

Ein weiterer Gesichtspunkt für die Wahl und Weiterentwicklung von Prozessmaßen ist, inwiefern die *Art der Auseinandersetzung mit der mathematischen Materie und die Art des pädagogischen Handelns der Fachkräfte* als Merkmale des Bildungsgeschehens erfasst werden. Essentielle Prozesskomponenten der Interaktionsforschung wie Momente gemeinsam geteilter Denkprozesse bzw. Momente des Sustained Shared Thinking (Gasteiger, 2007; Hardy & Steffensky, 2014; Siraj-Blatchford, 2009; Siraj-Blatchford et al., 2002), Scaffolding-Prozesse oder das Ausnutzen von Bildungsmomenten bzw. Teachable Moments (vgl. Hyun & Marshall, 2003). sind in der ECERS-R/-E nicht abgebildet. Die CLASS-Skala Conceptual development (Pianta et al., 2008) und die Sustained shared Thinking and Emotional Well-Being Skala (SSTEW; Siraj, Kingston & Melhuish, 2015) erfassen einige dieser Prozesskomponenten bereichsübergreifend und die COEMET erfasste diese speziell für mathematische Bildungsprozesse (Sarama & Clements, 2009).

11.2.3 Effektivitätsanalysen bestehender und zukünftiger Studien

Anhaltspunkte für weitere Effektivitätsanalysen und für die Planung zukünftiger Studien zur FIBB-Effektivität bieten die Befunde und Limitierungen der BiKS-Effektivitätsanalysen. Beispielsweise sollte zukünftig auf eine ausreichende *Größe und Repräsentativität der Stichproben* (der Einrichtungsstichprobe und der Stichprobe an Kindern pro Einrichtung) geachtet werden. Wie beim Sampling und bei Effektivitätsanalysen damit umzugehen ist, dass bildungsbenachteiligte Familien in Deutschland ihre Kinder später in die FIBB geben (Geier & Riedel, 2009; Schober & Spieß, 2012), muss eruiert werden. Interessante Forschungsfragen ergeben sich auch für die weiteren Ebenen des FIBB-Systems, z. B. Vergleich von Trägern oder (Bundes-)Ländern (vgl. Lenkeit, 2012). Zu achten ist auf ein *längsschnittliches oder ggfs. querschnittliches Design, welches die Erhebung von Kompetenzen und entwicklungsrelevanten Merkmalen* vorsieht. Für differenziertere Effektivitätsanalysen, z. B. die Betrachtung von Veränderungen in der Rangordnung der Einrichtungen über den zeitlichen Verlauf (Creemers et al., 2010) oder Wachstumsmodelle der Effektivität (Chapman et al., 2015; Creemers et al., 2010) sind mehrere Erhebungswellen unerlässlich.

Sowohl in EPPSE als auch in der schulischen Bildungsforschung werden *Mixed-Methods-Ansätze* verwendet, welche Qualitätserhebungen sowie vertiefende qualitative Analysen des pädagogischen Geschehens einschließen (vgl. Chapman et al., 2015; Racherbäumer et al.,

2013; Sammons et al., 2005; Siraj-Blatchford et al., 2006) und erheblich für Fortschritte der Qualitätsforschung dienen können, z. B. neue Prozessmaße (Siraj et al., 2015). Während für mehrere Qualitätskomponenten Standards vorliegen z. B. Strukturkomponenten (Viernickel & Schwarz, 2009) oder Prozessmaße (siehe Tabelle 2 im Anhang B), existieren *keine Standards zur Beurteilung des erzielten Bildungoutputs* der Gesamtstichprobe oder einzelner Einrichtungen. Ob die Effektivität zufriedenstellend ist, sollte vor dem Hintergrund angestrebter Bildungsziele beurteilt werden, die allerdings für die FIBB kaum konkret formuliert werden (siehe Kapitel 3.4 und 9.2).

11.2.4 Empirischer Vergleich zu anderen Entwicklungsbereichen

Ein umfassendes Verständnis mathematischer Bildung schließt das Wissen über weitere Entwicklungsbereiche ein, z. B. zur Effektivität in anderen Bereichen. In Anbetracht der belegten Differenzen in Ausmaß und Qualität der Förderung zwischen Bereichen ist von differenziellen Effektivitätsprofilen in der deutschen FIBB auszugehen: In EPPSE ergaben sich lediglich mittlere Zusammenhänge der Effektivität zwischen Bereichen der kognitiven Domäne (Sammons et al., 2002). Insbesondere der Frage, ob und welche Einrichtungen effektiv in Mathematik und sozio-emotionalen Bereichen sind (Sammons et al., 2005; Sammons et al., 2003), kommt in Anbetracht der großen Skepsis gegenüber früher Lernförderung und der vermuteten negativen sozio-emotionalen Folgen große Bedeutung zu (Anders, 2013; Loeb, Bridges, Bassok, Fuller & Rumberger, 2007; Melhuish et al., 2015; OECD, 2006). In diesem Zusammenhang wäre ebenso eine Metaanalyse der Evidenz zu bereichsspezifischen Prozesseffekten auf die sozio-emotionale Entwicklung und das Wohlbefinden von Kindern ein wichtiger nächster Schritt. Ebenfalls wäre es wichtig, dass in Forschungsprojekten die Rolle der bereichsspezifischen Anregung in anderen Bereichen z. B. Sprache oder Naturwissenschaften für die mathematische Entwicklung erforscht wird (vgl. Anders et al., 2012; Klibanoff et al., 2006).

11.2.5 Evidenz zu Interaktionseffekten

Mehr Evidenz wird zu kompensatorischen Effekten der FIBB-Regelversorgung benötigt (vgl. Anders, 2013; Burger, 2010; Melhuish et al., 2015). Beispielsweise wurde in EPPSE beobachtet, dass eine hohe Prozessqualität und mathematische Effektivität gerade für benachteiligte Kinder bedeutsam ist (u. a. niedriger elterlicher Bildungsabschluss und HLE; Sammons et al., 2014; Sylva et al., 2010). Ergebnisse der BiKS-Studie deuten an, dass benachteiligte Kinder in Deutschland nicht von einer hohen bereichsspezifischen Prozessqualität profitieren (niedriges HLE; Anders et al., 2012; Lehl et al., 2014). Allerdings gibt es nicht aus-

reichend vergleichbare Evidenz, um kompensatorische Effekte der Prozessqualität oder Effektivität metaanalytisch auszuwerten. Gleiches gilt für das Zusammenspiel von Prozessqualität und Effektivität mit quantitativen Merkmalen der FIBB wie Dauer oder Intensität der Betreuung (Anders, 2013; Melhuish et al., 2015; Nores & Barnett, 2010; Sylva et al., 2010) und mit neuen mathematischen Bildungserfahrungen z. B. in der Grundschule (vgl. Lehl et al., 2016; Sylva et al., 2010).

11.2.6 Differenzierung und Wirkweise der Struktur- und Orientierungskomponenten

Weitere relevante Ansatzpunkte für die Konkretisierung und Differenzierung von Qualitätsmodellen stellen Bedingungsfaktoren der Prozessqualität dar. Hier besteht weiterer Forschungsbedarf, wobei mehrere Punkte wichtig erscheinen:

Günstige strukturelle Rahmenbedingungen werden als wichtige Bedingungsfaktoren von Prozessqualität und damit Bildungsergebnissen diskutiert, scheinen diese jedoch nicht zu garantieren (vgl. Ulferts & Anders, 2015 zum fehlenden Gesamteffekt der Strukturkomponenten). Die Wirkweise der Strukturqualität scheint ungleich komplexer und vielmehr von der Interaktion wichtiger Komponenten und der Gruppenzusammensetzung abzuhängen (Slot, Lerkkanen, et al., 2015). Von enormer Relevanz für die Bildungspolitik und Praxis wären evidenzbasierte Aussagen zu geeigneten *Standards für Strukturkomponenten und prozess- und lernförderliche Kombinationen ausgewählter Komponenten* in Abhängigkeit von der Gruppenzusammensetzung (vgl. Burchinal, Vandergrift, Pianta & Mashburn, 2010).

Zudem wird eine eingehendere konzeptuelle und empirische Auseinandersetzung mit den spezifischen Struktur- und Orientierungskomponenten der einzelnen Entwicklungsbereiche benötigt, z. B. motivationale Orientierungen und Aus- und Fortbildung der Fachkräfte im Bereich Sprache/Literacy oder sozio-emotionalen Bereichen. Die Evidenz spricht dafür, dass gerade eine tiefgehende theoretische und empirische Auseinandersetzung mit fachkraftbezogenen Qualitätskomponenten von enormer Bedeutung ist. In den Analysen von Slot, Lerkkanen, et al. (2015) erwies sich die Qualifikation der Fachkraft als enorm bedeutungsvoll, um eine gute Prozessqualität auch bei herausfordernden Rahmenbedingungen für die pädagogische Arbeit wie ein hoher Migrantanteil oder große Gruppen zu realisieren. In einer Erweiterung der Metaanalyse resultierte nur für die Fachkraftqualifikation, nicht für strukturell-räumliche Komponenten der Prozessqualität ein signifikanter Effekt (u. a. Erzieher-Kind-Schlüssel oder Gruppengröße; Ulferts & Anders, 2015).

Die Auseinandersetzung sollte neben einer Bestandsaufnahme fachkraftbezogener Qualitätskomponenten in der FIBB-Regelversorgung die jeweilige Funktion für die pädagogische Arbeit und die Qualität realisierter Prozesse nachvollziehen. So gilt es sich mit der Anzahl und

Ausgestaltung mathematischer Aus- und Fortbildungsangebote (z. B. Intensität, Inhalte und Methoden) sowie ihrer Wirkung auf diverse Kompetenzfacetten auseinanderzusetzen. Wenig bekannt ist auch über die Motivation von Fachkräften sich eingehender zu qualifizieren bzw. über wirksame Anreize, um die Professionalisierungs- und Qualifikationsbereitschaft von Fachkräften zu erhöhen. Großer Forschungsbedarf lässt sich generell für die Orientierungskomponenten ausmachen wie zur Selbstwirksamkeit deutscher Fachkräfte gute mathematische Bildungsarbeit leisten zu können (Chen et al., 2014). Ebenfalls zentral sind die Fragen nach der Veränderbarkeit durch formale Professionalisierungsangebote und durch die Gestaltung des FIBB-Arbeitskontexts, so dass Raum für positive und motivierende informelle Erfahrungen mit der Mathematik oder früher mathematischer Bildung in Rahmen der beruflichen Tätigkeit geschaffen wird (Buchanan et al., 1998; Guo et al., 2011; Upadyaya et al., 2011).

11.2.7 Die kompetenzbezogene und soziokulturelle Zusammensetzung als Bedingungsfaktor

Wie die Evidenz aus TP2 und weiteren Studien zeigt (Becker, 2010; Becker & Schober, 2015; EACEA/Eurydice, 2009; Leseman & Slot, 2014; Pakarinen et al., 2011), handelt es sich bei der soziokulturellen und kompetenzbezogenen Zusammensetzung von Gruppen und Einrichtungen um ein wichtiges Merkmal FIBB. Konzeptuell scheint die Rolle der Zusammensetzung ebenfalls unzureichend geklärt, jedenfalls ist von einer Zuordnung zur Strukturqualität (z. B. Kuger & Kluczniok, 2008; Tietze et al., 2012) aus ethischen und bildungspolitischen Gründen eher abzuraten. So droht ein Auflisten des Migrantenanteils als Qualitätskomponente die Segregationsproblematik zu verstärken. Einzelne Prozessmaße berücksichtigen den Umgang mit multikultureller Vielfalt und Diversität in Prozessmaßen (z. B. DO-RESI-E-Ki, ECERS-R/-E). Die Vorschläge scheinen jedoch unzureichend, wenn eine pädagogisch herausfordernde Zusammensetzung noch immer ihren Niederschlag in Prozesswerten findet und zu einem ungünstigeren Gesamturteil führt (Kuger & Kluczniok, 2008; Kuger et al., 2016; Tietze et al., 2012). In Forschungsprojekten sollte die Auswirkungen der Zusammensetzung auf die mathematischen Bildungs- und Lernprozesse eingehender analysiert (vgl. Sammons et al., 2002) und zudem der Frage nachgegangen werden, wie den pädagogischen Herausforderungen begegnet werden kann, welche durch bestimmte Gruppenkompositionen entstehen, z. B. durch ein höheres Qualifikationsniveau oder eine spezifische Ausbildung der Fachkraft sowie kleinere Gruppen (Slot, Lerkkanen, et al., 2015).

11.3 Implikationen für das zukünftige Bündeln von Evidenz unterschiedlicher Studien

In Zukunft werden verschiedene Arten der Evidenzbündelung in der frühkindlichen Bildungsforschung an Bedeutung gewinnen aufgrund der weltweit zunehmenden Forschungstä-

tigkeit in diesem Feld sowie der gestiegenen Bestrebungen zu international vergleichender Forschung. Im Zuge der stärkeren Verfügbarkeit von Evidenz durch Public-Use-Files oder den Datenaustausch in großen Forschungsverbänden wird die quantitative Synthese von Rohdaten sowie die Durchführung von Sekundäranalysen zunehmen. Slot, Lerkkanen, et al. (2015) zeigten, wie parallelisierte Sekundäranalysen verschiedener Längsschnittstudien erste Erkenntnisse über das interaktive Zusammenwirken von Strukturkomponenten auf die Prozessqualität FIBB ermöglichen. In der Bildungsforschung sollte zukünftig auf *elaboriertere Methoden zur Harmonisierung und integrativen Datenanalyse bzw. komplexen Metaanalyse* zurückgegriffen werden, um komplexe Fragen zur Qualität FIBB mit der gebündelten Evidenz aus Rohdaten der Studien auszuwerten (Bainter & Curran, 2014; Becker, 2000; Becker & Wu, 2007; Cheung, 2014; Curran & Hussong, 2009).

Ausgangspunkt von Synthesen bildet eine gründliche Auseinandersetzung mit den Designs der Studien, um die Vergleichbarkeit, Aussagekraft und Limitierungen der Evidenz einschätzen zu können, wofür ein fundiertes Wissen über das Forschungsfeld notwendig ist, z. B. zu FIBB-Systemen und Prozessmaßen. In diesem Zusammenhang sollte die *Vergleichbarkeit der Evidenz anderer Forschungsrichtungen und Forschungsstränge* ausgewertet werden, z. B. die Übertragbarkeit von Erkenntnissen aus Evaluationsstudien zur Wirksamkeit von Förderansätzen und Bildungsoffensiven/-initiativen sowie Professionalisierungsmaßnahmen auf die Auswirkung der FIBB auf die kindliche Entwicklung und Professionalisierung von Fachkräften der Regelversorgung. Ebenfalls ist der Übertragbarkeit von Evidenz aus anderen Ländern Grenzen gesetzt, z. B. mit Blick auf landesspezifische Formen und Situationen der Benachteiligung, Bildungsansätze (z. B. zielgruppenspezifische vs. universelle kompensatorische Ansätze) oder Spezifika der pädagogischen Arbeit und Professionalisierung von Fachkräften in den jeweiligen Ländern. Um *länderübergreifende von -spezifischen Aspekten des Qualitätsverständnisses zu trennen*, bedarf es vermehrt internationalen und interdisziplinären Projekten, welche eine Plattform für den koordinierten Austausch über das Wissen zur Qualität und zu Systemen sowie von Daten und Forschung zur FIBB unter Experten darstellen (Broekhuizen et al., 2015; Cryer et al., 1999; Montie, Xiang & Schweinhart, 2006; OECD, 2011, 2015a; Slot, Lerkkanen, et al., 2015).

11.4 Bildungspolitische und praktische Implikationen für die FIBB-Regelversorgung

In den letzten zwei Dekaden stieg weltweit der Ausbau und die Investition in FIBB, allerdings ist gerade in den letzten Jahren der Druck zur Reduktion von Ausgaben auch im frühkindlichen Bildungssektor gewachsen (OECD, 2013, 2015a). Für weitere Investitionen und

einen weiteren Ausbau wird mit der Evidenz aus Metaanalysen zu nachhaltigen Effekten und ökonomischen Analysen zu hohen Bildungsrenditen von frühkindlicher Bildung argumentiert (Belfield, Nores, Barnett & Schweinhart, 2006; Heckman, 2006). Die Evidenz stammt vornehmlich aus zielgruppenspezifischen Offensiven und Initiativen, welche jedoch in dem Ausmaß in Europa kaum vorliegen. Die überwiegende Mehrheit der Kinder, ebenso Bildungsbenachteiligte, befindet sich in Angeboten regulärer FIBB. Damit haben Maßnahmen, welche die Effekte der Regelversorgung optimieren und so die Mehrheit aller Kinder in Europa erreichen, das Potenzial eine große Breitenwirkung zu entfalten und eine kompensatorische Bildung benachteiligter Kinder durch universell-hochwertige Bildung scheint gegenüber einer großflächigen Ausstattung mit zielgruppenspezifischen Angeboten weitere Vorteile zu bieten (siehe Kapitel 3.5). Die effektive Steuerung FIBB und gezielte Investition außerhalb spezifischer Offensiven und Initiativen ist allerdings schwierig und benötigt eigene empirisch abgesicherte Zielparameter. Das Qualitätskonzept bietet solche Zielparameter und die Qualitätsforschung die dazugehörige Evidenz (Anders, 2013; Burger, 2010; Dalli et al., 2011; Melhuish et al., 2015).

11.4.1 Fokus von Investitionen, Maßnahmen und Monitoring auf die Prozessqualität

Die gebündelte Evidenz zur FIBB-Regelversorgung in Europa aus TP1 spricht für die Prozessqualität als einen Zielparameter bildungspolitischer Steuerung und Investitionen. Gleichzeitig deuten die geringen Mittelwerte der Prozessqualität, die sich im TP1 in verschiedenen Erhebungen in Europa ergaben (vgl. Clifford et al., 2010; Kuger & Kluczniok, 2008; Slot, Lerkkanen, et al., 2015; Vermeer et al., 2016), insbesondere bereichsspezifischer Prozessqualität in Mathematik, auf ungenutztes Bildungspotenzial gerade im mathematischen Bereich hin, für den bereichsspezifische Prozesseffekte besonders stark ausfielen. Ein wichtiger Schritt zu einer flächendeckend guten Prozessqualität ist eine Verlagerung von der reinen Input- zur stärkeren Prozesssteuerung FIBB (OECD, 2015b; Tietze, 2008). FIBB-Monitoringsysteme sollten zusätzlich zu strukturellen Qualitätskomponenten und Sicherheitsstandards die Prozessqualität in Einrichtungen überprüfen wie in verpflichtenden Inspektionen oder im Rahmen nationaler oder internationaler Erhebungen (Bäumer et al., 2013; Moser & Studer, 2013; Network on Early Childhood Education and Care, 2013; OECD, 2015b; Tietze et al., 2012). Wie TP1-Befunde andeuten ist ein Fokus auf das interaktive Geschehen beim Monitoring ratsam wenn es um die Bildungsergebnisse FIBB geht. Von den vorliegenden Fragebogenverfahren, wie die Fachkräfteeinschätzungen im Rahmen des schottischen Monitorings (Bradshaw et al., 2014), ist in ihren aktuellen Fassungen abzuraten. Aus den fehlenden Effekten für Fragebogenverfahren ergibt sich zudem die Frage, ob Fachkräfte die eigene pädagogi-

sche Praxis hinreichend reflektieren und einschätzen können und wie Ergebnisse aus Prozessmonitorings als externes Feedback zur Bildungsarbeit für Professionalisierungszwecke dienen können. Erfahrungen zum Einsatz in Fachkräfteschulungen liegen für einzelne Prozessmaße mit starkem Fokus auf die Fachkraft-Kind-Interaktion vor (z. B. CLASS-Verfahren oder DORESI-E-Ki).

11.4.2 Monitoring mathematischer Bildungsergebnisse und Effektivitätsanalysen

Zudem wäre eine zusätzliche Output-Steuerung durch ein Monitoring von Bildungsergebnissen und ggfs. der Bildungseffektivität FIBB zu erwägen, vorausgesetzt das Monitoringdesign erfüllt die notwendigen Voraussetzungen. Für eine durchdachte Einführung als Monitoring- und Steuerungsinstrument in der FIBB sollte aus den bisherigen Erfahrungen in EPPSE, der Schuleffektivitätsforschung und ganz besonders aus dem wachsenden Gebrauch und Missbrauch in den USA gelernt werden (Chapman et al., 2015; Ditton; 2000; Sylva et al., 2010). Zu klären wäre in diesem Zusammenhang, wie damit umzugehen ist, dass kaum Angaben zu Zielgrößen vorliegen, ähnlich der Bildungsstandards und konkreten Lernziele in Curricula der Schulen.

11.4.3 Evidenzbasierte Maßnahmen für die Regelversorgung

Befunde der Evaluationsstudien zu Förderprogrammen und der Implementationsforschung zu Bildungsoffensiven und -initiativen können auch für die Regelversorgung wertvolle Erkenntnisse über Ansatzpunkte für die Verbesserung der mathematischen Bildungsqualität liefern, insbesondere wenn die Optimierung pädagogischer Prozesse oder der Bildungsarbeit sowie die Maximierung mathematischer Bildungsergebnisse evaluiert werden. Hieraus gilt es aber Maßnahmen zu entwickeln, mit denen flächendeckend für die Vielfalt an FIBB-Angeboten und -Konzepten eine gute mathematische Bildung sichergestellt ist.

In Deutschland gibt es einige mathematische Initiativen, Konzepte und Bildungsoffensiven, auch genießen mehrere Förderprogramme große Verbreitung, z. B. Zahlenland, für die Wirksamkeitsnachweise vorliegen (siehe Kapitel 4.2 und 4.3). Dies schlägt sich jedoch weder in einer großen Varianz der mathematischen Prozessqualität in deutschen Erhebungen nieder noch in Effektivitätsunterschieden in den zwei Bundesländern der BiKS-Studie des TP2. Die punktuelle Versorgung mit Angeboten scheint daher nicht ausreichend, um ein flächendeckend hohes Bildungspotenzial FIBB zu gewährleisten. Angesichts der niedrigen bis mittleren Prozesswerte in den meisten Ländern, kann ebenfalls angezweifelt werden, dass die gesetzlichen Vorgaben und Regulierungen zur Strukturqualität wie der Erzieher-Kind-Schlüssel oder die Gruppengröße ausreichend sind (European Commission/EACEA/Eurydice/Eurostat, 2014;

OECD, 2015b). Ob es sich um effektive Maßnahmen zur Sicherung und Optimierung des Bildungspotenzials handelt muss zudem wegen des fehlenden Gesamteffekts für viele Strukturkomponenten angezweifelt werden (Ulferts & Anders, 2015). Investitionen und Maßnahmen sollten sich vermehrt auf strukturelle Zielparameter richten, welche sich nachweislich auf die Prozessqualität und die kindliche Entwicklung auswirken, z. B. die Qualifikation von Fachkräften (Anders, 2013; Melhuish et al., 2015; Slot, 2014; Slot, Leseman, Verhagen & Mulder, 2015; Ulferts & Anders, 2015), und wegen nachgewiesener Interaktionseffekte dabei verstärkt auf die Abstimmung verschiedener Regulierungen achten (Slot, Lerkkanen, et al., 2015).

11.4.4 Intensivierung der Qualifizierung und mathematikspezifischen Professionalisierung

Zentraler Ansatzpunkte zur Optimierung des mathematischen Bildungspotenzials FIBB ist eine ausreichende Qualifikation und Professionalisierung der Fachkräfte (vgl. Aktionsrat Bildung, 2012; Fukkink & Lont, 2007; Slot et al., 2015; Ulferts & Anders, 2015), auch weil die pädagogische Arbeit der Fachkräfte hinter den gesteckten Erwartungen zurückbleibt. Obschon erfolgversprechende Offensiven und Initiativen sowie Konzepte der Fachkraftschulung in Mathematik existieren (z. B. Baroody et al., 2006; Burger, 2014; Chen & Chang, 2006; Gasteiger, 2012; Jungmann & Koch, 2017), signalisieren die TP3-Befunde einigen Nachholbedarf für die mathematikspezifische Aus- und Fortbildung in der FIBB-Regelversorgung. Zum einen sollten Regulierungen bzw. Anreize geschaffen werden, so dass Professionalisierungsangebote von Fachkräften in Anspruch genommen werden. Zudem bedarf es ausreichend qualitativ hochwertiger Angebote in Mathematik und eine regelmäßige Überprüfung von Inhalten und Methoden. Eine gute Grundlage für Fachkraftorientierungen könnten Angebote darstellen, welche neben der Wissensvermittlung Möglichkeiten für motivierende und positive Erfahrungen bei der Beschäftigung mit Mathematik und mathematischer Bildung unterbreiten und u. U. negative Schulerfahrungen thematisieren.

Als günstig könnte sich auch ein Anreichern der Berufserfahrung mit motivierenden mathematischen Erfahrungen erweisen. Dementsprechend könnten Arbeitskontext und pädagogischer Alltag in den Einrichtungen strukturiert werden, z. B. die Salienz und Bedeutung mathematischer Bildung durch eine Verankerung in den Einrichtungskonzeptionen als Bildungsschwerpunkt erhöhen, Förderung des fachlichen Austauschs der Fachkräfte über die mathematische Bildungsarbeit in Teammeetings, mathematische Bildung als festen Bestandteil interner Workshops und in Elterngesprächen etablieren.

11.4.5 Maßnahmen zur stärkeren Durchmischung bzw. zum pädagogischen Umgang

Die Befunde aus TP2 und weiteren Studien sprechen dafür, dass eine formal universelle Bildungsinfrastruktur Segregationsprozesse und den Ausschluss Bildungsbenachteiligter an (früher) und gleichberechtigter Teilhabe nicht verhindert (EACEA/Eurydice, 2009; Leseman & Slot, 2014; OECD, 2006). Bildungspolitische Maßnahmen zur frühen Inanspruchnahme FIBB durch Bildungsbenachteiligte und eine stärkere Durchmischung in Gruppen und Einrichtungen wie einkommensgestaffelte Beiträge bzw. Beitragsbefreiung oder -zuschüsse für Bildungsbenachteiligte sollten verstärkt werden (OECD, 2006; EACEA/Eurydice, 2009). Zusätzlich ist ein weiteres Optimieren der Rahmenbedingungen in Abhängigkeit von der Zusammensetzung zu erwägen, z. B. günstigere Erzieher-Kind-Schlüssel und Gruppengröße in Abhängigkeit von Anteil an Kindern mit Migrationshintergrund oder Bildungsbenachteiligung. Bislang scheinen die Bestrebungen in Deutschland und anderen Ländern diesbezüglich nicht auszureichend. Verstärkt sollte auch an wirksamen Konzepten und Schulungen für den Umgang mit soziokultureller und leistungsbezogener Diversität in FIBB und der Umsetzung inklusiver mathematischer Bildung gearbeitet werden.

Literaturverzeichnis

- Abreu-Lima, I. M. P., Leal, T. B., Cadima, J. & Gamelas, A. M. (2013). Predicting child outcomes from preschool quality in Portugal. *European Journal of Psychology of Education, 28*(2), 399-420.
- Ahn, S. & Becker, B. J. (2011). Incorporating quality scores in meta-analysis. *Journal of Educational and Behavioral Statistics, 36*(5), 555-585.
- Akay, H. & Boz, N. (2010). The effect of problem posing oriented analyses-II course on the attitudes toward mathematics and mathematics self-efficacy of elementary prospective mathematics teachers. *Australian Journal of Teacher Education, 35*(1), 59-75.
- Albayrak, M. & Unal, Z. A. (2011). The effect of methods of teaching mathematics course on mathematics teaching efficacy beliefs of elementary pre-service mathematics teachers. *International Journal of Humanities and Social Science, 16*(1), 183-190.
- Anders, Y. (2012). *Modelle professioneller Kompetenzen für frühpädagogische Fachkräfte*. München: KnoBLInGDESiGn.
- Anders, Y. (2013). Stichwort: Auswirkungen frühkindlicher institutioneller Betreuung und Bildung. *Zeitschrift für Erziehungswissenschaft, 16*(2), 237-275.
- Anders, Y. (2014). Literature review on pedagogy. Literature Review for the OECD, Paris: OECD.
- Anders, Y., Ballaschk, I. & Tietze, W. (2014). *Wissenschaftliche Untersuchungen zur Arbeit der Stiftung "Haus der kleinen Forscher"*. (No. 5). Schaffhausen: Schubi-Verl.
- Anders, Y., Grosse, C., Roßbach, H.-G., Ebert, S. & Weinert, S. (2013). Preschool and primary school influences on the development of children's early numeracy skills between the ages of 3 and 7 years in Germany. *School Effectiveness and School Improvement, 24*(2), 195-211.
- Anders, Y. & Roßbach, H.-G. (2015). Preschool Teachers' Sensitivity to Mathematics in Children's Play: The Influence of Math-Related School Experiences, Emotional Attitudes, and Pedagogical Beliefs. *Journal of Research in Childhood Education, 29*(3), 305-322.
- Anders, Y., Roßbach, H.-G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S. et al. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly, 27*(2), 231-244.

- Anderson, L. M., Shinn, C., Fullilove, M. T., Scrimshaw, S. C., Fielding, J. E., Normand, J. et al. (2003). The effectiveness of early childhood development programs. *American Journal of Preventive Medicine*, 24(3), 32-46.
- Andersson, B. E. (1989). Effects of public day-care: a longitudinal study. *Child Dev*, 60(4), 857-866.
- Andersson, B. E. (1992). Effects of day-care on cognitive and socioemotional competence of thirteen-year-old Swedish school children. *Child Dev*, 63(1), 20-36.
- Aktionsrat Bildung. (Hrsg.). (2012). *Professionalisierung in der Frühpädagogik: Qualifikationsniveau und-bedingungen des Personals in Kindertagesstätten*. Münster: Waxmann.
- Arnett, J. (1989). *Caregiver interaction scale*. Princeton, NJ: Educational Testing Service.
- Arnold, D. H. & Doctoroff, G. L. (2003). The early education of socioeconomically disadvantaged children. *Annual Review of Psychology*, 54, 517-545.
- Arnold, D. H., Fisher, P. H., Doctoroff, G. L. & Dobbs, J. (2002). Accelerating math development in Head Start classrooms. *Journal of Educational Psychology*, 94(4), 762-770.
- Aslan, D., Ogul, I. G. & Tas, I. (2013). The impacts of preschool teachers' mathematics anxiety and beliefs on children's mathematics achievement. *International Journal of Humanities and Social Science Invention*, 2(7), 45-49.
- Aunola, K., Leskinen, E., Lerkkanen, M.-K. & Nurmi, J.-E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology*, 96(4), 699-713.
- Bainter, S. A. & Curran, P. J. (2014). Advantages of integrative data analysis for developmental research. *Journal of Cognition and Development*, 16(1), 1-10.
- Ballou, D. & Springer, M. G. (2015). Using student test scores to measure teacher performance: Some problems in the design and implementation of evaluation systems. *Educational Researcher*, 44(2), 77-86.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. NJ: Prentice Hall: Englewood Cliffs.
- Baptist, P. & Raab, D. (2007). *Auf dem Weg zu einem veränderten Mathematikunterricht*. Universität Bayreuth: Zentrum zur Förderung des mathematisch-naturwissenschaftlichen Unterrichts.
- Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The future of children*, 5(3), 25-50.

- Barnett, W. S. (2011). Effectiveness of early educational intervention. *Science*, 333(6045), 975-978.
- Barnett, W. S., Brown, K. C. & Shore, R. (2004). *The universal vs. targeted debate: Should the United States have preschool for all?* New Brunswick, NJ: National Institute for Early Education Research.
- Baroody, A. J. (2004). The role of psychological research in the development of early childhood mathematics standards. In D. H. Clements & J. Samara (Eds.), *Standards for early childhood mathematics education* (pp. 149-172). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Baroody, A. J., Lai, M.-I. & Mix, K. (2006). The development of young children's early number and operation sense and its implications for early childhood education. In B. Spedek & O. N. Saracho (Eds.), *Handbook of research on the education of young children* (Vol. 2, S. 187-221). Mahwah, NJ / London: Lawrence Erlbaum Associates.
- Bates, A. B., Latham, N. I. & Kim, J.-A. (2013). Do I have to teach math? Early childhood pre-service teachers' fears or teaching mathematics. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, 5, 1-10.
- Baumeister, K. & Grieser, A. (2011). *Berufsbegleitende Fort- und Weiterbildung frühpädagogischer Fachkräfte - Analyse der Programmangebote: eine Studie der Weiterbildungsinitiative Frühpädagogische Fachkräfte (WiFF)*. München: DJI.
- Bäumer, T., Aßmann, C., von Maurice, J. & Blossfeld, H.-P. (2013). Möglichkeiten der Analyse von Kontexteffekten im Rahmen des Nationalen Bildungspanels. In R. Becker & A. Schulze (Hrsg.), *Bildungskontexte* (S. 61-83). Wiesbaden: Springer Fachmedien Wiesbaden.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469-520.
- Stiftung Bildungspakt Bayern (Hrsg.) (2007). *Das KiDZ-Handbuch. Grundlagen, Konzepte und Praxisbeispiele aus dem Modellversuch KiDZ – Kindergarten der Zukunft in Bayern*. Köln: Carl Link.
- Becker, B. (2010). Ethnische Unterschiede bei der Kindergartenselektion: Die Wahl von unterschiedlich stark segregierten Kindergärten in deutschen und türkischen Familien. In B. Becker & D. Reimer (Hrsg.), *Vom Kindergarten bis zur Hochschule. Die Generierung von ethnischen und sozialen Disparitäten in der Bildungsbiographie* (S. 17-47). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Becker, B. & Schober, P. S. (2015). Not just any day-care center? Social and ethnic disparities in the choice of high quality early education institutions. *Manuscript submitted for publication*.
- Becker, B. J. (2000). Multivariate meta-analysis. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 499-525). San Diego, CA: Academic Press.
- Becker, B. J. & Wu, M. J. (2007). The synthesis of regression slopes in meta-analysis. *Statistical Science*, 22(3), 414-429.
- Behr, K. & Walter, M. (2012). *Qualifikationen und Weiterbildung frühpädagogischer Fachkräfte: bundesweite Befragung von Einrichtungsleitungen und Fachkräften in Kindertageseinrichtungen ; zehn Fragen - zehn Antworten ; eine Studie der Weiterbildungsinitiative Frühpädagogische Fachkräfte (WiFF)*. München: DJI.
- Belfield, C. R., Nores, M., Barnett, S. & Schweinhart, L. (2006). The High/Scope Perry Preschool Program: Cost-benefit analysis using data from the age-40 Followup. *Journal of Human Resources*, 41(1), 162–190.
- Benz, C. (2012a). Attitudes of kindergarten educators about math. *Journal für Mathematik-Didaktik*, 33(2), 203-232.
- Benz, C. (2012b). Maths is not dangerous – attitudes of people working in German kindergarten about mathematics in kindergarten. *European Early Childhood Education Research Journal*, 20(2), 249-261.
- Bercovici, C. W. (2015). Early childhood math coaching: relationships and strategies. *Education Doctoral Theses 207* 1-109.
- Bischof, L. M., Hochweber, J., Hartig, J. & Klieme, E. (2013). Schulentwicklung im Verlauf eines Jahrzehnts – Erste Ergebnisse des PISA-Schulpanels. In N. Jude & E. Klieme (Hrsg.), *PISA 2009 – Impulse für die Schul- und Unterrichtsforschung* (S. 172-199). Weinheim: Beltz.
- Blevins-Knabe, B. & Musun-Miller, L. (1996). Number use at home by children and their parents and its relationship to early mathematical performance. *Early Development and Parenting*, 5(1), 35-45.
- Blok, H., Fukkink, R., Gebhardt, E. & Leseman, P. (2005). The relevance of delivery mode and other programme characteristics for the effectiveness of early childhood intervention. *International Journal of Behavioral Development*, 29(1), 35-47.

- Bong, M. & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, 15(1), 1-40.
- Boonen, A. J. H., Kolkman, M. E. & Kroesbergen, E. H. (2011). The relation between teachers' math talk and the acquisition of number sense within kindergarten classrooms. *Journal of School Psychology*, 49(3), 281-299.
- Bowman, N. A. (2012). Effect sizes and statistical methods for meta-analysis in higher education. *Research in Higher Education*, 53(3), 375-382.
- Bradbury, B., Corak, M., Waldfogel, J. & Washbrook, E. (2015). *Too many children left behind*. New York, NY: Russel Sage Foundation.
- Bradley, J. & Garrity, S. (2013). How to set up and arrange a Head Start classroom. In S. W. Gilford (Eds.), *Learning from Head Start: A teacher's guide to school readiness* (pp. 263-274). Lanham, MD: Rowman & Littlefield Education.
- Bradshaw, P., Lewis, G. & Hughes, T. (2014). *Growing Up in Scotland characteristics of pre-school provision and their association with child outcomes*. Edinburgh: Scottish Government.
- Brady, P. & Bowd, A. (2005). Mathematics anxiety, prior experience and confidence to teach mathematics among pre-service education students. *Teachers and Teaching: Theory and Practice*, 11(1), 37-46.
- Brand, B. R. & Wilkins, J. L. M. (2007). Using self-efficacy as a construct for evaluating science and mathematics methods courses. *Journal of Science Teacher Education*, 18(2), 297-317.
- Broberg, A. G., Hwang, C. P. & Chace, S. V. (1993, March). *Effects of day care on elementary school performance and adjustment*. Paper presented at the 60th Anniversary Meeting of the Society for Research in Child Development, New Orleans, Louisiana, USA.
- Broberg, A. G., Wessels, H., Lamb, M. E. & Hwang, C. P. (1997). Effects of day care on the development of cognitive abilities in 8-year-olds: A longitudinal study. *Developmental psychology*, 33(1), 62-69.
- Broekhuizen, M., Leseman, P., Moser, T. & van Trijp, K. (2015). *Stakeholders study: Values, beliefs and concerns of parents, staff and policy representatives regarding ECEC services in nine European countries: First report on parents*. Retrieved from CARE website: http://ecec-care.org/fileadmin/careproject/Publications/reports/CARE_WP6_D6_2_European_ECEC__Stakeholder_study_FINAL.pdf

- Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: Research perspectives. *Developmental Psychology*, 22(6), 723-742.
- Bronfenbrenner, U. & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner (Eds.), *Theoretical models of human development* (Handbook of child psychology, 6 ed., Vol. 2, 1 S. 793-828). Hoboken, NJ: Wiley.
- Brown, E. T. (2003). *The influence of teachers' efficacy and beliefs on mathematics instruction in the early childhood classroom*. Unpublished doctoral dissertation, University of Louisville, Louisville.
- Brown, E. T. (2005). The influence of teachers' efficacy and beliefs regarding mathematics instruction in the early childhood classroom. *Journal of Early Childhood Teacher Education*, 26(3), 239-257.
- Brown, E. T., Molfese, V. J. & Molfese, P. (2008). Preschool student learning in literacy and mathematics: Impact of teacher experience, qualifications, and beliefs on an at-risk sample. *Journal of Education for Students Placed at Risk (JESPAR)*, 13(1), 106-126.
- Bryk, A. S. & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis*. Newbury Park, CA: Sage.
- Buchanan, T. K., Burts, D. C., Bidner, J., White, V. F. & Charlesworth, R. (1998). Predictors of the developmental appropriateness of the beliefs and practices of first, second, and third grade teachers. *Early Childhood Research Quarterly*, 13(3), 459-483.
- Burchinal, M., Cryer, D., Clifford, R. M. & Howes, C. (2002). Caregiver training and classroom quality in child care centers. *Applied Developmental Science*, 6(1), 2-11.
- Burchinal, M., Vandergrift, N., Pianta, R. & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, 25(2), 166-176.
- 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(2), 140-165.
- Burger, K. (2014). Effective early childhood care and education: Successful approaches and didactic strategies for fostering child development. *European Early Childhood Education Research Journal*, 23(5), 743-760.
- Byrne, B. M. (2010). *Structural equation modeling with amos: Basic concepts, applications, and programming* (2nd). New York, NY/Hove, East Sussex: Routledge.

- Cabell, S. Q., DeCoster, J., LoCasale-Crouch, J., Hamre, B. K. & Pianta, R. C. (2013). Variation in the effectiveness of instructional interactions across preschool classroom settings and learning activities. *Early Childhood Research Quarterly*, 28(4), 820-830.
- Camehl, G. F., Schober, P. S. & Spieß, K. C. (2012). A SOEP-related study: Early childhood education and care quality in the Socio-Economic Panel (K2ID-SOEP). In S. Gerstorff & J. Schupp (Eds.), *SOEP Wave Report 2014* (pp. 31-34). Berlin: DIW Berlin.
- Camilli, G., Vargas, S., Ryan, S. & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *The Teachers College Record*, 112(3), 579-620.
- Card, N. (2012). *Applied meta-analysis for social science research*. New York: Guilford Publications.
- Carle, U. & Wenzel, D. (2007). Vorschulische Bildung im Kindergarten. In M. Haring, C. Rohlfis & C. Palentien (Hrsg.), *Perspektiven der Bildung* (S. 185–202). Wiesbaden: Springer-Verlag.
- Chambers, B., Cheung, A., Slavin, R. E., Smith, D. & Laurenzano, M. (2010). Effective early childhood education programs: A systematic review. In *Best Evidence Encyclopedia (BEE)* (pp. 1-60). Baltimore, MD: Johns Hopkins University, Center for Research and Reform in Education.
- Chapman, C., Muijs, D., Reynolds, D., Sammons, P. & Teddlie, C. (Eds.). (2015). *Routledge international handbook of educational effectiveness and improvement*. Abingdon, UK / New York, NY: Routledge.
- Chen, J.-Q. & Chang, C. (2006). Testing the whole teacher approach to professional development: A study of enhancing early childhood teachers' technology proficiency. *Early Childhood Research & Practice*, 8(1), 1-18.
- Chen, J.-Q. & McCray, J. (2012). The what, how, and why of effective teacher professional development in early mathematics education. *NHSA Dialog*, 15(1), 113–121.
- Chen, J.-Q., McCray, J., Adams, M. & Leow, C. (2014). A survey study of early childhood teachers' beliefs and confidence about teaching early math. *Early Childhood Education Journal*, 42(6), 367-377.
- Cheung, M. W. (2014). Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychol Methods*, 19(2), 211-229.
- Clements, D. H. & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. *Journal for Research in Mathematics Education*, 38(2), 136-163.

- Clements, D. H. & Sarama, J. (2011). Early childhood mathematics intervention. *Science*, 333(6045), 968-970.
- Clifford, R. (2005). Structure and stability of the early childhood environment rating scale. In H. Schonfeld, S. O'Brien & T. Walsh (Eds.), *Questions of quality* (pp. 12-21). Dublin, Ireland: The Centre for Early Childhood Development & Education, The Gate Lodge, St. Patrick's College.
- Clifford, R., Reszka, S. S. & Roßbach, H.-G. (2010). *Reliability and validity of the early childhood environment rating scale*. Retrieved from ersi website: <http://www.ersi.info/PDF/ReliabilityEcers.pdf>
- Cobanoglu, R. (2011). *Teacher self-efficacy and teaching beliefs as predictors of curriculum implementation in early childhood education*. Unpublished master's thesis, Middle East Technical University, Ankara.
- Cooper, H. (2009). *Research synthesis and meta-analysis: A step-by-step approach* (4th eds.). Thousand Oaks, CA: Sage Publications.
- Copley, J. V. (2004). The early childhood collaborative: A professional development model to communicate and implement the standard. In D. H. Clements & J. Samara (Eds.), *Engaging young children in mathematics. Standards for early childhood mathematics education* (pp. 401-414). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Copley, J. V. & Padrón, Y. (1999). Preparing teachers of young learners: Professional development of early childhood teachers in mathematics and science. In G. D. Nelson (Eds.), *Dialogue on early childhood science, mathematics, and technology education* (pp. 117-129). Washington, DC: American Association for the Advancement of Science.
- Creemers, B. & Kyriakides, L. (2009). Situational effects of the school factors included in the dynamic model of educational effectiveness. *South African Journal of Education*, 29(3), 293–315.
- Creemers, B., Kyriakides, L. & Sammons, P. (2010). *Methodological advances in educational effectiveness research*. Abingdon, UK / New York, NY: Routledge.
- Cross, C. T., Woods, T. A. & Schweingruber, H. (Eds.). (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. Washington, DC: National Academies Press.
- Cryer, D., Tietze, W., Burchinal, M., Leal, T. & Palacios, J. (1999). Predicting process quality from structural quality in preschool programs: A cross-country comparison. *Early Childhood Research Quarterly*, 14(3), 339–361.

- Curby, T. W., Grimm, K. J. & Pianta, R. C. (2010). Stability and change in early childhood classroom interactions during the first two hours of a day. *Early Childhood Research Quarterly*, 25(3), 373-384.
- Curran, P. J. & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods*, 14(2), 81-100.
- Dalli, C., White, E. J., Rockel, J., Duhn, I., Buchanan, E., Davidson, S. et al. (2011). *Quality early childhood education for under-two-year-olds: What should it look like? A literature review*. Wellington, New Zealand: Ministry of Education.
- Darling-Hammond, L., Chung, R. & Frelow, F. (2002). Variation in teacher preparation: How well do different pathways prepare teachers to teach? *Journal of Teacher Education*, 53(4), 286-302.
- de Haan, A., Elbers, E., Hoofs, H. & Leseman, P. (2013). Targeted versus mixed preschools and kindergartens: effects of class composition and teacher-managed activities on disadvantaged children's emergent academic skills. *School Effectiveness and School Improvement*, 24(2), 177-194.
- de Haan, A., Leseman, P. & Elbers, E. (2011). *Pilot gemengde groepen 2007-2010*. Utrecht: Universiteit Utrecht.
- Dehaene, S. (2011). *The number sense how the mind creates mathematics*. New York: Oxford University Press.
- Deppe, V. (2011). *Anforderungen an die Ausbildung von Erzieherinnen und Erziehern: Ergebnisse einer qualitativen Befragung von Fachschul- und Abteilungsleitungen*. München: DJI.
- Deutscher, T. & Selter, C. (2013). Frühe mathematische Bildung – Forschungsbefunde und Förderkonzepte. In M. Stamm & D. Edelmann (Hrsg.), *Handbuch frühkindliche Bildungsforschung* (S. 543-556). Wiesbaden: Springer Fachmedien Wiesbaden.
- DIN. (2005). *DIN EN ISO 9000:2005 - Qualitätsmanagementsysteme - Grundlagen und Begriffe*. Berlin: Beuth Verlag.
- Diskowski, D. (2009). Bildungspläne für Kindertagesstätten – ein neues und noch unbegriffenes Steuerungsinstrument. In H.-G. Roßbach & H.-P. Blossfeld (Hrsg.), *Frühpädagogische Förderung in Institutionen* (Sonderheft 11 der Zeitschrift für Erziehungswissenschaft, S. 47 - 61). Wiesbaden: VS Verlag für Sozialwissenschaften.
- 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.* (Beiheft Nr. 41 der Zeitschrift für Pädagogik, S. 73–92). Weinheim: Beltz.
- Dooley, T., Dunphy, E. & Shiel, G. (2012). Mathematics in early childhood and primary education (3-8 years). Dublin: National Council for Curriculum and Assessment.
- Döring, N. & Bortz, J. (2014). *Forschungsmethoden und Evaluation*. Heidelberg: Springer.
- Dunekacke, S., Jenßen, L., Baack, W., Tengler, M., Wedekind, H., Grassmann, M. et al. (2013). Was zeichnet eine kompetente pädagogische Fachkraft im Bereich Mathematik aus? Modellierung professioneller Kompetenz für den Elementarbereich. In G. Greefrath, F. Käpnick & M. Stein (Hrsg.), *Beiträge zum Mathematikunterricht 2013* (S. 280–283). Münster: WTM.
- Durda, T. (2015). *Eltern als Bildungspartner: Effekte von Elternarbeit in Kindertageseinrichtungen auf den sprachlichen Entwicklungsstand von Kindern*. Unveröffentlichte Masterarbeit, Otto-Friedrich-Universität Bamberg, Bamberg.
- EACEA/Eurydice. (2009). *Early childhood education and care in europe: Tackling social and cultural inequalities*. Brussels: Eurydice.
- Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S. et al. (2006). Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly*, 21(2), 174-195.
- Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D. et al. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child development*, 78(2), 558–580.
- Ebert, S., Lockl, K., Weinert, S., Anders, Y., Kluczniok, K. & Roßbach, H.-G. (2013). Internal and external influences on vocabulary development in preschool children. *School Effectiveness and School Improvement*, 24(2), 138-154.
- Eccles, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L. et al. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Eds.), *Achievement and achievement motivation* (pp. 75-146). San Francisco, CA: Freeman.
- ECCRN, N. (2002a). Child-care structure -> process -> outcome: Direct and indirect effects of child-care quality on young children's development. *Psychological Science*, 13(3), 199-206.
- ECCRN, N. (2002b). Early child care and children's development prior to school entry: Results from the NICHD study of early child care. *American Educational Research Journal*, 39(1), 133-164.

- ECCRN, N. (2003). Does quality of child care affect child outcomes at age 4 1/2? *Developmental Psychology*, 39(3), 451-469.
- ECCRN, N. (2005). Early child care and children's development in the primary grades: Follow-up results from the NICHD study of early child care. *American Educational Research Journal*, 42(3), 537-570.
- Elliot, C., Smith, P. & McCulloch, K. (1996). *British ability scales second edition (BAS II)*. Windsor, Berkshire: NFER-NELSON Publishing Company.
- Enders, C. K. (2001). The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educational and Psychological Measurement*, 61(5), 713-740.
- European Commission/EACEA/Eurydice/Eurostat. (2014). *Key data on early childhood education and care in Europe. 2014 Edition*. Eurydice and Eurostat Report. Luxembourg: Publications Office of the European Union.
- Evangelou, M., Sylva, K., Kyriacou, M., Wild, M. & Glenny, G. (2009). *Early years learning and development: Literature review*. London: Department for Children, Schools and Families.
- Evans, B. R., Leonard, J., Krier, K. & Ryan, S. (2013). The influence of a reform-based mathematics methods course on preservice teachers' beliefs. *Journal of Educational Research and Practice*, 3(1), 79-92.
- Ferla, J., Valcke, M. & Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences*, 19(4), 499-505.
- Fives, H. & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer & M. Zeidner (Eds.), *Individual differences and cultural and contextual factors*. (APA educational psychology handbook Vol. 2, pp. 471-499). Washington, DC: American Psychological Association.
- HIGH/SCOPE. (2008). *The High/Scope preschool educational approach: A prospectus for pre-kindergarten programs*. Ypsilanti, MI: High/Scope Press.
- Fried, A. (2011, Februar). *Mathematische Grundbildung in niedersächsischen Kindertageseinrichtungen*. Beitrag präsentiert auf der 45. Jahrestagung der Gesellschaft für Didaktik der Mathematik, Freiburg.

- Fried, A. (2012). Mathematische Erfahrungen im Kindergarten. Eine Fragebogenstudie in niedersächsischen Kindertageseinrichtungen. Hildesheim: Franzbecker.
- Fried, L., Hoeft, M., Isele, P., Stude, J. & Wexeler, W. (2012). *Schlussbericht zur Wissenschaftlichen Flankierung des Verbundprojekts „TransKiGs – Stärkung der Bildungs- und Erziehungsqualität in Kindertageseinrichtungen und Grundschule – Gestaltung des Übergangs“*. Dortmund: TU Dortmund.
- Friedrich, G. & Munz, H. (2004). *Projekt- und Evaluationsbericht „Komm mit ins Zahlenland“: Ein „ganzheitliches“ Frühförderkonzept am Beispiel elementarer Mathematik*. Zugriff am 08. Juni 2016 unter <http://www.ifvl.de/material/Projektbericht-Zahlenland.pdf>
- Friedrich, G. & Munz, H. (2011). Förderung schulischer Vorläuferfähigkeiten durch das didaktische Konzept „Komm mit ins Zahlenland“. *Psychologie in Erziehung und Unterricht*(2), 134-146.
- Fröbel, F. (1850). Kinderbeschäftigung. Anleitung zum Papierfalten. Eine entwickelnd-erziehende und unterhaltend-belehrende Kinderbeschäftigung für Kinder von 5 bis 7 Jahren und darüber, unter eingehender Mitwirkung von leitenden Erwachsenen. *Friedrich Fröbels Wochenschrift*, 371-388.
- Fröhlich-Gildhoff, K., Nentwig-Gesemann, I. & Pietsch, S. (2011). *Kompetenzorientierung in der Qualifizierung frühpädagogischer Fachkräfte: eine Expertise der Weiterbildungsinitiative Frühpädagogische Fachkräfte*. München: DJI.
- Fthenakis, W. E., Schmitt, A., Daut, M., Eitel, A. & Wendell, A. (2009). *Natur-Wissen schaffen. Band 2: Frühe mathematische Bildung*. Troisdorf: Bildungsverlag EINS.
- Fukkink, R. G. & Lont, A. (2007). Does training matter? A meta-analysis and review of caregiver training studies. *Early Childhood Research Quarterly*, 22(3), 294-311.
- Ganzeboom, H. B., De Graaf, P. M. & Treiman, D. J. (1992). A standard international socioeconomic index of occupational status. *Social science research*, 21(1), 1-56.
- Gasteiger, H. (2007). *Stand der mathematischen Kompetenzdiagnosen am Übergang von Kindertagesstätten und Grundschule und zukünftige Perspektiven*. Berlin: Senatsverwaltung für Bildung, Wissenschaft und Forschung.
- Gasteiger, H. (2010). *Elementare mathematische Bildung im Alltag der Kindertagesstätte*. Münster: Waxmann.
- Gasteiger, H. (2012). Fostering early mathematical competencies in natural learning situations: Foundation and challenges of a competence-oriented concept of

- mathematics education in kindergarten. *Journal für Mathematik-Didaktik*, 33(2), 181-201.
- Gasteiger, H. (2014). Professionalization of early childhood educators with a focus on natural learning situations and individual development of mathematical competencies: Results from an evaluation study. In U. Kortenkamp, B. Brandt, C. Benz, G. Krummheuer, S. Ladel & R. Vogel (Eds.), *Early mathematics learning* (pp. 275-290). New York, NY: Springer.
- Gasteiger, H. (2017). Frühe mathematische Bildung – sachgerecht, kindgemäß, anschlussfähig. In S. Schuler, C. Streit & G. Wittmann (Hrsg.), *Perspektiven mathematischer Bildung im Übergang vom Kindergarten zur Grundschule* (S. 9-26). Wiesbaden: Springer Fachmedien Wiesbaden.
- Geier, B. & Riedel, B. (2009). Ungleichheiten der Inanspruchnahme öffentlicher frühpädagogischer Angebote. Einflussfaktoren und Restriktionen elterlicher Betreuungsentscheidungen. In H.-G. Roßbach & H.-P. Blossfeld (Hrsg.), *Frühpädagogische Förderung in Institutionen* (S. 11–28). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Geiser, C. (2011). *Datenanalyse mit Mplus: Eine anwendungsorientierte Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Ghazvini, A. & Mullis, R. L. (2002). Center-based care for young children: Examining predictors of quality. *The Journal of Genetic Psychology*, 163(1), 112-125.
- Ginsburg, H. P. & Golbeck, S. L. (2004). Thoughts on the future of research on mathematics and science learning and education. *Early Childhood Research Quarterly*, 19(1), 190-200.
- Ginsburg, H. P., Lee, J. S. & Boyd, J. S. (2008). Mathematics education for young children: What it is and how to promote it. *Society for Research in Child Development*, 22(1), 3–22.
- Goldstein, H. (2011). *Multilevel statistical models* (4). West Sussex, United Kingdom: John Wiley & Sons, Ltd.
- Goodwin, B. (2011). *Simply better: Doing what matters most to change the odds for student success*. Denver, CO: McREL.
- Gordon, R. A., Fujimoto, K., Kaestner, R., Korenman, S. & Abner, K. (2013). An assessment of the validity of the ECERS-R with implications for measures of child care quality and relations to child development. *Developmental Psychology*, 49(1), 146-160.

- Greenes, C., Ginsburg, H. P. & Balfanz, R. (2004). Big math for little kids. *Early childhood research quarterly*, 19(1), 159-166.
- Griffin, S. (2004). Building number sense with Number Worlds: a mathematics program for young children. *Early Childhood Research Quarterly*, 19(1), 173-180.
- Griffin, S. & Case, R. (1997). Re-thinking the primary school math curriculum: An approach based on cognitive science. *Issues in Education*, 3(1), 1-49.
- Guarino, C. M., Hamilton, L. S., Lockwood, J. R. & Rathbun, A. H. (2006). *Teacher qualifications, instructional practices, and reading and mathematics gains of kindergartners*. (NCES 2006-031). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Guo, Y., Justice, L. M., Sawyer, B. & Tompkins, V. (2011). Exploring factors related to preschool teachers' self-efficacy. *Teaching and Teacher Education*, 27(5), 961-968.
- Halle, T., Vick Whittaker, J. E. & Anderson, R. (2010). *Quality in early childhood care and education settings: A compendium of measures*. Washington, DC: Child Trends.
- Hamre, B., Hatfield, B., Pianta, R. & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: Associations with preschool children's development. *Child Development*, 85(3), 1257-1274.
- Hardy, I. & Steffensky, M. (2014). Prozessqualität im Kindergarten: Eine domänenspezifische Perspektive. *Unterrichtswissenschaft*, 42(2), 101-116.
- Harms, T., Clifford, R. M. & Cryer, D. (1998). *Early Childhood Rating Scale, Revised Edition (ECERS-R)*. Vermont: Teachers College Press.
- Harms, T., Clifford, R. M. & Cryer, D. (2014). *Early Childhood Environment Rating Scale (ECERS-3)*. New York/London: Teachers College Press.
- Hasemann, K. & Gasteiger, H. (2014). Mathematiklernen im Übergang Kindergarten – Grundschule. In F. Padberg (Hrsg.), *Anfangsunterricht Mathematik* (S. 43-61). Berlin, Heidelberg: Springer.
- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), 1900-1902.
- Hellmich, F. (2007). Möglichkeiten der Förderung mathematischer Vorläuferfähigkeiten im vorschulischen Bereich. *Bildungsforschung*, 4(1), 1-16.
- Hembree, R. (1999). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, 21(1), 33-46.
- Henschen, E. & Teschner, M. (2013). Von Kindergärten, Kindheitspädagoginnen und der Mathematik mit Bauklötzen. In J. Sprenger, A. Wagner & M. Zimmermann (Hrsg.),

- Mathematik lernen, darstellen, deuten, verstehen* (S. 53-67). Wiesbaden: Springer Fachmedien.
- Herbison, P., Hay-Smith, J. & Gillespie, W. J. (2006). Adjustment of meta-analyses on the basis of quality scores should be abandoned. *J Clin Epidemiol*, 59(12), 1249-1256.
- Higgins, J. P. T. & Green, S. (Eds.). (2008). *Cochrane handbook for systematic reviews of interventions*. Chichester, UK: John Wiley & Sons.
- Hofer, K. G. (2010). How measurement characteristics can affect ECERS-R scores and program funding. *Contemporary Issues in Early Childhood*, 11(2), 175-191.
- Hox, J. J. (2010). *Multilevel analysis. Techniques and applications* (2). New York, NY Hove, East Sussex: Routledge.
- Hyun, E. & Marshall, J. (2003). Teachable-moment-oriented curriculum practice in early childhood education. *Journal of Curriculum Studies*, 35(1), 111-127.
- Isele, P. (2014). *Entwicklung und Validierung eines Beobachtungsverfahrens zur Erfassung von förderrelevanter Prozessqualität beim Übergang vom Elementar in den Primarbereich*. Unveröffentlichte Dissertation, Technische Universität Dortmund, Dortmund.
- Janssen, R. (2010). *Die Ausbildung frühpädagogischer Fachkräfte an Berufsfachschulen und Fachschulen: Eine Analyse im Ländervergleich* München: DJI.
- JMK & KMK. (2004). *Gemeinsamer Rahmen der Länder für die frühe Bildung in Kindertageseinrichtungen. Beschluss der Jugendministerkonferenz vom 13./14.05.2004. Beschluss der Kultusministerkonferenz vom 03./04.06.2004*. Bonn: Kultusministerkonferenz.
- Jordan, N. C., Kaplan, D., Ramineni, C. & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, 45(3), 850-867.
- Jörns, C., Schuchardt, K., Mähler, C. & Grube, D. (2013). Alltagsintegrierte Förderung numerischer Kompetenzen im Kindergarten. *Frühe Bildung*, 2(2), 1-8.
- Jungmann, T. & Koch, K. (Hrsg.). (2017). *Professionalisierung pädagogischer Fachkräfte in Kindertageseinrichtungen*. Wiesbaden: Springer Fachmedien.
- Juni, P., Witschi, A., Bloch, R. & Egger, M. (1999). The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA*, 282(11), 1054-1060.
- Kalder, R. S. & Lesik, S. A. (2011). A classification of attitudes and beliefs towards mathematics for secondary mathematics pre-service teachers and elementary pre-

- service teachers: An exploratory study using latent class analysis. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, 5, 1-21.
- Kane, T. J. & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Seattle, WA: MET Project, Bill & Melinda Gates Foundation.
- Karoly, L. A., Kilburn, R. M. & Cannon, J. S. (2005). Proven benefits of early childhood interventions. Santa Monica, CA: RAND Corporation.
- Katz, L. (1996). Die Qualität der Früherziehung in Betreuungseinrichtungen - Fünf Perspektiven. In W. Tietze (Hrsg.), *Früherziehung. Trends, internationale Forschungsergebnisse, Praxisorientierungen* (S. 226-239). Neuwied: Luchterhand.
- Kaufmann, S. (2011). *Handbuch für die frühe mathematische Bildung*. Braunschweig: Schroedel.
- Kelley, W. P. & Tomhave, W. K. (1985). A study of math anxiety/math avoidance in preservice elementary teachers. *The Arithmetic Teacher*, 32(5), 51-53.
- Kim, R. S. (2011). *Standardized regression coefficients As indices of effect sizes in meta-analysis*. Unpublished doctoral dissertation, Florida State University, Tallahassee, Florida.
- Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M. & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk". *Developmental Psychology*, 42(1), 59-69.
- Kluczniok, K., Anders, Y. & Ebert, S. (2011). Fördereinstellungen von Erzieherinnen - Einflüsse auf die Gestaltung von Lerngelegenheiten im Kindergarten und die kindliche Entwicklung am Beispiel von frühen Rechenfertigkeiten. *Frühe Bildung*, 0(1), 13-21.
- Kluczniok, K., Anders, Y., Sechtig, J. & Roßbach, H.-G. (2016). Influences of an academically oriented preschool curriculum on the development of children – are there negative consequences for the children's socio-emotional competencies? *Early Child Development and Care*, 186(1), 117-139.
- Kluczniok, K., Lehrl, S., Kuger, S. & Roßbach, H.-G. (2013). Quality of the home learning environment during preschool age – Domains and contextual conditions. *European Early Childhood Education Research Journal*, 21(3), 420-438.
- Kluczniok, K. & Roßbach, H.-G. (2014). Conceptions of educational quality for kindergartens. *Zeitschrift für Erziehungswissenschaft*, 17(6), 145-158.
- KMK. (2004). *Bildungsstandards im Fach Mathematik für den Primarbereich*. Beschluss der Kultusministerkonferenz vom 15.10.2004. Bonn: Luchterhand.

- KMK. (2009). *Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung*. Beschluss der Kultusministerkonferenz vom 10.12.2009. Köln: Wolters Kluwer.
- König, A. (2007). Dialogisch-entwickelnde Interaktionsprozesse als Ausgangspunkt für die Bildungsarbeit im Kindergarten. *bildungsforschung*, 4(1), 1-21.
- Konstantopoulos, S. (2006). Trends of school effects on student achievement: Evidence from NLS: 72, HSB: 82, and NELS: 92. *Teachers College Record*, 108(12), 2550-2581.
- Kowalski, K., Pretti-Frontczak, K. & Johnson, L. (2001). Preschool teachers' beliefs concerning the importance of various developmental skills and abilities. *Journal of Research in Childhood Education*, 16(1), 5-14.
- Krajewski, K., Grübing, M. & Peter-Koop, A. (2009). Die Entwicklung mathematischer Kompetenzen bis zum Beginn der Grundschulzeit. In A. Heinze & M. Grübing (Hrsg.), *Mathematiklernen vom Kindergarten bis zum Studium* (S. 17-34). Münster/New York/München/Berlin: Waxmann.
- Krajewski, K., Nieding, G. & Schneider, W. (2008). Kurz- und langfristige Effekte mathematischer Frühförderung im Kindergarten durch das Programm „Mengen, zählen, Zahlen“. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40(3), 135-146.
- Krajewski, K., Renner, A., Nieding, G. & Schneider, W. (2009). Frühe Förderung von mathematischen Kompetenzen im Vorschulalter. In H. G. Rossbach & H. P. Blossfeld (Hrsg.), *Frühpädagogische Förderung in Institutionen* (S. 91–103). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Krapp, A. & Ryan, R. M. (2002). Selbstwirksamkeit und Lernmotivation. Eine kritische Betrachtung der Theorie von Bandura aus der Sicht der Selbstbestimmungstheorie und der pädagogisch-psychologischen Interessentheorie. In M. Jerusalem & D. Hopf (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen* (S. 54–82). Weinheim: Beltz.
- Krüger, H.-H. & Grunert, C. (Hrsg.). (2010). *Handbuch Kindheits- und Jugendforschung* (2. Aufl.). Weisbaden: VS Verlag für Socialwissenschaften.
- Krumm, V., Wetzel, G., Tietze, W., Hundertmark-Mayser, J., Roßbach, H. G., Bairrao, J. et al. (1997). *European Child Care and Education Study - Cross national analyses of the quality of effects of early childhood programmes on children's development*. Brussels: European Union.

- Kuger, S. & Kluczniok, K. (2008). Prozessqualität im Kindergarten - Konzept, Umsetzung und Befunde. In H. G. Roßbach & H. P. Blossfeld (Hrsg.), *Frühpädagogische Förderung in Institutionen* (S. 159–178). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kuger, S., Kluczniok, K., Kaplan, D. & Roßbach, H.-G. (2016). Stability and patterns of classroom quality in German early childhood education and care. *School Effectiveness and School Improvement*, 27(3), 418-440.
- Kuger, S., Sechtig, J. & Anders, Y. (2012). Kompensatorische (Sprach-)Förderung. *Frühe Bildung*, 1(4), 181-193.
- Kwon, Y.-I. (2002). Changing curriculum for early childhood education in England. *Early Childhood Research & Practice*, 4(2), 2-13.
- La Paro, K. M., Thomason, A. C., Lower, J. K., Kintner-Duffy, V. L. & Cassidy, D. J. (2012). Examining the definition and measurement of quality in early childhood education: A review of studies using the ECERS-R from 2003 to 2010. *Early Childhood Research & Practice*, 14(1), 1-17.
- Lake, V. E. & Kelly, L. (2014). Female preservice teachers and mathematics: Anxiety, beliefs, and stereotypes. *Journal of Early Childhood Teacher Education*, 35(3), 262-275.
- Larra, F. (2005). Ansätze zur Steuerung pädagogischer Qualität in vorschulischen Einrichtungen. In Sachverständigenkommission 12. Kinder und Jugendbericht (Hrsg.), *Entwicklungspotenziale institutioneller Angebote im Elementarbereich* (S. 235-268). München: Verlag Deutsches Jugendinstitut.
- Leal, T., Abreu-Lima, I. M. P. & Cadima, J. (2013, August). *Associations between dimensions of quality and children outcomes in preschool and first grade – A Portuguese study*. Paper presented at the 15th Biennial Conference of the European Association for Research in Learning and Instruction (EARLI), Munich.
- Ledig, M. (2011). *Fort- und Weiterbildung von Lehrkräften an Fachschulen für Sozialpädagogik: Ergebnisse einer Interviewstudie mit Schulleitungen* (No. 978-3-86379-029-5). München: DJI.
- Lee, J. S. & Ginsburg, H. P. (2007). Preschool teachers' beliefs about appropriate early literacy and mathematics education for low- and middle-socioeconomic status children. *Early Education & Development*, 18(1), 111-143.
- LeFevre, J.-A., Skwarchuk, S.-L., Smith-Chant, B. L., Fast, L., Kamawar, D. & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early

- school years. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 41(2), 55-66.
- Lehrl, S. (2013). Die häusliche Lernumwelt im Vorschulalter–wie Eltern die kindliche Kompetenzentwicklung unterstützen. In G. Faust (Hrsg.), *Einschulung: Ergebnisse aus der Studie 'Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul-und Schulalter (BiKS)'* (S. 51–68). Münster/New York/München/Berlin: Waxmann.
- Lehrl, S., Kluczniok, K. & Roßbach, H.-G. (2016). Longer-term associations of preschool education: The predictive role of preschool quality for the development of mathematical skills through elementary school. *Early Childhood Research Quarterly*, 36(3), 475-488.
- Lehrl, S., Kuger, S. & Anders, Y. (2014). Soziale Disparitäten beim Zugang zu Kindergartenqualität und differenzielle Konsequenzen für die vorschulische mathematische Entwicklung. *Unterrichtswissenschaft*, 42(4), 132-151.
- Lenkeit, J. (2012). How effective are educational systems? A value-added approach to measure trends in PIRLS. *Journal for educational research online*, 4(2), 143–173.
- Lenkeit, J. (2013). Effectiveness measures for cross-sectional studies: a comparison of value-added models and contextualised attainment models. *School Effectiveness and School Improvement*, 24(1), 1-25.
- Lerkkanen, M.-K., Kiuru, N., Pakarinen, E., Viljaranta, J., Poikkeus, A.-M., Rasku-Puttonen, H. et al. (2012). The role of teaching practices in the development of children's interest in reading and mathematics in kindergarten. *Contemporary Educational Psychology*, 37(4), 266-279.
- Leseman, P. P. M. & Slot, P. L. (2014). Breaking the cycle of poverty: challenges for European early childhood education and care. *European Early Childhood Education Research Journal*, 22(3), 314-326.
- Leyendecker, B., Agache, A. & Madsen, S. (2014). Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit (NUBBEK) – Design, Methodenüberblick, Datenzugang und das Potenzial zu Mehrebenenanalysen. *Zeitschrift für Familienforschung*, 26(2), 244-258.
- Linder, S. M. (2012). Building content and communities: Developing a shared sense of early childhood mathematics pedagogy. *Journal of Early Childhood Teacher Education*, 33(2), 109-126.

- Locuniak, M. N. & Jordan, N. C. (2008). Using kindergarten number sense to predict calculation fluency in second grade. *Journal of Learning Disabilities, 41*(5), 451-459.
- 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. *Economics of Education Review, 26*(1), 52-66.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau, 58*(2), 103-117.
- Maas, C. J. & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology, 1*(3), 86-92.
- Maas, C. J., Hox, J. J. & Lensvelt-Mulders, G. J. (2004). Longitudinal meta-analysis. *Quality and Quantity, 38*(4), 381-389.
- Manning, M., Homel, R. & Smith, C. (2010). A meta-analysis of the effects of early developmental prevention programs in at-risk populations on non-health outcomes in adolescence. *Children and Youth Services Review, 32*(4), 506-519.
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D. et al. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child development, 79*(3), 732-749.
- McCray, J. S. & Chen, J.-Q. (2012). Pedagogical content knowledge for preschool mathematics: Construct validity of a new teacher interview. *Journal of Research in Childhood Education, 26*(3), 291-307.
- McMullen, M. B. (1997). The effects of early childhood academic and professional experience on self perceptions and beliefs about developmentally appropriate practices. *Journal of Early Childhood Teacher Education, 18*(3), 55-68.
- Melchers, P. & Preuß, U. (2003). *Kaufman-Assessment Batterie for Children (K-ABC), Deutsche Version. Individualtest zur Messung von Intelligenz und Fertigkeiten bei Kindern*. Göttingen: Hogrefe.
- Melhuish, E. (2001). The quest for quality in early day care and preschool experience continues. *International Journal of Behavioral Development, 25*(1), 1-6.
- Melhuish, E. (2011). Preschool matters. *Science, 333*(6040), 299-300.
- Melhuish, E., Ereky-Stevens, K., Petrogiannis, K., Ariescu, A., Penderi, E., Rentzou, K. et al. (2015). *A review of research on the effects of early childhood education and care (ecec) upon child development*.
- Melhuish, E., Quinn, L., Sylva, K., Sammons, P., Siraj-Blatchford, I. & Taggart, B. (2013). Preschool affects longer term literacy and numeracy: results from a general population

- longitudinal study in Northern Ireland. *School Effectiveness and School Improvement*, 24(2), 234-250.
- Ministerium für Bildung, Jugend & Sport. (2015). *Thüringer Bildungsplan bis 18 Jahre: Bildungsansprüche von Kindern und Jugendlichen*. Erfurt: Thüringer Ministerium für Bildung, Jugend und Sport.
- Mischo, C., Wahl, S., Hendler, J. & Strohmmer, J. (2012). Pädagogische Orientierungen angehender frühpädagogischer Fachkräfte an Fachschulen und Hochschulen. *Frühe Bildung*, 1(1), 34-44.
- Mischo, C., Wahl, S., Strohmmer, J. & Hendler, J. (2012). Knowledge orientations of prospective early childhood teachers: A study of students' scientific versus subjective orientations in teacher education courses in Germany. *Journal of Early Childhood Teacher Education*, 33(2), 144-162.
- Montie, J. E., Xiang, Z. P. & Schweinhart, L. J. (2006). Preschool experience in 10 countries: Cognitive and language performance at age 7. *Early Childhood Research Quarterly*, 21(3), 313-331.
- Morawiak, U., Schulz, A., Jungmann, T. & Koch, K. (2017). Professionalisierungsangebote im KOMPASS-Projekt. In T. Jungmann & K. Koch (Hrsg.), *Professionalisierung pädagogischer Fachkräfte in Kindertageseinrichtungen* (S. 29-64). Wiesbaden: Springer Fachmedien.
- Moseley, C. & Utley, J. (2006). The effect of an integrated science and mathematics content-based course on science and mathematics teaching efficacy of preservice elementary teachers. *Journal of Elementary Science Education*, 18(2), 1-12.
- Moser, U. & Studer, C. (2013). Bildungsstatistik und Bildungsmonitoring. In M. Stamm & D. Edelmann (Hrsg.), *Handbuch frühkindliche Bildungsforschung* (S. 875-885). Wiesbaden: Springer Fachmedien.
- Moss, P. & Pence, A. (Eds.). (1994). *Valuing quality in early childhood services: New approaches to defining quality*. London: Sage Publications.
- NAEYC & NCTM. (2010). *Early childhood mathematics: Promoting good beginnings*. Washington, DC: NAEYC.
- NCTM. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- Network on Early Childhood Education and Care. (2013). *Literature review on monitoring quality in early childhood education and care (ECEC)*. Paris: OECD.

- Neumann, I., Duchhardt, C., Grüßing, M., Heinze, A., Knopp, E. & Ehmke, T. (2013). Modeling and assessing mathematical competence over the lifespan. *Journal for educational research online*, 5(2), 80-109.
- Niklas, F. & Schneider, W. (2012). Einfluss von „Home Numeracy Environment“ auf die mathematische Kompetenzentwicklung vom Vorschulalter bis Ende des 1. Schuljahres. *Zeitschrift für Familienforschung*, 24(2), 134-147.
- Niklas, F., Segerer, R., Schmiedeler, S. & Schneider, W. (2012). Findet sich ein „Matthäus-Effekt“ in der Kompetenzentwicklung von jungen Kindern mit oder ohne Migrationshintergrund? *Frühe Bildung*, 1(1), 26-33.
- Nores, M. & Barnett, W. S. (2010). Benefits of early childhood interventions across the world: (Under) Investing in the very young. *Economics of Education Review*, 29(2), 271-282.
- Oberhuemer, P. (2012). Fort- und Weiterbildung frühpädagogischer Fachkräfte im europäischen Vergleich. München: Deutsches Jugendinstitut.
- OECD. (2001). *Starting strong: Policy implications for early childhood education and care in the U.S.* Paris: OECD Publishing.
- OECD. (2006). *Starting strong II: early childhood education and care.* (No. 92-64-03545-1 978-92-64-03545-4). Paris: OECD Publishing.
- OECD. (2011). *Starting strong III: A quality toolbox for early childhood education and care.* Paris: OECD Publishing.
- OECD. (2013). *How do early childhood education and care (ecec) policies, systems and quality vary across OECD countries?* Paris: OECD Publishing.
- OECD. (2015a). *Education at a glance 2015.* Paris: OECD Publishing.
- OECD. (2015b). *Starting strong IV: Monitoring quality in early childhood education and care.* Paris: OECD Publishing.
- Oppermann, E., Anders, Y. & Hachfeld, A. (2016). The influence of preschool teachers' content knowledge and mathematical ability beliefs on their sensitivity to mathematics in children's play. *Teaching and Teacher Education*, 58, 174-184.
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of educational research*, 62(3), 307-332.
- Pajares, M. F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (S. 1-49). New York: JAI Press.

- Pakarinen, E., Lerkkanen, M. K., Poikkeus, A. M., Siekkinen, M. & Nurmi, J. E. (2011). Kindergarten teachers adjust their teaching practices in accordance with children's academic pre-skills. *Educational Psychology, 31*(1), 37-53.
- Palmer, A. (2009). 'I'm not a "maths-person"!': Reconstituting mathematical subjectivities in aesthetic teaching practices. *Gender and Education, 21*(4), 387-404.
- Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (2013). *IQB-Ländervergleich 2012 Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. Münster: Waxmann.
- Pauen, S. & Herber, V. (Hrsg.). (2009). *Vom Kleinsein zum Einstein - Offensive Bildung*. Düsseldorf: Cornelsen Verlag.
- Penn, H. (2011). *Quality in early childhood education services - An international perspective*. Berkshire: McGraw-Hill Education.
- Peter-Koop, A. (2009). Orientierungspläne Mathematik für den Elementarbereich – ein Überblick. In A. Heinze & M. Grübing (Hrsg.), *Mathematiklernen vom Kindergarten bis zum Studium* (S. 47-51). Münster: Waxmann.
- Phillipsen, L. C., Burchinal, M. R., Howes, C. & Cryer, D. (1997). The prediction of process quality from structural features of child care. *Early childhood research quarterly, 12*(3), 281-303.
- Pianta, R. C. & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*(2), 109-119.
- Pianta, R. C., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D. et al. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science, 9*(3), 144–159.
- Pianta, R. C., La Paro, K. M. & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS) - Manual Pre-K*. Baltimore: Brookes.
- Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K. & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction, 31*, 2-12.
- Presser, A. L., Clements, M., Ginsburg, H. P. & Ertle, B. (2015). Big Math for Little Kids: The effectiveness of a preschool and kindergarten mathematics curriculum. *Early Education and Development, 26*(3), 399-426.

- Quaiser-Pohl, C. (2008). Förderung mathematischer Vorläuferfähigkeiten im Kindergarten mit dem Programm „Spielend Mathe“. In F. Hellmich & H. Köster (Hrsg.), *Vorschulische Bildungsprozesse in Mathematik und Naturwissenschaften*. (S. 103-125). Bad Heilbrunn: Klinkhardt.
- Quinn, R. J. (1997). Effects of mathematics methods courses on the mathematical attitudes and content knowledge of preservice teachers. *The Journal of educational research*, 91(2), 108–114.
- Racherbäumer, K., Funke, C., van Ackeren, I. & Clausen, M. (2013). Schuleffektivitätsforschung und die Frage nach guten Schulen in schwierigen Kontexten. In R. Becker & A. Schulze (Hrsg.), *Bildungskontexte* (S. 239-267). Wiesbaden: Springer Fachmedien.
- Rademacher, J., Trautewig, N., Günther, A., Lehmann, W. & Quaiser-Pohl, C. (2005). Wie können mathematische Fähigkeiten im Kindergarten gefördert werden? Ein Förderprogramm und seine Evaluation. *Report Psychologie*, 30(9), 366-376.
- Ramm, G., Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D. et al. (2006). *PISA 2003: Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Rammstedt, B. (2013). *Grundlegende Kompetenzen Erwachsener im internationalen Vergleich: Ergebnisse von PIAAC 2012*. Münster: Waxmann.
- Rathgeb-Schnierer, E. (2013). Kleine Kinder spielen und lernen mit bunten Perlen. In J. Sprenger, A. Wagner & M. Zimmermann (Hrsg.), *Mathematik lernen, darstellen, deuten, verstehen* (S. 37-51). Wiesbaden: Springer Fachmedien.
- Rechsteiner, K., Hauser, B. & Vogt, F. (2012). Förderung der mathematischen Vorläuferfertigkeiten im Kindergarten: Spiel oder Training? In M. Ludwig & M. Kleine (Hrsg.), *Beiträge zum Mathematikunterricht 2012* (S. 677-680). Münster: WTM Verlag für wissenschaftliche Texte und Medien.
- Relich, J. (1996). Gender, self-concept and teachers of mathematics: Effects on attitudes to teaching and learning. *Educational Studies in Mathematics*, 30(2), 179-195.
- Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C. et al. (2014). Educational effectiveness research (EER): a state-of-the-art review. *School Effectiveness and School Improvement*, 25(2), 197-230.
- Richter, D., Lehl, S. & Weinert, S. (2016). Enjoyment of learning and learning effort in primary school: the significance of child individual characteristics and stimulation at home and at preschool. *Early Child Development and Care*, 186(1), 96-116.

- Rosenfeld, D. (2010). Increasing perceived efficacy for teaching mathematics: An exploratory study. *Journal of Mathematics Education at Teachers College*, 1(1), 25-35.
- Roßbach, H.-G., Kluczniok, K. & Kuger, S. (2008). Auswirkungen eines Kindergartenbesuchs auf den kognitiv-leistungsbezogenen Entwicklungsstand von Kindern. *Zeitschrift für Erziehungswissenschaft*, 10, 139-156.
- Roßbach, H.-G., Sechtig, J. & Freund, U. (2010). *Empirische Evaluation des Modellversuchs "Kindergarten der Zukunft in Bayern – KiDZ"*. Bamberg: University of Bamberg Press.
- Roßbach, H.-G. & Weinert, S. (2008). *Kindliche Kompetenzen im Elementarbereich: Förderbarkeit, Bedeutung und Messung*. Berlin: BMBF.
- Roux, S. & Tietze, W. (2007). Effekte und Sicherung von (Bildungs-) Qualität in Kindertageseinrichtungen. *ZSE: Zeitschrift für Soziologie der Erziehung und Sozialisation*, 27(4), 367–384.
- Royar, T. (2007). Mathematik im Kindergarten: Kritische Anmerkungen zu den neuen „Bildungsplänen“ für Kindertageseinrichtungen. *mathematica didactica*, 30(1), 29-48.
- Saçkes, M., Flevares, L. M., Gonya, J. & Trundle, K. C. (2012). Preservice early childhood teachers' sense of efficacy for integrating mathematics and science: Impact of a methods course. *Journal of Early Childhood Teacher Education*, 33(4), 349-364.
- Sälzer, C. (2010). Schuleffektivität als pädagogische Wirksamkeit von Schulen oder: Eingangsvoraussetzungen und Kontexteffekte. In C. Sälzer (Hrsg.), *Schule und Absentismus* (S. 43-61). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Sammons, P., Hall, J., Sylva, K., Melhuish, E., Siraj-Blatchford, I. & Taggart, B. (2013). Protecting the development of 5–11-year-olds from the impacts of early disadvantage: the role of primary school academic effectiveness. *School Effectiveness and School Improvement*, 24(2), 251-268.
- Sammons, P., Siraj-Blatchford, I., Sylva, K., Melhuish, E., Taggart, B. & Elliot, K. (2005). Investigating the effects of preschool provision: Using mixed methods in the EPPE Research. *International Journal of Social Research Methodology*, 8(3), 207-224.
- Sammons, P., Sylva, K., Melhuish, E., Siraj, I., Taggart, B., Toth, K. et al. (2014). *Influences on students' GCSE attainment and progress at age 16: Effective Pre-School, Primary & Secondary Education (EPPSE) Project research report*. London: Department for Education.

- Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B. & Elliot, K. (2002). *Measuring the impact of pre-school on children's cognitive progress over the pre-school period*. London: Department for Education.
- Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B. & Elliot, K. (2003). *Measuring the impact of pre-school on children's social/behavioural development over the pre-school period*. London: Department for Education.
- Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B., Elliot, K. et al. (2004). *The Effective Provision of Pre-School Education (EPPE) project: Technical paper 9 - Report on age 6 assessment*. London: Institute of Education, University of London.
- Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B., Grabbe, Y. et al. (2007). *Effective pre-school and primary education 3-11 Project (EPPE 3-11): Summary report: Influences on children's attainment and progress in key stage 2*. London: Department for Education.
- Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B., Toth, K. et al. (2011). *Influences on students' attainment and progress in key stage 3: Academic outcomes in english, maths and science in year 9*. London: Department for Education.
- Saracho, O. N. & Spodek, B. (2007). Early childhood teachers' preparation and the quality of program outcomes. *Early Child Development and Care*, 177(1), 71-91.
- Sarama, J. & Clements, D. H. (2004). Building Blocks for early childhood mathematics. *Early Childhood Research Quarterly*, 19(1), 181-189.
- Sarama, J. & Clements, D. H. (2009). *Learning and teaching early math: The learning trajectories approach*. New York, NY: Routledge.
- Sarama, J., Clements, D. H., Starkey, P., Klein, A. & Wakeley, A. (2008). Scaling up the implementation of a pre-kindergarten mathematics curriculum: Teaching for understanding with trajectories and technologies. *Journal of Research on Educational Effectiveness*, 1(2), 89-119.
- Saxe, G. B., Guberman, S. R., Gearhart, M., Gelman, R., Massey, C. M. & Rogoff, B. (1987). Social processes in early number development. *Monographs of the Society for Research in Child Development*, 52(2), i-162.
- Scarr, S. & Eisenberg, M. (1993). Child care research: Issues, perspectives, and results. *Annual Review of Psychology*, 44, 613-644.
- Scarr, S., Eisenberg, M. & Deater-Deckard, K. (1994). Measurement of quality in child care centers. *Early Childhood Research Quarterly*, 9(2), 131-151.

- Scheerens, J. (1992). Effective schooling: Research, theory and practice. *School Effectiveness and School Improvement*, 4(3), 230-235.
- Scheerens, J. & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Scheerens, J. & Creemers, B. (1989). Conceptualizing school effectiveness. *International Journal of Educational Research*, 13(7), 691-706.
- Schneider, W., Küspert, P. & Krajewski, K. (2013). *Die Entwicklung mathematischer Kompetenzen*. Paderborn: Schöningh.
- Schober, P. S. & Spieß, K. C. (2012). Frühe Förderung und Betreuung von Kindern: bedeutende Unterschiede bei der Inanspruchnahme besonders in den ersten Lebensjahren. *DIW Wochenbericht*, 79(43), 17-28.
- Schuler, S. (2009). Was können Spiele zur frühen mathematischen Bildung beitragen? Chancen, Bedingungen und Grenzen. In M. Neubrand (Hrsg.), *Beiträge zum Mathematikunterricht 2009* (S. 399-402). Münster: WTM-Verlag
- Schuler, S. (2013). Spielend Mathematik lernen? In J. Sprenger, A. Wagner & M. Zimmermann (Hrsg.), *Mathematik lernen, darstellen, deuten, verstehen* (S. 69-76). Wiesbaden: Springer Fachmedien Wiesbaden.
- Schuler, S., Pelzer, M., Wittkowski, A. & Wittmann, G. (2015). Zwischen Interessen des Kindes und Schulvorbereitung: Überzeugungen von ErzieherInnen zu mathematischer Bildung im Kindergarten und im Übergang zur Grundschule. *Frühe Bildung*, 4(4), 196-202.
- Schuler, S., Streit, C. & Wittmann, G. (Hrsg.). (2017). *Perspektiven mathematischer Bildung im Übergang vom Kindergarten zur Grundschule*. Wiesbaden: Springer Fachmedien.
- Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R. & Nores, M. (2005). *Lifetime effects: the High/Scope Perry Preschool study through age 40*. Ypsilanti, MI: High/Scope Press
- Schweinhart, L. J. & Weikart, D. P. (1997). The High/Scope preschool curriculum comparison study through age 23. *Early childhood research quarterly*, 12(2), 117-143.
- Senatorin für Soziales, Kinder, Jugend & Frauen der Freien Hansestadt Bremen. (2012). *Rahmenplan für Bildung und Erziehung im Elementarbereich*. Bremen: Ministerium für Soziales, Kinder, Jugend und Frauen.
- Sherman, H. J. & Christian, M. (1999). Mathematics attitudes and global self-concept: An investigation of the relationship. *College Student Journal*, 33(1), 95-101.

- Silinskas, G., Pakarinen, E., Poikkeus, A. M., Lerkkanen, M. K. & Nurmi, J. E. (2015, August). *Longitudinal effects of kindergarten quality on the development of literacy skills in Grade 1*. Paper presented at the 16th Biennial EARLI Conference for Research on Learning and Instruction, Limassol, Cyprus.
- Siraj, I., Kingston, D. & Melhuish, E. (2015). *Assessing quality in early childhood education and care: Sustained Shared Thinking and Emotional Well-Being SSTEWS scale for 2-5 year-olds provision*. London: Institute of Education Press.
- Siraj-Blatchford, I. (2009). Conceptualising progression in the pedagogy of play and sustained shared thinking in early childhood education: A Vygotskian perspective. *Education and Child Psychology, 26*(2), 77-89.
- Siraj-Blatchford, I., Sammons, P., Taggart, B., Sylva, K. & Melhuish, E. (2006). Educational research and evidence-based policy: The mixed-method approach of the EPPE project. *Evaluation & Research in Education, 19*(2), 63-82.
- Siraj-Blatchford, I., Sylva, K., Muttock, S., Gilden, R. & Bell, D. (2002). *Researching effective pedagogy in the early years*. London: Department of Educational Studies, University of Oxford.
- Slot, P. L. (2014). *Early childhood education and care in the Netherlands: Quality, curriculum, and relations with child development*. Unpublished dissertation, University, Utrecht.
- Slot, P. L., Lerkkanen, M.-K. & Leseman, P. (2015). *The relations between structural quality and process quality in European early childhood education and care provisions: Secondary analyses of large scale studies in five countries*. Retrieved from CARE website: http://ecec-care.org/fileadmin/careproject/Publications/reports/CARE_WP2_D2__2_Secondary_data_analyses.pdf
- Slot, P. L., Leseman, P. P. M., Verhagen, J. & Mulder, H. (2015). Associations between structural quality aspects and process quality in Dutch early childhood education and care settings. *Early Childhood Research Quarterly, 33*, 64–76.
- Slot, P. L., Mulder, H., Verhagen, J. & Leseman, P. P. M. (2014). *Domain-general and domain-specific quality characteristics of early childhood education and care predict growth of two-year-olds' vocabulary and attention skills over one year*. Manuscript submitted for publication.
- Smith, M., Dickinson, D., Sangeorge, A. & Anastasopoulos, L. (2002). *Early literacy and language classroom observation scale (ELLCO)*. Baltimore, MD: Paul Brookes.

- Snider, M. H. & Fu, V. R. (1990). The effects of specialized education and job experience on early childhood teachers' knowledge of developmentally appropriate practice. *Early Childhood Research Quarterly*, 5(1), 69–78.
- Sommerlatte, A., Lux, M., Meiering, G. & Führlich, S. (2009). *Lerndokumentation Mathematik - Anregungsmaterialien*. Berlin: Senatsverwaltung für Bildung, Wissenschaft und Forschung.
- Sparr, M., Chen, J. & McCray, J. (2011). Unpacking teacher confidence in early mathematics instruction. Poster session presented at the biannual meeting of the Society for Research in Child Development, Montreal, Canada.
- Stamm, M. & Edelmann, D. (Hrsg.). (2012). *Handbuch frühkindliche Bildungsforschung*. Wiesbaden: Springer Fachmedien.
- Statistisches Bundesamt. (2013). *Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01.03.2013*. (Statistiken der Kinder- und Jugendhilfe No. 5225402137004). Wiesbaden: Statistisches Bundesamt.
- Steinweg, A. S. (2009). *Lerndokumentation Mathematik*. Berlin: Senatsverwaltung für Bildung, Wissenschaft und Forschung.
- Susperreguy Jorquera, M. I. (2013). *"Math talk" in families of preschool-aged children: Frequency and relations to children's early math skills across time*. Unpublished dissertation, University of Michigan, Ann Arbor.
- Swars, S. L. (2005). Examining perceptions of mathematics teaching effectiveness among elementary preservice teachers with differing levels of mathematics teacher efficacy. *Journal of Instructional Psychology*, 32(2), 139-147.
- Sweeting, K. (2011). *Early years teachers' attitudes towards mathematics*. Unpublished master's thesis, Queensland University of Technology, Brisbane, Australia.
- Sylva, K., Ereky-Stevens, K. & Ariescu, A. M. (2015). *Overview of european ECEC curricula and curriculum template*. Retrieved from CARE website: http://ecec-care.org/fileadmin/careproject/Publications/reports/CARE_WP2_D2_1_European_ECEC_Curricula_and_Curriculum_Template.pdf
- Sylva, K., Melhuish, E., Sammons, P., Siraj, I., Taggart, B., Smees, R. et al. (2014). *Effective Pre-school, Primary and Secondary Education (EPPSE 3-16) Project: Students' educational and developmental outcomes at age 16*. London: Department for Education.

- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. & Taggart, B. (2010). *Early childhood matters: Evidence from the effective pre-school and primary education project*. Oxford/New York: Routledge.
- Sylva, K., Sammons, P., Chan, L. L. S., Melhuish, E., Siraj-Blatchford, I. & Taggart, B. (2013). The effects of early experiences at home and pre-school on gains in English and mathematics in primary school: a multilevel study in England. *Zeitschrift für Erziehungswissenschaft*, 16(2), 277-301.
- Sylva, K., Siraj-Blatchford, I., Taggart, B., Sammons, P., Melhuish, E., Elliot, K. et al. (2006). Capturing quality in early childhood through environmental rating scales. *Early Childhood Research Quarterly*, 21(1), 76-92.
- Sylva, K., Siraj-Blatchford, I. & Taggart, B. (2003). *Assessing quality in the early years - Early Childhood Environment Rating Scale Extension (ECERS-E)*. Stoke on Trent/Sterling: Trentham Books.
- Thiel, O. (2009). *Prozessqualität mathematischer Bildung im Kindergarten*. Berlin: Gesellschaft für Didaktik der Mathematik.
- Thiel, O. (2010). Teachers' attitudes towards mathematics in early childhood education. *European Early Childhood Education Research Journal*, 18(1), 105-115.
- Tietze, W. (2008). Qualitätssicherung im Elementarbereich. In E. Klieme & R. Tippelt (Hrsg.), *Qualitätssicherung im Bildungswesen* (S. 16-35). Weinheim/Basel: Beltz.
- Tietze, W., Becker-Stoll, F., Bensel, J., Eckhardt, A. G., Haug-Schnabel, G., Kalicki, B. et al. (2012). *NUBBEK. Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit. Fragestellungen und Ergebnisse im Überblick*. Weimar/Berlin: verlag das netz.
- Tietze, W., Hundertmark-Mayser, J. & Roßbach, H. (1999). *European Child Care and Education (ECCE)-Study Group: School-age assessment of child development: Long-term impact of pre-school experiences on school success, and family-school relationships*. Brussel, Belgium: European Union.
- Tietze, W., Meischner, T., Gänsfuß, R., Grenner, K., Schuster, K., Völkel, P. et al. (1998). *Wie gut sind unsere Kindergärten? Eine Untersuchung zur pädagogischen Qualität in deutschen Kindergärten*. Weinheim: Beltz.
- Tietze, W., Roßbach, H.-G. & Grenner, K. (2005). *Kinder von 4 bis 8 Jahren - Zur Qualität der Erziehung und Bildung in Kindergarten, Grundschule und Familie*. Weinheim: Beltz.

- Tietze, W., Schlecht, D. & Wellner, B. (2005). *Kindergarten-Skala Erweiterung (KES-R-E). Forschungsversion*. Berlin: Freie Universität.
- Tietze, W., Schuster, K.-M., Grenner, K. & Roßbach, H.-G. (2007). *Kindergarten-Skala (KES-R). Revidierte Fassung. Feststellung und Unterstützung pädagogischer Qualität in Kindergärten*. Berlin: Cornelsen Verlag Scriptor.
- Tirosh, D., Tsamir, P., Levenson, E., Tabach, M., Barkai, R., Roesken, B. et al. (2011, October). *Prospective and practicing preschool teachers' mathematics knowledge and self-efficacy: Identifying two and three dimensional figures*, Paper presented at the 17th annual international conference on Mathematical Views (MAVI), Bochum.
- Tobias, S. (1987). *Overcoming math anxiety*. New York, NY: W. W. Norton.
- Tooke, D. J. & Lindstrom, L. C. (1998). Effectiveness of a mathematics methods course in reducing math anxiety of preservice elementary teachers. *School Science and Mathematics*, 98(3), 136–139.
- Tsakiri, T. (2012). *A study of early childhood education and care and child development in Greece*. Unpublished dissertation, University, London.
- Tsamir, P., Tirosh, D., Levenson, E., Tabach, M. & Barkai, R. (2014). Employing the CAMTE framework: Focusing on preschool teachers' knowledge and self-efficacy related to students' conceptions. In U. Kortenkamp, B. Brandt, C. Benz, G. Krummheuer, S. Ladel & R. Vogel (Eds.), *Early mathematics learning* (pp. 291-306). New York, NY: Springer.
- Ulferts, H. & Anders, Y. (2015). *Effects of ECEC on academic outcomes in literacy and mathematics: Meta-analysis of European longitudinal studies*. Retrieved from CARE website: http://ecec-care.org/fileadmin/careproject/Publications/reports/CARE_WP4_D4_2_Metaanalysis_public.pdf
- Upadyaya, K., Viljaranta, J., Lerkkanen, M.-K., Poikkeus, A.-M. & Nurmi, J.-E. (2011). Cross-lagged relations between kindergarten teachers' causal attributions, and children's interest value and performance in mathematics. *Social Psychology of Education*, 15(2), 181-206.
- van Oers, B. (2002). Teachers' epistemology and the monitoring of mathematical thinking in early years classrooms. *European Early Childhood Education Research Journal*, 10(2), 19-30.
- van Oers, B. (2010). Emergent mathematical thinking in the context of play. *Educational Studies in Mathematics*, 74(1), 23-37.

- Vandell, D. L., Belsky, J., Burchinal, M., Vandergrift, N. & Steinberg, L. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth development. *Child Development, 81*(3), 737-756.
- Vartuli, S. (1999). How early childhood teacher beliefs vary across grade level. *Early Childhood Research Quarterly, 14*(4), 489–514.
- Vermeer, H. J., van IJzendoorn, M. H., Cárcamo, R. A. & Harrison, L. J. (2016). Quality of child care using the environment rating scales: A meta-analysis of international studies. *International Journal of Early Childhood, 48*(1), 33-60.
- Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics, 261*- 293.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *J Stat Softw, 36*(3), 1–48.
- Viernickel, S. & Schwarz, S. (2009). *Schlüssel zu guter Bildung, Erziehung und Betreuung – Wissenschaftliche Parameter zur Bestimmung der pädagogischen Fachkraft-Kind-Relation*. Berlin: GEW.
- von Maurice, J., Artelt, C., Blossfeld, H. P., Faust, G., Roßbach, H.-G. & Weinert, S. (2007). *Bildungsprozesse, Kompetenzentwicklung und Formation von Selektionsentscheidungen im Vor-und Grundschulalter: Überblick über die Erhebungen in den Längsschnitten BiKS-3-8 und BiKS-8-12 in den ersten beiden Projektjahren*. Bamberg: Universität.
- Vural, D. E. & Gunhan, B. C. (2011). Development of self-efficacy scale regarding prospective pre-school teacher toward teaching mathematics. *International Journal of Contributions to Mathematics Teaching, 1*(1-2), 35-41.
- Weikart, D. P. & Schweinhart, L. (2013). The High/Scope curriculum for early childhood care and education. In J. L. Roopnarine & J. E. Johnson (Eds.), *Approaches to early childhood education* (6th ed., pp. 277-293). New York, NY: Pearson.
- Weinert, F. E. (2001). *Leistungsmessungen in Schulen*. Weinheim, Basel: Beltz.
- Weinert, S., Ebert, S., Lockl, K. & Kuger, S. (2012). Disparitäten im Wortschatzerwerb: Zum Einfluss des Arbeitsgedächtnisses und der Anregungsqualität in Kindergarten und Familie auf den Erwerb lexikalischen Wissens. *Unterrichtswissenschaft, 40*(1), 4-25.
- Wessels, H., Lamb, M. E. & Hwang, C. P. (1996). Cause and causality in daycare research: An investigation of group differences in Swedish child care. *European Journal of Psychology of Education, 11*(2), 231-245.

- Whitebook, M. (2003). Early education quality: Higher teacher qualifications for better learning environments - A review of the literature. Berkeley, CA: Institute for Industrial Relations.
- Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review*, 6(1), 49–78.
- Wigfield, A. & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68-81.
- Wigfield, A. & Eccles, J. S. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation*. New York, NY: Academic Press.
- Wilkins, J. L. M. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education*, 11(2), 139-164.
- Wilkins, J. L. M. (2009). Elementary school teachers' attitudes toward different subjects. *The Teacher Educator*, 45(1), 23-36.
- Wilkins, J. L. M. & Brand, B. R. (2004). Change in preservice teachers' beliefs: An evaluation of a mathematics methods course. *School Science and Mathematics*, 104(5), 226-232.
- Wittmann, E. C. (2004). Design von Lernumgebungen zur mathematischen Frühförderung. In G. Faust (Hrsg.), *Anschlussfähige Bildungsprozesse im Elementar- und Primarbereich* (S. 49-63). Bad Heilbrunn: Klinkhardt.
- Zacharos, K. & Ravanis, K. (2000). The transformation of natural to geometrical concepts, concerning children 5–7 years old. The case of measuring surfaces. *European Early Childhood Education Research Journal*, 8(2), 63-72.

Anhang A – Manuskript zur Metaanalyse (TP1)

Originalmanuskript: Ulferts, H. & Anders, Y. (2016). *Impact of pedagogical processes in early childhood education and care on academic outcomes: A meta-analysis of European Longitudinal Studies*. Manuscript in preparation.

Abstract

This meta-analysis studies the impact of pedagogical processes in early childhood education and care on outcomes in two academic domains: literacy and mathematics. It synthesizes evidence from 17 European longitudinal studies comprising 16,461 children in regular center-based care spanning the period between ages three and 16. A three-level meta-analysis revealed significant positive overall effects for global process quality ($ES = .11$, Cohen's $d = .22$) and early learning promotion ($ES = .10$, Cohen's $d = .20$). The effects do not decline with age and phase of academic career. Effects vary by outcome domain, type of process measure, and differences in adopted analysis. Results suggest that improving quality of regular provision has persistent developmental impact on academic development.

Keywords: meta-analysis, early childhood education and care, literacy, mathematics, process quality

Theoretical Background

The impact of early childhood education and care (ECEC) has received growing scientific and public interest with the expansion of and increasing investments into ECEC services in many countries around the world (Organisation for Economic Co-operation and Development [OECD], 2013; 2015) and the growing awareness of its potential developmental benefits. By now several reviews and meta-analyses have aggregated evidence of beneficial effects of attending ECEC, especially for outcomes in cognitive domains (e.g., Anders, 2013; Burger, 2010; Camilli, Vargas, Ryan, & Barnett, 2010; Melhuish et al., 2015; Nores & Barnett, 2010), and of economic benefits of investing in ECEC (e.g., Heckman, 2006). These summaries also show that ECEC represents an effective antidote against social inequalities.

At the age of four on average 88% of children in the OECD countries, and as high as 91% in Europe, attend some form of childcare, mainly center-based (OECD, 2015). Previous meta-analyses studied the benefits of attending versus not attending ECEC and the effectiveness of specific programs and interventions (e.g., Blok, Fukkink, Gebhardt, & Leseman, 2005; Camilli, et al., 2010; Karoly, Kilburn, Cannon, 2005; Manning, Homel, & Smith, 2010; Nores &

Barnett, 2010). With the vast majority of children attending regular provision, however, it is important to accumulate knowledge about developmental gains of improving important aspects of regular ECEC in order to effectively channel investments and policy efforts. Reviews summarized the evidence not only on the effectiveness of programs and interventions, but also on the developmental consequences of optimizing the quality and quantity of regular ECEC (Anders, 2013; Burger, 2010; Dalli, White, Rockel, & Duhn, 2011; Melhuish et al., 2015). Corresponding quantitative syntheses could provide a more compressed knowledge basis; however, these are still missing.

Theory and research imply that processual characteristics such as central quality aspects play key roles in explaining differences in developmental impact (Kluczniok & Roßbach, 2014; Pianta & Hamre, 2009).

The current meta-analysis, therefore, aims for a quantitative synthesis of evidence for the effects of two important characteristics of pedagogical processes: global process quality and the extent of stimulation and promotion in various domains. The meta-analysis focuses on outcomes in two important academic domains, i.e., literacy and mathematics, which are officially part of educational goals or areas of experience of center-based ECEC in many countries (OECD, 2006). Main objective of this meta-analysis is to find out if improvements in pedagogical processes of the major part of ECEC system, i.e. center-based regular care, could lead to substantial academic gain, hence, representing effective means of providing majority of young children with a good start for their further academic career.

Limitations of Previous Meta-Analyses on the Impact of ECEC

Several attempts to synthesize evidence on the developmental impact of ECEC already exist and they evidence in overall beneficial effects of ECEC on child outcomes (e.g., Blok et al., 2005; Camilli, et al., 2010; Nores & Barnett, 2010). Generally, the effects tend to be stronger and more homogenous for outcomes in cognitive domains, including literacy and math outcomes, than for those in other domains. Several limitations of these syntheses call for a revision of the aggregation approach instead of a simple integration of new study results. Firstly, previous meta-analyses primarily study the effect of ECEC attendance per se or of participation in specific ECEC program or intervention. Overall results are normally based on evidence from group contrast designs. Included studies compare children attending to those not attending ECEC, or children participating in different programs and interventions. Some standardized measure of difference in means between the samples, such as Cohen's *d* or Hedges *g*, usually serves as an effect size in these syntheses (Card, 2012; Cooper, 2009). While relying on group contrast designs, most meta-analyses discard important evidence on regular provisi-

on, as those studies do not compare children of different programs or interventions and many recent studies do not sample children not attending ECEC. Due to high attendance rates, this disproportionally impacts European evidence. Hence, even when aiming at an international scope, the existing meta-analyses are mainly based on US studies (e.g., Blok et al., 2005; Norens & Barnett, 2010).

Furthermore, to a great extent, aggregated effects encompass information about the effectiveness of specific programs or interventions, which often provided additional support for families, such as parent-trainings or home-visits. The generalizability of findings to regular provision is questionable. First of all, the informative value of the effectiveness of specific programs and interventions for the effects of regular provision is rather limited and the size and nature of pure ECEC effects remain unclear. If the final goal is to optimize the ECEC system in general, rather than to supplement it with specific programs and interventions, meta-analysis should extract the knowledge about the relative impact of naturally occurring variations of central characteristics of regular ECEC.

In addition, previous meta-analyses report that overall effects vary substantially. One possible explanation is that they subsume effects for outcomes of very heterogeneous nature into one overall effect for the cognitive domain. Overall effects often span outcomes from intelligence and basic cognitive skills to academic outcomes including skills. The stimulation and promotion of early literacy and mathematics form educational core tasks of center-based ECEC in most ECEC curricula or guidelines currently in place (OECD, 2006), and, therefore, represent reasonable foci of meta-analyses on ECEC effects (Burger, 2010).

The Role of Pedagogical Processes for the Academic Impact of ECEC

Studies commonly assess several indicators of quality and quantity to further characterize the ECEC experience of children. After a vast quantitative expansion of ECEC service in most countries, the improvement and impact of quality is currently the central point of scientific and public attention, and, considering the low to medium overall quality of service in most countries, a major challenge for ECEC policy (OECD, 2011; 2013). Despite common goals and characteristics, ECEC systems around the world differ (Eurydice, 2009; 2014; OECD, 2006; 2011), but research syntheses can accumulate evidence on the developmental impact of ECEC quality across countries, as the quality framework and measures are applicable in different countries with differing ECEC systems and programs (Pianta et al., 2005; Kluczniok & Roßbach, 2014). Reviews on the impact of ECEC have shown that quality characteristics partly

explain heterogeneous findings and help understand how ECEC affects child development (e.g., Anders, 2013; Burger, 2010; Dalli et al., 2011; Melhuish et al., 2015).

Though slightly different conceptualizations and operationalization exist, researchers and experts share a common understanding of the concept of quality and appropriate ways of its assessments (Pianta et al., 20015; Kluczniok & Roßbach, 2014). A common conceptualization is structural-processual, covering aspects of structural quality such as teacher-child ratio, group size, and structural arrangements, as well as aspects of process quality. Process quality captures the quality of pedagogical processes and daily interactions in ECEC. Within quality frameworks, processual characteristics represent key factors to explain differences in developmental impact of ECEC on child development, as theory and research imply their direct link to child outcomes; other quality characteristics are considered to influence development indirectly through their influence on pedagogical processes (Kluczniok & Roßbach, 2014; Pianta et al., 2005). Because of their key role in explaining the developmental impact of ECEC, pedagogical processes represent promising foci for a meta-analysis.

Commonly, global aspects of process quality are separated from domain-specific aspects (Kuger & Kluczniok, 2008; Hamre, Hatfield, Pianta, & Jamil, 2014). Global process quality measures general aspects of the quality of interactions between caregiver and child, among children, with parents, as well as with material and spatial surroundings. Examples of such aspects are the socioemotional climate, the warmth of interactions, and the responsiveness of caregivers. Domain-specific aspects of process quality refer to the extent and quality of promotion and stimulation of early learning in various domains, including literacy and mathematics. Studies and reviews collected evidence for the key role of pedagogical processes: in general children's development in literacy and mathematics seem to benefit from a higher global process quality and a higher extent of early learning promotion (Anders, 2013; Dalli et al., 2011; Melhuish et al., 2015). The British study Effective Pre-school, Primary Education and Secondary Education (EPPSE) found academic benefits of higher process quality of ECEC even at 16 years of age (Melhuish et al, 2015).

Commonalities and Differences of ECEC Systems and Research in Europe and the US

ECEC is widely acknowledged as the first step into the phase of institutional learning, where learning opportunities for pre-academic skills usually target the three to six year olds, as most children enter ECEC at around three to four years of age (OECD, 2015), with a recent trend of earlier enrollments. Besides, due to the predominance of systems split between “early education” and “childcare” services (European Commission/EACEA/Eurydice/Eurostat, 2014; OECD, 2011), pedagogical work and research on early learning promotion and stimulation in

ECEC centers, as well as most literacy and math programs (Baarody & Lai, 2006; Burger, 2014; Neumann & Dickinson, 2011), have been mainly restricted to older children. Only as of very recently new efforts are being made in ECEC research and practice to find ways to support younger children in their pre-academic learning (Anders, 2013; Anders et al., 2015, Dalli et al., 2011; OECD, 2011).

As in the US, combating social and economic disadvantage is a major goal of educational efforts in Europe, and ECEC's potential to foster early learning is similarly seen as an effective means to achieve this goal (European Commission/EACEA/Eurydice/Eurostat, 2014). Instead of supplementing the systems with large-scale targeting programs for disadvantaged children, such as Head Start in the US (Heckman, 2006; Barnett, 2011), the leading approach in Europe is the enhancement of the developmental impact of regular provision, so that all children, including the disadvantaged, can benefit (Eurydice, 2009). Most governments aim for inclusive ECEC systems with universal access to high quality education for children from all backgrounds, which can compensate for a potential lack of stimulation and promotion at home.

Expanding services and granting subsidies for childcare, often to parents who could not afford it otherwise, represents a common policy measure (OECD, 2006; 2013). ECEC policies additionally strive to assure and increase the quality of ECEC services in general, as well as the extent and quality of promotion and stimulation of early learning in particular, because it is argued that both, but especially the latter, still widely remain on low to mediocre levels (Kuger & Kluczniok, 2008; Slot et al., 2015). Governments have increasingly introduced assessments of quality and child outcomes in ECEC to guide reforms, legislations and initiatives by installing regular inspection or monitoring systems or as part of research projects.

Compared to the US Europe has a rather short history of ECEC research, in particular large-scale longitudinal research (Anders, 2013). Around two decades ago, all around Europe large-scale projects on the impact of ECEC on child development evolved (see Anders, 2013; Burger, 2010; Dalli et al., 2011; Melhuish et al., 2015). The vast majority of studies tries to capture the developmental impact of naturally occurring variations in regular provision by investigating indicators of quality and quantity instead of evaluating the implementation of a specific program or intervention. These studies follow the developmental trajectories of children attending regular ECEC, additionally assessing further child and family background characteristics, as well as characteristics of the institutional and home environments. By adopting longitudinal and complex designs and analytical approaches studies can disentangle the insti-

tutional impact from other influences on child outcomes (Burger, 2010). A few studies evaluate rather unspecific and broad initiatives aimed at increasing quality and quantity in regular provision.

Like in the US, disadvantage is a major research topic in Europe, but the type and dynamics of disadvantage differ slightly across countries. In accordance with the predominant inclusive approach, large-scale research in Europe usually tries to address the topic by assessing culture-specific or country-specific indicators of disadvantage and studying their role in effects of regular ECEC. Thereby, scientists want to find out if and how disadvantaged children can benefit from ECEC in regular provision as much as their peers or more, in order to close developmental gaps (compensatory effects; Eurydice, 2009; Kluczniok & Roßbach, 2014; Vandell et al., 2010). Common goal is to gather empirical knowledge for an evidence-based improvement of the system of regular provision for all children, including the disadvantaged. However, there are also some European evaluation studies of targeting programs for disadvantaged children, such as the Dutch studies pre-Cool (Slot, Mulder, Verhagen, & Leseman, 2014) and Utrecht Mixed Preschool Groups (de Haan, Elbers, Hoofs, & Leseman, 2013).

The Rationale for Potential Moderators of Effects

Despite a general tendency of academic benefits of enhanced pedagogical processes in regular provision reviews evidence substantial variation of processual effects between studies and also within studies, for example for certain outcomes or measurement points (Anders, 2013, Dalli et al., Melhuish et al., 2015). There are several important differences between existing evidence, which can potentially explain variations and, hence, should be examined as potential moderators of processual effects, starting with the time of outcome measurement.

Studying the time of outcome measurement as a moderator provides insights into the *persistence of processual effects*. Previous meta-analyses found that effects of ECEC program and intervention on cognitive outcomes fade as children grow older (e.g., Blok et al., 2005; Camilli et al., 2010; Nores & Barnett, 2010). So it is likely that the associations to characteristics such as pedagogical processes decline gradually with age and increased distance to the ECEC phase. Another plausible explanation for a decline of effects is the exposure to new institutional influences, for example when children enter primary schools which bare certain qualities (Anders, Grosse, Roßbach, Ebert, & Weinert, 2013; Lehl, Kluczniok, & Roßbach, 2016; Sylva et al., 2010). Separating effects according to the phase of outcome measurement represents a plausible alternative decrease pattern of ECEC effects (Camilli et al., 2010).

Another important difference is whether reported findings refer to *attainment at a specific time point or growth in outcome over a certain period*. Ultimate goal of early pre-academic education is to boost children's growth. Research has indicated that children within ECEC already differ in literacy and math skills and probably have entered ECEC with differing skill levels depending on their family background (Anders, 2013; Anders et al., 2012; Arnold & Doctoroff, 2003; Ebert et al., 2013; Eurydice, 2009; Melhuish et al., 2015). Though aiming at boosting children's growth in outcomes, there are alternative explanations for observed associations between pedagogical processes and children's attainment. For example, it might be easier to achieve high quality of pedagogical processes in groups with a low percentage of disadvantage children or in groups high mean levels of skills (cf. Kuger & Kluczniok, 2008; Slot, Lerkkanen, & Leseman, 2015). Associations can also reflect selection effects, i.e. selective placement of children sharing certain background characteristics to ECEC settings of particular quality. For some ECEC systems selection effects reflect targeting policy, e.g. similarly to Head Start in the US, only Dutch children with a low socioeconomic background are entitled to the education-oriented VVE-programs (Slot et al., 2014). In other systems selective placement according to family background naturally occurs as a result of social segregation (cf. Eurydice, 2009; OECD, 2006). A longitudinal design of studies enables exploring influences on growth by entering previous attainments as predictors (e.g., Sylva et al., 2010, Anders et al., 2013) or by using growth curve modeling (e.g., Anders et al., 2012, 2013; de Haahn et al., 2013), which is, however, only possible for later waves of outcome assessments.

Previous meta-analyses on the impact of ECEC subsumed findings for outcomes of very heterogeneous nature under one effect for cognitive outcomes. None of them, so far, studied *the domain-specificity of effects* by considering domain as a moderator. Research and ECEC policy, however, stress the importance of domain-specific learning processes: children need to acquire specific skills and knowledge to perform well in mathematics and literacy (Burger, 2014; Harme, Hatfield, Pianta, & Jamil, 2014; Kluczniok & Roßbach, 2014; Kuger & Kluczniok, 2008; Sylva et al., 2015; OECD, 2006; 2011). For example, research shows that children's math development benefits particularly from early numeracy skills, such as number recognition, counting or basic arithmetic skills (Baroody, Lai, & Mix, 2006; Burger, 2014). Basic language skills such as phonological awareness or knowledge about the alphabet form early life foundations for later literacy development (Burger, 2014; Neuman & Dickinson, 2011). ECEC curricula specify learning goals and experiences separately for domains (OECD, 2006; 2011). Research on pedagogical processes in ECEC considered the extent and quality of promotion and stimulation by creating specific scales or subscales for different domains (Hal-

le, Vick, & Anderson, 2010; Kluczniok & Roßbach, 2014; Sylva, Siraj-Blatchford, & Taggart, 2003).

It is also likely that the *type of measure used to assess pedagogical processes* is related to the observed effects. Most frequently observational measures serve to capture pedagogical processes within ECEC groups or classes. Adopted scales overlap in their understanding of beneficial aspects of pedagogical processes, but slightly differ in considered facets. Some measures of global process quality include *ratings of material and spatial surroundings* in their overall rating (e.g., the Early Childhood Environment Rating Scale-Revised Edition [ECERS-R]; Harms et al. 1998), whereas others *focus on the quality of interaction* (e.g., the Caregiver Interaction Scale [CIS], Arnett, 1989; the Classroom Assessment Scoring System [CLASS], Pianta et al., 2008). The question if certain materials and structural arrangements are necessary to achieve developmentally beneficial interactions remains unanswered. *Observational measures* are also the standard for assessing the extent of early learning promotion (Halle et al., 2010; Sylva et al., 2006), e.g., rating scales like the Extension of the Early Childhood Environment Rating Scale (ECERS-E; Sylva et al., 2003; see also Halle et al., 2010). Large-scale assessments in research and practice, however, have developed *staff questionnaires* to capture processes as a more economical alternative to relatively time- and cost-intensive observations (e.g., Durda, 2015; Silinskas et al., 2015; Slot et al., 2014). Yet, it is not clear if they equally well reflect the quality of interaction beneficial for child development.

Studies on regular provision generally do not meet the criteria of random-assignment and controlled experimental studies (cf, Higgins & Green, 2008), because they study naturally occurring variations in ECEC systems. Some studies adopt quasi-experimental designs, contrasting existing „intervention groups“, but most studies rely on quasi-experimental methods, and use multivariate analysis to estimate ECEC effects and control for further factors relevant for child development (Burger, 2010; Nores & Barnett, 2010). *Adopted analytical approach* is likely to be related to obtained results (Becker & Wu, 2007; Bowman, 2012; Burger, 2010; Cooper, 2009). Hence, it is important to study its central characteristics such as the *number and type of considered predictors* as well as further details as moderators of effects.

Family characteristics such as low socioeconomic status, income, or migration background, besides *child characteristics*, are important predictors of child development, which can partly be explained through their influence on the quality of the *HLE or home learning environment* (Anders, 2013; Arnold & Doctoroff, 2003; Lehl et al., 2016; Sylva et al., 2010). Apart from child and family characteristics, scientific literature discusses the role of *structural quality*, as well as *quantity of ECEC*, e.g. hours per week or duration in years (e.g., Anders,

2013; Burger, 2010; Melhuish et al., 2015). Considering further relevant factors in the analysis is necessary to entwine ECEC effects from other influences and from one another, and to get a better estimate of the unique effect of specific ECEC characteristics, e.g. pedagogical processes.

The complex designs of studies also allow for diverse analytical approaches, ranging from correlations and multiple regressions to growth curve analysis and structural equation modeling, which differ in important details. ECEC research increasingly adopts *multilevel approaches* to account for the hierarchical or clustered structure of the data, i.e. children nested in groups nested in centers (Anders et al., 2012; Sylva et al., 2010). Not adjusting for clustered data can produce misleading results, for example significance tests with inflated alpha levels (Hox, 2002). *Latent approaches* can partly account for the dimensionality of constructs, e.g. measures of child outcomes and pedagogical processes, as well as measurement error, which leads to an increase in the strength of existing relations between constructs (Byrne, 2009).

Research Aims of the Current Meta-Analysis

The aim of this meta-analysis is a quantitative synthesis of longitudinal European evidence on the impact of ECEC on academic outcomes in literacy and mathematics. Instead of studying the effectiveness of specific programs or interventions, the meta-analysis explores the impact of variations in regular center-based care of two key quality aspects: global process quality and the extent of early learning promotion. Based on previous reviews (e.g., Anders, 2013, Dalli et al., Melhuish et al., 2015), we anticipated positive effects for both processual characteristics. As considered longitudinal studies report multiple effect sizes for similar effects, e.g. for different time-points or outcome measures, it is likely that effects vary between and within studies. Additionally, we expected characteristics of outcome and process measures (i.e., age and educational phase of outcome assessment, growth versus attainment of outcome, outcome domain, and type of process measure), as well as adopted analytical approach (i. e., number and type of considered variables, multilevel and latent approach) to explain existing variations in effects.

Method

Search and Selection

To identify studies and relevant references we conducted a systematic search procedure. We explored various electronic databases, such as EBSCOhost, Web of Science, PsycINFO, and Pubmed, and search engines such as Google Scholar, using various combination of search

terms related to ECEC (e.g., “preschool”, “ECEC”, “ECCE”, “childcare”, “early childhood education”) along with a term related to the outcome (e.g., “literacy”, “language”, “writing”, “mathematics”, “numeracy”, “academic”). We restricted the search to Europe (e.g., “Europe”, “European”, “UK”, “Finnish”) and, if necessary, to longitudinal designs (e.g., “longitudinal”, “development”, “growth”). We also cross-checked our results with reference lists of related reviews and meta-analyses (e.g., Anders, 2013; Anderson et al., 2003; Blok et al., 2005; Burger, 2010; Dalli et al., 2011; Melhuish et al., 2015; Mitchell, Wylie, & Carr, 2008; Nores & Barnett, 2010) and searched websites of research institutes and ministries in Europe for relevant research projects. We additionally asked European ECEC experts to identify studies or references missing in our reference list at the time, with an emphasis on unpublished and untranslated work. All in all, we identified 207 potentially relevant references.

To be included studies had to meet several criteria. Studies had to study primarily center-based childcare in Europe with a longitudinal design and a beginning after 1990. The report of effect sizes for the relation of at least one measure reflecting pedagogical processes, either global process quality or the extent of early learning promotion, and one relevant outcome measure (e.g., arithmetic or numeracy skills, language, (pre-) reading or writing skills) was mandatory for selection. Studies with risk groups as samples, for example hyperactive children or children with special needs, were excluded. We also excluded studies evaluating specific programs or interventions, which did not reflect the experience of children in regular provision. We did, however, include studies evaluating rather broad and unspecific initiatives with the goal of enhancing quality or quantity of regular provision. We excluded references which only provided effect sizes for comparisons between samples experiencing differing levels of process quality and a sample not attending childcare (i.e., some references from EPPSE), because effects should reflect the developmental impact of variations in process quality for children attending ECEC. Twenty-seven references, marked with asterisks in the reference list, met the inclusion criteria.

Coding

Coders extracted information on the *ESs* (for effect sizes), potential moderators and further study characteristics. As effect sizes we selected results of the model with the most predictors and gave priority to results referring to growth in outcome as opposed to results not controlling for prior attainments. Generally, we preferred findings for the total scores of measures, though if not reported, we used findings for subscales as approximations (e.g., subscales of ECERS-R). Findings of the same study representing unique information were included as separate *ES*, i.e., differing in studied outcome (e.g., vocabulary and pre-reading

skills as literacy outcomes), process measure (e.g., CIS and ECERS-R for global process quality), subsample (e.g., country-specific subsamples of the European Child Care and Education Study [ECCE]; Tietze, Hundertmark-Mayser, & Roßbach, 1999), or measurement point.

Coders independently extracted information from references according to a heavily structured coding scheme in SPSS. All coders were PhD or master students in psychology or human factors, and possessed deep knowledge of child development, multivariate analysis, and complex and longitudinal research designs. To guarantee transparency and reliability of coding, coders received intensive training, a coding manual, sample codings for six references in the SPSS file, and continuous support throughout the coding process. A sample of studies were coded twice by two independent coders (71% of the data). The average interrater agreement, measured by Cohen's kappa, across the categorical variables was $\kappa = .80$. The average interclass correlation coefficient for continuous variables was $ICC = .93$. The raters discussed disagreements until they reached a consensus; the consensus then became the code. If crucial information was missing the authors of primary studies were contacted. Thereby, the meta-analysis was able to cover unpublished information on study characteristics for several studies and newly computed effect sizes for three studies. All in all, our overall effects are based on 143 effect sizes (see Table 5 of the Supporting Online Material).

Analytical Approach for Building Overall Effects

To build the overall effects we used standardized coefficients for the relation of pedagogical processes to child outcomes as *ES* (Card, 2012; Bowman, 2012; Kim, 2011). Though based on various analyses, ranging from correlations, multiple regressions to latent growth curve modelling, all *ESs* belonged to the correlation family and were standardized so that they ranged from -1 to 1, with an *ES* of 0 indicating no processual effect. In case of unstandardized coefficients and unavailability of necessary information for post-hoc standardization, we asked the authors for (standardized) results or excluded the coefficients from the analysis. When theory suggested reversed relationships for effects, we built the inverse of reported effects post hoc, for example for the CIS subscales detachment and punitiveness (Arnett, 1989).

We conducted all analyses with the statistic program R using longitudinal multilevel meta-analysis (Maas, Hox, & Lensvelt-Mulders, 2004) in the R-package metafor with restricted maximum likelihood (REML) estimators (Viechtbauer, 2010). Longitudinal multilevel meta-analysis takes into account the dependency of *ESs*, where *ESs* for different process and outcome measures (level 1) are nested within measurement points or ages of children (level 2) within studies (level 3). We weighted each *ES* by the inverse of its variance (Card, 2012; see

Figures 1 and 2 of the Supporting Online Material), so that *ESs* of larger studies received more weights.

We built two separate overall effects for global process quality and early learning promotion (see Table 5 of the Supporting Online Material). The overall effect for *global process quality* includes ECERS-R, CLASS or CIS and their subscales, and two of the inspectorate self-evaluation scales of the Social Care and Social Work Improvement Scotland (SCSWIS-scales; Bradshaw et al., 2014). The overall effect of *early learning promotion* includes ECERS-E and subscales, ELLCO, the extended version of the Dortmunder Rating Scale (DORESI-E-Ki, Fried et al., 2012), specific subscales of ECERS-R referring to the promotion of early learning (i.e., activities, language and reasoning and interactions), and a self-constructed observational instruments of teacher-managed literacy and math activities (de Haan et al., 2013). It also entails measures based on staff questionnaires about the frequency of reading activities (Pakarinen, Lerkkanen et al., 2015; Silinskas, Pakarinen, Poikkeus, Lerkkanen, & Nurmi, 2015) or developmentally stimulating activities (Durda, 2015), as well as self-evaluative quality indicators (QI) of Education Scotland (Bradshaw et al., 2014).

Test of Moderators and Evaluating Potential Bias

We studied potential moderators using mixed-effects model based on the multilevel-models described earlier (Viechtbauer, 2010). We studied age as a continuous moderator and grouped *ES* to build categorical moderators for educational phase of academic career at outcome assessment (in ECEC vs. in school). We also compared *ESs* referring to growth vs. attainment in outcome, and studied if effects vary by outcome domain (literacy vs. mathematics), and by type of process measure (interaction-focused vs. additional ratings of material surrounding for global process quality, and observations vs. questionnaires for early learning promotion). As additional moderators regarding the adopted analytical approach we introduced the number (below vs. above the median of five predictors) and type of predictors (i.e., if analyses considered characteristics of the child, family, HLE, or structural quality and quantity of ECEC). As recommended by Cooper (2009), we introduced the moderators separately to avoid a severe loss of statistical power and multicollinearity.

Despite a thorough review of literature, our sample may not include all relevant findings. We applied different approaches to test the possibility and consequences of missing findings (cf., Higgins & Green, 2008). First of all, to test for publication bias, we tested if *ESs* derived from peer-reviewed articles differed in size from *ESs* from non-peer reviewed references. In addition, we visually inspected funnel plots of the overall effects for asymmetries and computed rank correlation tests for funnel plot asymmetries. By using the file drawer analysis (faile-

safe N calculation) with a significance level of $p = .05$, we tested if overall effects were robust against missing null findings.

Results

Descriptive Overview on Evidence Basis

Our meta-analysis is based on evidence of 17 longitudinal studies conducted in nine European countries. Five studies evaluate initiatives aimed at enhancing the quality of ECEC. Results summarize findings for 16,461 children in 2,887 centers and evidence spans a period from age three to age 16, although the majority of findings do not exceed the primary school phase (65%). Studies differ in sample sizes, adopted analysis, and further design characteristics. Mostly studies assess and control several child and family background characteristics. Table 4 in the Supporting Online Material depicts central characteristics of all included evidence. The included studies used different measures of global process quality and the extent of early learning promotion, most commonly the ECERS-R, ECERS-E and the CLASS, for which on average low to mediocre levels were observed, especially for early learning promotion (see Table 3 in the Supporting Online Material).

Overall Effects and Heterogeneity

Aggregation of evidence from 13 studies and 73 ES and 11 studies and 70 ES, respectively, resulted in significant positive effects for global process quality ($ES = .11$, $se = .04$, $p < .01$, $CI_{95\%} = .04 - .18$) and early learning promotion ($ES = .10$, $se = .03$, $p < .001$, $CI_{95\%} = .05 - .15$). Transforming the correlational results into Cohen's d resulted in $d = .22$ and $d = .20$. Our three-level models showed that ES varied within ($\sigma_{age} = .023$ and $.01$) and between studies ($\sigma_{studies} = .005$ and $.002$). Cochrane's Q -tests of homogeneity (Higgins & Green, 2008) indicated that observed variations were beyond chance: $Q_{ES}(69) = 264.55$ and $Q_{ES}(72) = 2482.15$ ($p < .001$). The forest plots available as Figures 1 and 2 of the Supporting Online Material provide a visual display of the heterogeneity of the overall effects.

Moderators of Processual Effects

Table 1 displays the significance of each tested *moderator referring to outcome and process measures* ($df = 1$ for all Q_M) and if substantial variation in effects remained after including the moderator (Q_E and corresponding df_E). The table also shows the effect sizes, corresponding sample sizes (k_{ES} and k_{study}), standard errors (se), and confidence intervals ($CI_{95\%}$) for each sub-effect. Effects did not decline when children left ECEC. Neither did results indicate that effects decline with age when tested as a continuous moderator ($Q_M = 0.33$, $p = .56$ and $Q_M = 0.004$, $p = .95$). For global process quality, only the effect on growth in child outcome was on average significant ($Q_M = 3.93$, $p < .05$; $ES = .14$, $p < .001$ vs. $ES = .05$, $p = .33$). Outcome do-

main emerged as a significant moderator of both effects: $Q_M = 36.04, p < .001$ and $Q_M = 6.32, p < .05$. Though significantly related to outcomes in both domains, the effect of global process quality was stronger for literacy than mathematics ($ES = .13, p < .001$ vs. $ES = .07, p < .05$). By contrast, the effect of early learning promotion was stronger in mathematics than literacy ($ES = .11, p < .01$ vs. $ES = .08, p < .001$).

The type of process measure was a moderator of both effects ($Q_M = 9.35, p < .01$ and $Q_M = 4.8, p < .05$). We observed stronger effects for interaction-focused measures of global process quality than for those measures additionally rating the environmental surrounding ($ES = .19, p < .001$ vs. $ES = .09, p < .05$). Child outcomes were only associated with observational measures of early learning promotion ($ES = .12, p < .001$), whereas no significant overall relation to staff questionnaires emerged ($ES = .01, p = .78$).

Additionally, we tested the *adopted analytical approaches* as moderators of effects ($df = 1$ for all Q_M). Table 3 in the Supporting Online Material displays the results of the mixed-effects models. The number of predictors in the analysis did not moderate processual effects. Whether studies considered further child characteristics or not moderated both effects ($Q_M = 3.6, p < .10$ and $Q_M = 4.64, p < .05$) and only controlled effects were significant ($ES = .12, p < .001$ vs. $ES = .003, p = .96$ and $ES = .13, p < .01$ vs. $ES = .04, p = .48$). Whether studies considered family background or not moderated the effect for early learning promotion ($Q_M = 5.97, p < .05$) and similarly only controlled effects were significant ($ES = .15, p < .01$ vs. $ES = .05, p = .39$). For global process quality, only effects which considered HLE characteristics were significant ($Q_M = 4.37, p < .05$; $ES = .18, p < .001$ vs. $ES = .07, p = .11$). Though even without entering aspects of structural quality a significant effect for early learning promotion emerged, it was marginally higher if structural aspects were included ($Q_M = 3.6, p < .10$; $ES = .14, p < .001$ vs. $ES = .07, p < .01$). Introducing aspects of quantity as additional predictors not moderate the effects. Furthermore, we found effect sizes to be lower if analyses considered the multilevel structure of the data for global process quality ($Q_M = 2.72, p < .10$; $ES = .08, p < .05$ vs. $ES = .17, p < .01$) and early learning promotion ($Q_M = 7.41, p < .01$; $ES = .05, p < .10$ vs. $ES = .16, p < .001$). Studies observed stronger effects for early learning promotion if they used a latent approach ($Q_M = 6.15, p < .05$; $ES = .19, p < .001$ vs. $ES = .06, p < .05$). None of the studied moderators explained the heterogeneity in effect sizes completely, because a significant amount of residual variance remained ranging from $Q_E(71) = 2139.65$ to 2476.44 for global process quality and from $Q_E(68) = 226.07$ to 264.53 for early learning promotion.

Evaluating Possible Bias

Comparing effect sizes from peer-reviewed articles to effect sizes from other references revealed a marginally significant publication bias only for early learning promotion with non peer-reviewed effect sizes being higher ($Q_M(I) = 3.34, p < .10; ES = .13, p < .001$ vs. $ES = .05, p = .27$; see also Table 3 of the Supporting Online Material). Funnel plots showed slight asymmetry as displayed in Figures 3 and 4 of the Supporting Online Material. The rank correlation test for funnel plot asymmetry implied a significant correlation between effect estimates and sampling variances for early learning promotion (Kendall's $\tau = .25, p < .01$). Nonetheless, fail-safe N calculations suggest the robustness of findings: For the overall effects of global process quality and early learning promotion 1,884 and 3,128 null findings, respectively, are necessary to outreach a significance level of $p < .05$.

Discussion

This meta-analysis found that pedagogical processes as the key factors of ECEC have impact children's development in two important academic domains, i.e. literacy and mathematics. Overall effects for 17 longitudinal studies representing findings for 16,461 children in 2,887 ECEC centers emerged across nine European countries with differing ECEC systems. Processes are significantly related to academic outcomes of children from various family backgrounds across domains, ages and different phases of children's educational career. In line with the ECEC quality frameworks, observed correlational results for global process quality ($ES = .11, d = .22$) and the extent of early learning promotion ($ES = .10, d = .20$) are positive, but small in size. Our estimates of effects are slightly lower as compared to recent related meta-analytic findings for cognitive outcomes ($d = .23 - .34$; Blok et al., 2005; Camilli et al., 2010; Karoly et al., 2005; Manning et al., 2010; Nores & Barnett, 2010). Our findings, however, are not directly comparable, because included studies and the type of evidence diverge from other meta-analyses in several ways, with important implications for ECEC policy. Previously aggregated evidence for the effectiveness of specific programs and multi-facets interventions is not easily generalizable regular ECEC. Though effective, these programs and interventions which mostly target disadvantaged children are not uncontroversial. They are cost- and resource-intensive and are criticized for potentially increasing social segregation and stigmatization, and for failing to reach a large percentage of families most in need (OECD; 2006; Penn, 2011, Barnett, Brown, & Shore, 2004). Moreover, Bradbury, Corak, Waldfogel and Washbrook (2015) critically note that supplementing ECEC systems with programs and interventions for low-income families may actually be detrimental to middle-income families. Furthermore, with the majority of children attending ECEC at a certain age, the pressing ques-

tion of today's ECEC policy is not whether to send children to ECEC, but how to improve ECEC to maximize the developmental benefits of attendance. Comparison of children attending to those not attending ECEC, as previously often included in meta-analyses, yields no answer to the question about the developmental impact of variations within ECEC. Camilli et al. (2010) found a significant overall effect of .23 for cognitive outcomes only for studies comparing participants of ECEC programs or interventions to children with no comparable intervention, whereas no significant effect emerged ($ES = .07$) when comparing participants of interventions and programs with differing characteristics. Our meta-analysis found that pedagogical processes as key quality characteristics of regular provision beneficially impact children's academic development. ECEC policies which aim at increasing academic benefits for majority of children, should address global process quality and the extent of promotion in various domains in regular center-based care. The low to mediocre average levels of global process quality and even more so the extent of early learning promotion in ECEC centers in most countries leave room for improvement (Kuger & Kluczniok, 2008; Slot et al., 2015). The literature provides some suggestions and examples of policies and initiatives, which can assure and enhance the quality of pedagogical processes in ECEC (e.g., Anders et al., 2015; Fried et al., 2012; OECD, 2006; 2011; 2014; Slot et al., 2014).

Characteristics of Outcomes and Process Measures as Moderators

We observed significant variations in overall effects, which could partly be explained by the characteristics of outcomes and process measures. Though significantly related to outcomes in both domains, we found domain-specific accentuation of effects. Global process quality was more strongly associated with literacy outcomes, which might reflect the large share of language-based interaction inherent in global process ratings (Pianta et al., Harms et al. Halle REF). Thus, ECEC settings of high global quality provide beneficial learning opportunities and stimulation for language development to a certain extent, but not necessarily to an equal extent for early math development. It is common understanding in research, policy, and practice, that children need domain-specific stimulation and promotion to acquire essential knowledge and skills in pre-academic domains like early mathematics, science, as well as for pre-writing and -reading skills (Burger, 2014; Harme, Hatfield, Pianta, & Jamil, 2014; Kluczniok & Roßbach, 2014; Kuger & Kluczniok, 2008; OECD, 2006; 2011). We found stronger association of the extent of early learning promotion with math outcomes, which might indicate stronger dependency on specific stimulation and promotion.

Though evaluated programs, curricular approaches, and concepts of teacher trainings of early math education exist (e.g. Baarody & Lai, 2006; Burger, 2014; Neumann & Dickinson,

2011), the extent of early learning promotion in regular provision is still insufficient, in particular for domains other than literacy (Kuger & Kluczniok, 2008; Slot et al., 2015).

In contrast to other syntheses (e.g., Blok et al., 2005; Burger, 2010; Camilli et al., 2011) we found no significant gradual decline of effects with age and no decrease after children left ECEC and were exposed to the institutional influences in primary and secondary school (cf. Nores & Barnett, 2010). We found measures of global process quality with an interaction-focus to be more strongly related to academic outcomes than measures including ratings of the material surrounding. Rating the quality of the surroundings indicates that ECEC settings are safe and well-equipped to prevent children from potential emotional or physical harm and to guarantee the pre-conditions necessary for pre-academic learning (Harms et al., 1998; Kluczniok & Roßbach, 2014). Interactions, however, capture if and how learning opportunities are used, which forms part of pedagogical processes essential for pre-academic learning in ECEC settings. A significant effect of the extent of early learning promotion emerged only if assessed with observations as recommended by Sylva et al. (2006) and Pianta & Hamre (2009). With increased use in large-scale assessments in research and practice, e.g. as essential parts of large-scale longitudinal projects (Durda, 2015) or regular inspections and quality monitoring (Bradshaw et al., 2014; OECD, 2015), intensified efforts to develop more economical alternatives for capturing developmentally beneficial processes in ECEC settings are needed, such as reliable and valid staff questionnaires. In addition, both processual aspects affect growth in outcome and, hence, the effects cannot completely be accounted for by alternative explanations, e.g. ease of achieving high levels of pedagogical processes in high-achieving groups, or selective placement. Considering prior attainments in analysis seems particularly important for the effects of global process quality.

Characteristics of Adopted Analysis as Moderators

Included studies adopted complex designs and diverse analytical approaches. Differences in adopted analyses can partly explain variations in effects (Becker & Wu, 2007; Bowman, 2012; Burger, 2010; Cooper, 2009). The majority of studies used some type of multivariate analysis to estimate ECEC effects while controlling for further factors relevant for child development, which is necessary to obtain an estimate of the unique ECEC effect. Thus, meta-analysis on ECEC effects needs to overcome the common skepticism of basic meta-analytic literature (Cooper, 2009; Becker & Wu, 2007) about combining multivariate effect sizes such as coefficients from multiple regressions or structural equation modeling (Bowman, 2012; Card, 2012; Kim, 2011). One main concern with regard to combining multivariate effect sizes such as coefficients from multiple regressions or structural equation modeling (Cooper, 2009;

Becker & Wu, 2007) is that analyses diverge in number and type of included predictors and, hence, are not directly comparable. Using bivariate correlations in meta-analysis, as often proposed, would not do justice to the complexity of the field, and the quality of aggregated results would be questionable. Recently, very sophisticated meta-analytical approaches evolved to combine regression coefficients, but these approaches require specific information which is seldom provided in references; hence these secondary analyses basically need access to primary data (Becker and Wu, 2007). If a secondary analysis of original data is not possible, meta-analysis needs to rely on the available information on effects and study the characteristics of the adopted analysis as the moderators of the effects (Nores & Barnett, 2010). Thereby, meta-analysis can actually set and underline analytical standards for research in the field. We found that type rather than number of predictors included in the analysis moderated effects of pedagogical processes. We found that considering characteristics of child, the family, and home learning environment, as well as aspects of structural quality as additional predictors related positively to observed effects. The addition of aspects of ECEC quantity as predictors was not associated with effects. Research and syntheses showed, however, that variations in quantity of childcare (e.g. intensity or duration of provision) directly and in interaction with pedagogical processes affect child development (Anders, 2013; Burger, 2010; Melhuish et al., 2015). Hence, future studies should include these aspects in their design and analysis. Furthermore, latent approaches resulted in higher effect sizes for early learning promotion, whereas approaches considering the clustered nature of data led to lower and more conservative estimates for both effects. Researchers should consider recent developments in educational methods for appropriate state-of-the-art analytical approaches, and should carefully consult recent primary research and syntheses to identify relevant factors to be assessed, and to choose suitable analytical approaches.

Limitations and Future Research

Despite successfully including a large proportion of grey literature and unpublished results, our results may be biased. We found slight indications of bias, but overall effects were robust against potential null findings. Included long-term evidence is restricted to the school phase and seldom goes beyond primary school, which is partly caused by our time criterion (study begin after 1990). We did, however, only identify one Swedish studies meeting all other criteria except for the time criterion (i.e., the Goteborg Child Care Study; Broberg, Wessels, Lamb, & Hwang, 1997). Other meta-analyses cover longer developmental periods in their long-term effects (e.g. Blok et al., 2005; Camilli et al., 2012); however, the ECEC systems and research have tremendously changed over the past decades. By eliminating studies that began

before 1990, we were still able to study persistency of effects, but findings reflect the academic benefits of pedagogical processes in ECEC nowadays.

Moreover, it is important to underline that number of effect sizes and studies varied substantially in our moderator analyses with important implications for the reliability of sub-effects. More studies and more timely publication of results are needed to provide a more comprehensive picture of beneficial characteristics of recent/today's ECEC services. This meta-analysis does not cover the interaction of pedagogical processes with factors such as aspects of the quantity of provision or indicators of disadvantage. Though evidence on interaction effects is increasing, there is still not enough comparable evidence for a quantitative synthesis, but recent systematic summaries already exists (Anders, 2013; Melhuish et al., 2015). This meta-analysis is limited to Europe, discarding evidence from other countries, especially the US. As described ECEC systems and research in Europe differs from the US and several meta-analyses have already synthesized the US evidence (e.g., REF). These syntheses led to growing public and scientific awareness of the potential of early center-based education and additional services through offering programs and interventions, especially for disadvantaged children. By aggregating evidence on naturally occurring variations in regular provision, our meta-analysis complements these syntheses and suggests that increasing the general quality of pedagogical processes in ECEC centers can have academic benefits for the majority of children. Both types of research and synthesis are needed for the continuous and systematic development of ECEC systems. A major field for future research and synthesis is the growing sector of private ECEC services such as private day care or playgroups, which diverge from center-based care in important aspects. Private services are, for example, usually less regulated, require less qualification of staff, are attended by younger children with different needs, and are often not obliged to follow curricula or guidelines in their educational work. Finally, the logical next step of this meta-analysis is the expansion to other developmental areas, exploring impact on children's current and future academic and life paths including health, well-being and socioemotional outcomes.

Conclusion

Our meta-analysis on the impact of variations in regular provision in Europe indicates beneficial effects of increasing the quality of pedagogical processes in regular ECEC centers on children's development in two important academic domains, i.e. literacy and mathematics, . ECEC initiatives, research, and governmental financing and regulation should focus on assuring and enhancing global process quality and the stimulation and promotion in various developmental domains.

References

- *Abreu-Lima, I. M. P., Leal, T. B., Cadima, J., & Gamelas, A. M. (2013). Predicting child outcomes from preschool quality in Portugal. *European Journal of Psychology of Education, 28*, 399-420.
- Anders, Y. (2013). Stichwort: Auswirkungen frühkindlicher institutioneller Betreuung und Bildung [Keyword: Impact of centre-based early childhood education and care]. *Zeitschrift für Erziehungswissenschaft, 16*(2), 237-275.
- *Anders, Y., Ballaschk, I., Dietrichkeit, T., Flöter, M., Groeneveld, I., Lee, H.-J.,... Wieduwilt, N. (2015). Evaluation des Bundesprogramms „Schwerpunkt-Kitas Sprache & Integration“ – Abschlussbericht [Evaluation of the federal program „Priority Childcare Centres Language & Integration“ – Final report]. In H.-G. Roßbach, Y. Anders & W. Tietze (Hrsg.), *Evaluation des Bundesprogramms „Schwerpunkt-Kitas Sprache & Integration“ – Abschlussbericht*. Bamberg und Berlin.
- *Anders, Y., Große, C., Roßbach, H.-G., Ebert, S., & Weinert, S. (2013). Preschool and primary school influences on the development of children's early numeracy skills between the ages of 3 and 7 years in Germany. *School Effectiveness and School Improvement, 24*, 195–211. <http://doi.org/10.1080/09243453.2012.749794>
- *Anders, Y., Roßbach, H. G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S., & von Maurice, J. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly, 27*(2), 231-244. <http://doi.org/10.1016/j.ecresq.2011.08.003>
- Anderson, L. M., Shinn, C., Fullilove, M. T., Scrimshaw, S. C., Fielding, J. E., Normand, J., & Carande-Kulis, V. G. (2003). The effectiveness of early childhood development programs. *American Journal of Preventive Medicine, 24*(3), 32–46. [http://doi.org/10.1016/S0749-3797\(02\)00655-4](http://doi.org/10.1016/S0749-3797(02)00655-4)
- Arnett, J. (1989). *Caregiver interaction scale*. Princeton, NJ: Educational Testing Service.
- Arnold, D. H., & Doctoroff, G. L. (2003). The early education of socioeconomically disadvantaged children. *Annual Review of Psychology, 54*(1), 517-545.
- Barnett, W. S. (2011). Effectiveness of early educational intervention. *Science, 333*(6045), 975–978. <http://doi.org/10.1126/science.1204534>
- Barnett, W. S., Brown, K. C., & Shore, R. (2004). *The universal vs. targeted debate: Should the United States have preschool for all?* Retrieved from <http://www.bpichicago.org/wp-content/uploads/2014/01/The-Universal-vs-Targeted-Debate.pdf>

- Baroody, A. J., & Lai, M. (2006). The development of young children's early number and operation sense and its implications for early childhood education. In B. Spedek & O. N. Saracho (Eds.), *Handbook of Research on the Education of Young Children* (2nd ed.). Mahwah, NJ / London: Lawrence Erlbaum Associates.
- Becker, B. J., & Wu, M. J. (2007). The synthesis of regression slopes in meta-analysis. *Statistical Science*, 22(3), 414-429. doi:Doi 10.1214/07-Sts243
- Blok, H., Fukkink, R. G., Gebhardt, E. C., & Leseman, P. P. M. (2005). The relevance of delivery mode and other programme characteristics for the effectiveness of early childhood intervention. *International Journal of Behavioral Development*, 29(1), 35-47. doi:Doi 10.1080/01650250444000515
- Bowman, N. A. (2012). Effect sizes and statistical methods for meta-analysis in higher education. *Research in Higher Education*, 53(3), 375-382. doi:10.1007/s11162-011-9232-5
- Bradbury, B., Corak, M., Waldfogel, J., & Washbrook, E. (2015). *Too many children left behind*. New York: Russel Sage Foundation.
- *Bradshaw P., Lewis G., & Hughes T. (2014). *Growing Up in Scotland: Characteristics of preschool provision and their associations with child outcomes*. Edinburgh: Scottish Government.
- Broberg, A. G., Wessels, H., Lamb, M. E., & Hwang, C. P. (1997). Effects of day care on the development of cognitive abilities in 8-year-olds: A longitudinal study. *Developmental Psychology*, 33, 62-69.
- 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(2), 140-165. doi:10.1016/j.ecresq.2009.11.001
- Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum.
- *Cadima, J. (2015). [Analyses of the effects of preschool quality on literacy and numeracy outcomes for the contexts and transition study]. Unpublished analysis.
- Camilli, 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(3), 579-620.
- Card, N. A. (2012). *Applied meta-analysis for social science research*. New York: Guilford Publications.

- Cooper, H. (2009). *Research synthesis and meta-analysis: A step-by-step approach* (Vol. 2). Thousand Oaks, CA: Sage Publications.
- Dalli, C., White, E. J., Rockel, J., & Duhn, I. (2011). *Quality early childhood education for under-two-year-olds: what should it look like? A literature review*. Wellington, New Zealand: Ministry of Education.
- *de Haan, A., Elbers, E., Hoofs, H., & Leseman, P. (2013). Targeted versus mixed preschools and kindergartens: effects of class composition and teacher-managed activities on disadvantaged children's emergent academic skills. *School Effectiveness and School Improvement*, 24(2), 177-194. doi:10.1080/09243453.2012.749792
- *Durda, T. (2015). *Eltern als Bildungspartner: Effekte von Elternarbeit in Kindertageseinrichtungen auf den sprachlichen Entwicklungsstand von Kindern* [Parents as education-partners: Effects of parent work in preschools on children's language competence]. (Unpublished master's thesis). Otto-Friedrich-Universität, Bamberg.
- *Ebert, S., Lockl, K., Weinert, S., Anders, Y., Kluczniok, K., & Roßbach, H. G. (2013). Internal and external influences on vocabulary development in preschool children. *School Effectiveness and School Improvement*, 24(2), 138-154. doi:Doi 10.1080/09243453.2012.749791
- European Commission/EACEA/Eurydice/Eurostat. (2014). *Key data on early childhood education and care in Europe. 2014 Edition. Eurydice and Eurostat Report*. Luxembourg: Publications Office of the European Union.
- Eurydice. (2009). *Early childhood education and care in Europe: Tackling social and cultural inequalities*. Brussels: Education, Audiovisual and Culture Executive Agency.
- *Flöter, M., Weigel, S., & Schmerse, D. (2015, March). *Effekte in der sprachlichen Bildung von Kindern durch den Besuch von Kindertageseinrichtungen* [Effects of preschool on children's language competence]. Paper presented at the 3rd annual conference of the Association of Empirical Educational Research (GEBF), Bochum, Germany.
- *Fried, L., Hoefl, M., Isele, P., Stude J. & Wexeler, W. (2012). *Schlussbericht zur Wissenschaftlichen Flankierung des Verbundprojekts „TransKiGs – Stärkung der Bildungs- und Erziehungsqualität in Kindertageseinrichtungen und Grundschule – Gestaltung des Übergangs“* [Final report of the scientific evaluation of the cooperative project „TransKiGS – Strengthening the education and care quality of preschools and primary schools – Shaping the transition]. Dortmund: TU Dortmund.
- Halle, T., Vick Whittaker, J. E., & Anderson, R. (2010). *Quality in early childhood care and education settings: A compendium of measures*. Washington, DC: Child Trends.

- Hamre, B., Hatfield, B., Pianta, R., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: associations with preschool children's development. *Child Development*, 85(3), 1257–1274.
- Harms, T., Clifford, M., & Cryer, D. (1998). *Early Childhood Rating Scale, Revised Edition (ECERS-R)*. Vermont: Teachers College Press.
- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), 1900-1902.
- Higgins, J. P. T., & Green, S. (Eds.) (2008). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley.
- Hox, J. (2002). *Multilevel analysis. Techniques and applications*. Mahwah, NJ: Erlbaum.
- Karoly, L.A., Kilburn, R. and Cannon, J.S. (2005). *Early childhood interventions: Proven results, future promise*. St. Monica, California: RAND Corporation.
- Kim, R. S. (2011). *Standardized regression coefficients as indices of effect sizes in meta-analysis*. (Unpublished doctoral dissertation), Florida State University.
- Kluczniok, K., & Roßbach, H.-G. (2014). Conceptions of educational quality for kindergartens. *Zeitschrift Für Erziehungswissenschaft*, 17(6), 145–158.
<http://doi.org/10.1007/s11618-014-0578-2>
- Kuger, S., & Kluczniok, K. (2008). Prozessqualität im Kindergarten—Konzept, Umsetzung und Befunde [Process quality in preschool – Concept, Application, and Evidence]. *Zeitschrift Für Erziehungswissenschaft*, 10, 159–178.
- *Leal, T., Abreu-Lima, I. M. P., & Cadima, J. (2013, August 2013). *Associations between dimensions of quality and children outcomes in preschool and first grade –a Portuguese study*. Paper presented at the 15th Biennial Conference of the European Association for Research in Learning and Instruction (EARLI), Munich, Germany.
- *Lehrl, S., Kluczniok, K., & Rossbach, H.-G. (2016). Longer-term associations of preschool education: The predictive role of preschool quality for the development of mathematical skills through elementary school. *Early Childhood Research Quarterly*, 36, 475–488.
<http://doi.org/10.1016/j.ecresq.2016.01.013>
- Maas, C. J., Hox, J. J., & Lensvelt-Mulders, G. J. (2004). Longitudinal meta-analysis. *Quality and Quantity*, 38(4), 381-389.
- Manning, M., Homel, R., & Smith, C. (2010). A meta-analysis of the effects of early developmental prevention programs in at-risk populations on non-health outcomes in adoles-

science. *Children and Youth Services Review*, 32(4), 506–519.

<http://doi.org/10.1016/j.chilyouth.2009.11.003>

*Melhuish, E. (2015). [Analyses of the effects of preschool quality on literacy and numeracy outcomes for the Effective Pre-School, Primary and Secondary Education Study]. Unpublished analysis.

Melhuish, E., Ereky-Stevens, Petrogiannis, K., Ariescu, A., Penderi, E., Rentzou, K., ...

Leseman, P. (2015). *A review of research on the effects of early childhood education and care (ECEC) on children development*. Retrieved from CARE website: http://ecec-care.org/fileadmin/careproject/Publications/reports/CARE_WP4_D4__1_review_of_effects_of_ecec.pdf.

*Melhuish, E., Quinn, L., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B. & Shields, C. (2002) *The Effective Pre-school Provision in Northern Ireland Project, Technical Paper 5: Pre-school experience and cognitive development at the start of primary school*. Belfast, N.I.: Stranmillis University Press.

*Melhuish, E., Quinn, L., Hanna, K., Sylva, K., Siraj-Blatchford, I., Sammons, P. & Taggart, B. (2004). *The Effective Pre-school Provision in Northern Ireland Project, Technical Paper 10: Pre-school experience and literacy and numeracy development at the end of year 2 of primary school*. Belfast, N.I.: Stranmillis University Press.

Mitchell, L., Wylie, C., & Carr, M. (2008). *Outcomes of early childhood education: Literature review*. New Zealand: Council for Educational Research.

Neuman, S. B., & Dickinson, D. K. (Eds.). (2011). *Handbook of early literacy research* (Vol. 3). New York/London: Guilford Press.

Nores, M., & Barnett, W. S. (2010). Benefits of early childhood interventions across the world: (Under) Investing in the very young. *Economics of Education Review*, 29(2), 271-282. doi:DOI 10.1016/j.econedurev.2009.09.001

OECD. (2006). *Starting strong II: early childhood education and care*. Paris: OECD.

Organisation for Economic Co-operation and Development (2011). *Starting strong III: A quality toolbox for early childhood education and care*. Paris: OECD Publishing.

OECD. (2013). *How do early childhood education and care (ECEC) policies, systems and quality vary across OECD countries?* (Education indicators in focus 11). Retrieved from <http://www.oecd->

[libra-](http://www.oecd-)

[ry.org/docserver/download/5k49czkz4bq2.pdf?expires=1457865862&id=id&accname=guest&checksum=EFC23ED8673B7F30ECAD994FADA2FA69](http://www.oecd-)

- Organisation for Economic Co-operation and Development (2015). *Education at a glance 2015*. Paris: OECD Publishing.
- *Pakarinen, E., Kiuru, N., Lerkkanen, M. K., Poikkeus, A. M., Siekkinen, M., & Nurmi, J. E. (2010). Classroom organization and teacher stress predict learning motivation in kindergarten children. *European Journal of Psychology of Education, 25*(3), 281-300. doi:DOI 10.1007/s10212-010-0025-6
- *Pakarinen, E., Kiuru, N., Lerkkanen, M. K., Poikkeus, A. M., Ahonen, T., & Nurmi, J. E. (2011). Instructional support predicts children's task avoidance in kindergarten. *Early Childhood Research Quarterly, 26*(3), 376-386.
<http://doi.org/10.1016/j.ecresq.2010.11.003>
- *Pakarinen, E., Lerkkanen, M. K., Poikkeus, A. M., Siekkinen, M., & Nurmi, J. E. (2011). Kindergarten teachers adjust their teaching practices in accordance with children's academic pre-skills. *Educational Psychology, 31*(1), 37-53.
doi:10.1080/01443410.2010.517906
- *Pakarinen, E., Lerkkanen, M. K., Silinskas, G., Salminen, J., Poikkeus, A. M., Siekkinen, M., & Nurmi, J. E. (2015, August). *Longitudinal associations between kindergarten classroom interactions and Finnish children's reading skills*. Paper presented at the 16th Biennial EARLI Conference for Research on Learning and Instruction, Limassol, Cyprus.
- Penn, H. (2011). *Quality in early childhood education services - An international perspective*. UK: Berkshire: McGraw-Hill Education.
- 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*(3), 144-159.
- Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*(2), 109-119.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS) Manual: K-3*. Baltimore: Paul H. Brookes Publishing Company.
- *Pinto, A. I., Pessanha, M., & Aguiar, C. (2013). Effects of home environment and center-based child care quality on children's language, communication, and literacy outcomes. *Early Childhood Research Quarterly, 28*(1), 94-101. doi:DOI 10.1016/j.ecresq.2012.07.001

- *Quinn, L., Melhuish, E., Hanna, K., Sylva, K., Siraj-Blatchford, I., Sammons, P., . . . Doyle, A. (2006). *Effective Pre-school Provision Northern Ireland (EPPNI), Technical paper no. 12: Pre-school experience and literacy and numeracy development at the end of key stage 1*. Retrieved from <http://www.bbk.ac.uk/psychology/our-staff/academic/edward-melhuish/documents/EPPNIP4paper12.pdf>
- *Quinn, L., Melhuish, E., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., . . . Sweeney, G. (2003). *Effective Pre-school Provision Northern Ireland (EPPNI) - Pre-school experience and cognitive development at the end of year 1*. Retrieved from http://anu.stran.ac.uk/eppni/documents/CognitiveDevelopmentandProgressat-theEndofYear1_003.pdf
- *Roßbach, H. G., Sechtig, J., & Freund, U. (2010). *Empirische Evaluation des Modellversuchs "Kindergarten der Zukunft in Bayern-KiDZ"* [Empirical evaluation of the model program preschools of the future in Bavaria - KiDZ]. Bamberg: University of Bamberg Press.
- *Silinkas, G., Pakarinen, E., Poikkeus, A. M., Lerkkanen, M. K., & Nurmi, J. E. (2015, August). *Longitudinal effects of kindergarten quality on the development of literacy skills in Grade 1*. Paper presented at the 16th Biennial EARLI Conference for Research on Learning and Instruction, Limassol, Cyprus.
- Slot, P.L., Lerkkanen, M.-K., & Leseman, P.P.M. (2015). *The relations between structural quality and process quality in European early childhood education and care provisions: Secondary analyses of large scale studies in five countries*. Retrieved from http://ecec-care.org/fileadmin/careproject/Publications/reports/CARE_WP2_D2__2_Secondary_data_analyses.pdf
- *Slot, P. L., Mulder, H., Verhagen, J., & Leseman, P. P. M. (2014). *Domain-general and domain-specific quality characteristics of early childhood education and care predict growth of two-year-olds' vocabulary and attention skills over one year*. Manuscript submitted for publication.
- *Sylva, K., Siraj-Blatchford, I., Taggart, B., Sammons, P., Melhuish, E., Elliot, K., & Totsika, V. (2006). Capturing quality in early childhood through environmental rating scales. *Early Childhood Research Quarterly*, 21(1), 76-92. doi:10.1016/j.ecresq.2006.01.003
- Sylva, K., Siraj-Blatchford, I., & Taggart, B. (2003). *Assessing quality in the early years: Early childhood environment rating scale: Extension (ECERS-E), Four curricular subscales*. London: Trentham Books.

- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (Eds.). (2010). *Early Childhood Matters: Evidence from the Effective Pre-school and Primary Education Project*. Oxford: Routledge.
- *Tietze, W., Hundertmark-Mayser, J., & Roßbach, H. (1999). *European child care and education study. School-age assessment of child development: Long-term impact of pre-school experiences on school success and family-school relationships*. Brussels: Research and Development RTD Action: Targeted Socio-Economic Research.
- *Tsakiri, T. (2012). A study of early childhood education and care and child development in Greece. (Unpublished doctoral dissertation), University of London, London.
- Vandell, D., Belsky, J., Burchinal, M., Vandergrift, N., & Steinberg, L., (2010). Do effects of early child care extend to age 15 years? Results From the NICHD Study of Early Child Care and Youth Development: Age 15 follow-up. *Child Development*, *81*(3), 737–756. <http://doi.org/10.1111/j.1467-8624.2010.01431.x>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1-48.

Table 1
Results of the meta-regression for outcome and process measure characteristic

Moderator	Test of moderator	Residual variance		Sample		Effect size of subgroups		
	Q_M^a	Q_E	df_E	k_{ES}	k_{Study}	ES	se	$CI_{95\%}$
Educational phase ^b								
Global - in ECEC	1.78	2332.6***	71	48	11	.09#	.04	.00 - .17
Global - in school				25	6	.17**	.05	.06 - .28
Early learning - in ECEC	1.12	263.75***	68	45	10	.12***	.03	.05 - .18
Early learning - in school				25	6	.07#	.04	.00 - .14
Outcome domain								
Global - Literacy	36.04***	2346.47***	71	46	12	.13***	.04	.06 - .21
Global - Math				27	7	.07*	.04	.00 - .15
Early learning - Literacy	6.32*	235.83***	68	43	10	.08**	.03	.03 - .14
Early learning - Math				27	9	.11***	.03	.06 - .17
Growth vs attainment								
Global - attainment	3.93*	2195.97***	71	32	6	.05	.05	-.05 - .15
Global - growth				41	13	.14***	.04	.07 - .22
Early learning - attainment	1.17	263.22***	68	39	6	.12***	.03	.05 - .19
Early learning - growth				31	10	.08*	.03	.02 - .14
Process measures								
Global - interaction-focus	9.35**	2139.65***	71	24	5	.19***	.05	.09 - .29
Global - surrounding				49	10	.09*	.04	.01 - .18
Early learning - observation	4.80*	226.07***	68	55	8	.12***	.03	.07 - .18
Early learning - questionnaire				15	3	.01	.04	-.07 - .10

Note. *** $p < .001$; ** $p < .01$; * $p < .05$; # $p < .10$. ^a $df = 1$ for all Q_M . ^bage as a continuous predictor did not moderate effects.

Anhang B - Zusätzliches Online-Material zur Metaanalyse (TP1)

Table 1

Overview of included studies

Project	Title of study	Country	Project period	Age	
				First assessment ^a	Last follow-up ^b
Attiki	A Study of Early Childhood Education and Care and Child Development in Greece	Greece	2000 – 2002	3 years	4 years
BiKS	Bildungsprozesse, Kompetenzentwicklung und Formation von Selektionsentscheidungen im Vor- und Grundschulalter	Germany	2005 – ongoing	4 years	9 years
Contexts and Transition	Contexts and Transition	Portugal	2005 – 2008	4 and 5 years	6 years
Early Chances	Early Chances	Germany	2012 - 2016	2 years	5 years
ECCE	European Child Care and Education Study	Germany Spain	2004 – 2013	4 years	6 years
Engagement Study	Longitudinal Study of Children’s Engagement and Adaptation	Portugal	2001 - 2009	< 1 year	5 years
EPPNI	Effective Pre-school Provision in Northern Ireland	Northern Ireland	1998 – 2010	3 years	8 years
EPPSE	Effective Pre-School, Primary and Secondary Education	England	1997 - 2013	3 years	16 years
First Steps	Interaction and Learning within Children-Parent-Teacher Triangle	Finland	2006 - 2016	6 years	10 years
GUS	Growing up in Scotland - Birth cohort 1	Scotland	2005 – ongoing	< 1 year	4 years

(continued)

Table 1 (continued)
Overview of included studies

Project	Title of study	Country	Project period	Age	
				First assessment ^a	Last follow-up ^b
KiDZ	Kindergarten der Zukunft in Bayern	Germany	2004 – 2013	4 years	6 years
NEPS	National Educational Panel Study – Kindergarten cohort	Germany	2010 – ongoing	4 years	7 years
Pre-COOL	Cohort study Pre-COOL	Netherlands	2009 – ongoing	2 years	3 years
TransKiGs	Stärkung der Bildungs- und Erziehungsqualität in Kindertageseinrichtungen und Grundschule – Gestaltung	Germany	2005 - 2009	5 years	7 years
Utrecht Mixed Preschool Groups	Utrecht Mixed Preschool Groups - Preschool and kindergarten cohort	Netherlands	2008 – 2010	3 and 4 years	5 and 6 years

Notes. ^aRefers to the first child assessment not necessarily included in meta-analysis. ^bRefers to the last follow-up included in meta-analysis.

Table 2

Means, standard deviations, and metric of pedagogical process measure for each included study

Global process quality				Early learning promotion			
Measure	<i>M</i>	<i>SD</i>	Metric	Measure	<i>M</i>	<i>SD</i>	Metric
Attiki							
ECERS-R	2.56	0.94	1 – 7 ^a	ECERS-E	1.57	0.43	1 – 7 ^a
BiKS							
ECERS-R	3.73	0.58	1 – 7 ^a	ECERS-E	2.88	0.58	1 – 7 ^a
				ECERS-E literacy	3.23	0.75	1 – 7 ^a
				ECERS-E numeracy	2.60	0.80	1 – 7 ^a
Contexts and Transition							
ECERS-R	3.44	0.72	1 – 7 ^a	ECERS-E literacy	3.26	0.89	1 – 7 ^a
				ECERS-R language	3.50	1.03	1 – 7 ^a
				ECERS-R Intraaction	4.47	1.26	1 – 7 ^a
				ECERS-E numeracy	2.05	0.93	1 – 7 ^a
				ELLCO checklist	16.44	4.30	1 – 41
				ELLCO observation	2.80	0.51	1 – 5 ^b
Early Chances							
ECERS-R (selected items)	4.69	0.75	1 – 7 ^a	ECERS-E (selected items)	2.92	0.86	1 – 7 ^a
CIS (without permissiveness)	3.58	0.37	1 – 4 ^c	DO-RESI (selected items)	4.13	0.98	1 – 7 ^a
ECCE - Germany							
ECERS	4.51	0.71	1 – 7 ^a				
ECCE - Spain							
ECERS	4.05	0.88	1 – 7 ^a				
Engagement Study							
ECERS-R	3.32	0.89	1 – 7 ^a				
EPPNI							
ECERS-R	4.61	0.69	1 – 7 ^a	ECERS-E	2.26	0.66	1 – 7 ^a
EPPSE							
ECERS-R	4.34	1.00	1 – 7 ^a	ECERS-E	3.12	1.00	1 – 7 ^a
First Steps							
CLASS – Emotional support	5.12	0.52	1 – 7 ^a	Reading-related activities	3.13	0.52	1 – 5 ^d
CLASS – Classroom organisation	5.34	0.38	1 – 7 ^a				
CLASS – Instructional support	3.96	0.72	1 – 7 ^a				

(continued)

Table 2 (continued)

Means, standard deviations, and metric of pedagogical process measure for each included study

Global process quality				Early learning promotion			
Measure	<i>M</i>	<i>SD</i>	Metric	Measure	<i>M</i>	<i>SD</i>	Metric
GUS							
SCSWIS Care and support	4.70	0.66	1 – 6 ^e	QI Improvement in performance	4.30	1.98	1 – 6 ^e
SCSWIS Environment	4.50	0.99	1 – 6 ^e	QI Children's experiences	4.50	1.98	1 – 6 ^e
				QI Meeting learning needs	4.30	2.31	1 – 6 ^e
KiDZ							
ECERS-R - TG	4.30	N/R	1 – 7 ^a	ECERS-E TG	3.90	N/R	1 – 7 ^a
ECERS-R - CG	4.10	N/R	1 – 7 ^a	ECERS-E - CG	3.20	N/R	1 – 7 ^a
NEPS							
				Stimulating activities in target child	6.10	0.60	0 – 6 ^f
Pre-COOL							
CLASS – Emotional support	4.94	0.70	1 – 7 ^a	ECERS-E	2.24	0.84	1 – 7 ^a
CLASS – Behavioral support	5.85	0.48	1 – 7 ^a				
CLASS – Support for learning	3.23	0.78	1 – 7 ^a				
TransKiGs							
ECERS-R	3.00	0.62	1 – 7 ^a	ECERS-E	1.13	0.84	1 – 7 ^a
CIS (without permissiveness)	3.04	0.23	1 – 4 ^c	DO-RESI-E-KiGs	2.67	0.82	1 – 7 ^a
Utrecht Mixed Preschools Groups – Preschool cohort							
				Teacher-managed literacy activities - TG	8.50	6.40	0 – 100 ^g
				Teacher-managed literacy activities - CG	7.40	5.90	0 – 100 ^g
				Teacher-managed math activities - TG	1.40	1.20	0 – 100 ^g
				Teacher-managed math activities - CG	0.03	0.05	0 – 100 ^g
Utrecht Mixed Preschools Groups – Kindergarten cohort							
				Teacher-managed literacy activities - TG	13.70	2.70	0 – 100 ^g
				Teacher-managed literacy activities - CG	17.00	6.60	0 – 100 ^g
				Teacher-managed math activities - TG	2.10	0.90	0 – 100 ^g
				Teacher-managed math activities - CG	1.40	1.00	0 – 100 ^g

Notes. TG = Treatment Group; CG = Control group; N/R = not reported.

Ranging from ^a1 = *inadequate* to 7 = *excellent*; ^b1 = *deficient* 5 = *exemplary*; ^c1 = *not at all* to 4 = *very much characteristic*; ^d1 = *never* to 5 = *many times per day*; ^e1 = *unsatisfactory* to 6 = *excellent*; ^f0 = *never* to 6 = *daily*; ^g0 to 100 % of observation.

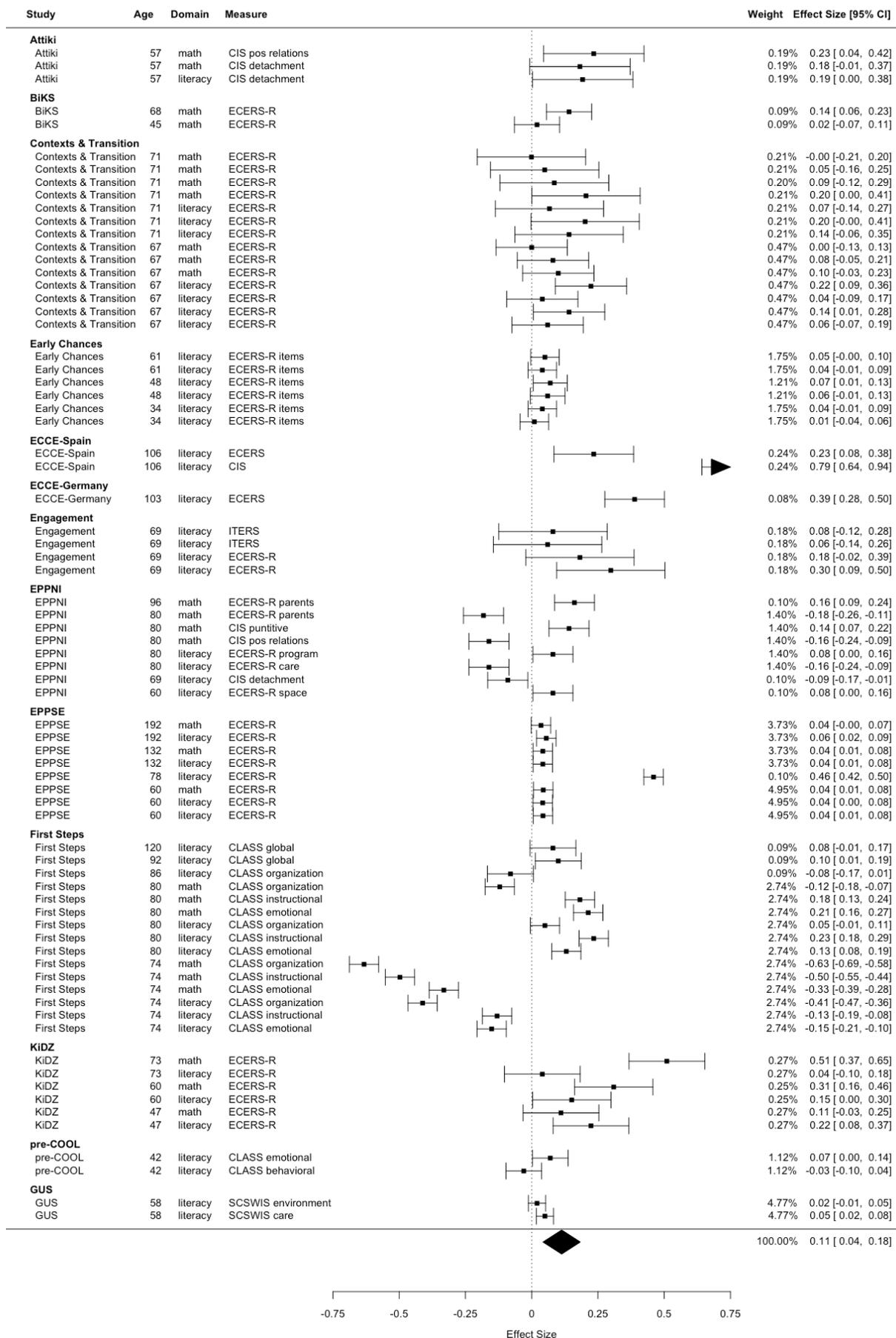


Figure 1. Forest plot for the effect of global process quality. Each tick mark represents an individual effect size and the line its corresponding confidence interval. The diamond shape at the bottom shows the overall effect size (.11**) and the dotted line marks a null effect.

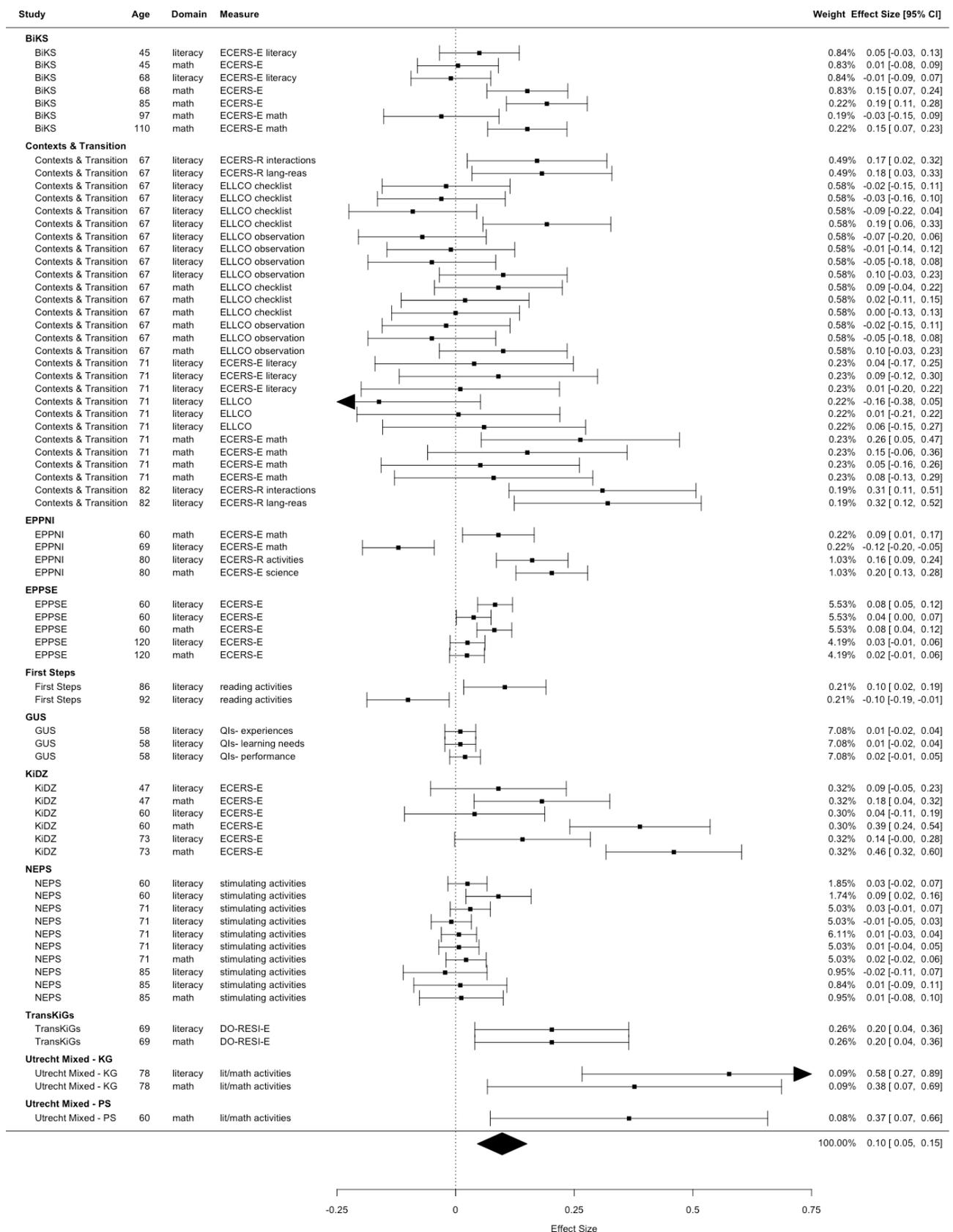


Figure 2. Forest plot for the effect of early learning promotion. Each tick mark represents an individual effect size and the line its corresponding confidence interval. The diamond shape at the bottom shows the overall effect size (.10***) and the dotted line marks a null effect.

Table 3

Mixed-effects-meta-regressions for characteristics of adopted analysis as moderators

Subeffect	Test of moderator	Residual variance		Sample		Effect		
	Q_M^a	Q_E	df_E	k_{ES}	$k_{Studien}$	ES	SE	$CI_{95\%}$
Number of predictors								
Global – low	0.73	2194.50***	69	37	5	.08	.06	-.03 - .20
Global – high				34	10	.14**	.04	.05 - .22
Early learning – low	0.003	252.53***	65	38	6	.11*	.05	.02 - .20
Early learning – high				29	7	.11**	.04	.03 - .18
Child characteristics								
Global – no	3.60#	2166.37***	71	23	4	.003	.06	-.11 - .12
Global – yes				50	13	.12***	.03	.06 - .19
Early learning – no	4.64*	264.53***	68	20	5	.04	.05	-.07 - .14
Early learning – yes				50	8	.13**	.04	.05 - .21
Family characteristics								
Global – no	2.37	2166.45***	71	26	5	.05	.06	-.07 - .16
Global – yes				47	13	.13**	.04	.06 - .21
Early learning – no	5.97*	262.55***	68	24	6	.05	.05	-.06 - .15
Early learning – yes				46	8	.15**	.05	.05 - .24
HLE								
Global – no	4.37*	2432.97***	71	53	9	.07	.04	-.02 - .15
Global – yes				20	9	.18***	.05	.09 - .28
Early learning – no	0.06	263.73***	68	49	10	.10**	.04	.03 - .17
Early learning – yes				21	4	.11*	.05	.02 - .20
Structural quality								
Global – no	0.05	2471.17***	71	60	9	.11*	.04	.02 - .20
Global – yes				13	6	.12*	.06	.00 - .25
Early learning – no	3.60#	256.05***	68	57	7	.08**	.03	.02 - .14
Early learning – yes				13	6	.14***	.04	.07 - .22
ECEC quantity								
Global – no	0.58	2476.44***	71	63	12	.13**	.04	.05 - .20
Global – yes				10	4	.07	.07	-.08 - .21
Early learning – no	0.48	254.54***	68	59	11	.11**	.03	.05 - .17
Early learning – yes				11	3	.07	.05	-.03 - .17
Latent approach								
Global – no	0.10	2471.77***	71	65	11	.11**	.04	.03 - .19
Global – yes				8	4	.13#	.07	-.01 - .28
Early learning – no	6.15*	233.76***	68	57	9	.06*	.03	.01 - .12
Early learning – yes				13	4	.19***	.05	.10 - .28
Multilevel approach								
Global – no	2.72#	2362.73***	71	51	9	.17**	.05	.07 - .27
Global – yes				22	7	.08*	.03	.01 - .15
Early learning – no	7.41**	264.31***	68	36	6	.16***	.03	.10 - .23
Early learning – yes				34	7	.05#	.03	-.01 - .10

Note. *** $p < .001$. ** $p < .01$. * $p < .05$. # $p < .10$. ^a $df = 1$ for all Q_M .

Table 4

Information on included reference, effect sizes, and moderators for each study

Source		Information on effect					Outcome measure			Process measure			Analytical approach							
Project	Reference	Peer-Review? Effect	ES	N	Age	In school? Growth?	Title	Math?	Title	surrounding? questionnaire?	Type	N of predictors	High N pred.?	Child char.?	Family char.? HLE?	Structural q.?	Quantity?	Multilevel?	Latent?	
Attiki	Tsakiri (2012)	○	G	-.19 ^a	110	57	○ ●	BAS verbal comprehension	○	CIS detachment	○ -	Reg	5	○ ●	○ ○	○ ○	○ ○	○ ○	○ ○	
		○	G	-.18 ^a	110	57	○ ●	BAS early number concept	●	CIS detachment	○ -	Reg	9	● ●	● ○	○ ○	○ ○	○ ○	○ ○	
		○	G	.23	110	57	○ ●	BAS early number concept	●	CIS positive relations	○ -	Reg	9	● ●	● ○	○ ○	○ ○	○ ○	○ ○	
	Anders et al. (2012)	●	EL	-.03 ^b	532	45	○ ○	K-ABC arithmetics	●	ECERS-E	- ●	LGM - I	16	● ●	● ●	● ●	● ●	● ●	● ●	● ●
		●	G	.02	532	45	○ ○	K-ABC arithmetics	●	ECERS-R	● -	LGM - I	16	● ●	● ●	● ●	● ●	● ●	● ●	● ●
		●	EL	.15	532	68	○ ●	K-ABC arithmetics	●	ECERS-E	- ●	LGM - S	18	● ●	● ●	● ●	● ●	● ●	● ●	● ●
		●	G	.14	532	68	○ ●	K-ABC arithmetics	●	ECERS-R	● -	LGM - S	18	● ●	● ●	● ●	● ●	● ●	● ●	● ●
	Anders et al. (2013)	●	EL	.04 ^b	532	45	○ ○	K-ABC arithmetics	●	ECERS-E	- ●	LGM - I	16	● ●	● ●	● ●	● ●	● ●	● ●	● ●
		●	EL	.19	532	85	● ●	K-ABC arithmetics	●	ECERS-E	- ●	LGM - S	17	● ●	● ●	● ●	● ●	● ●	● ●	● ●
	Ebert et al. (2013)	●	EL	.05	547	45	○ ○	Receptive Vocabulary	○	ECERS-E literacy	- ●	LGM - I	15	● ●	● ●	● ●	● ●	● ●	● ●	● ●
		●	EL	-.01	547	68	○ ●	Receptive Vocabulary	○	ECERS-E literacy	- ●	LGM - S	17	● ●	● ●	● ●	● ●	● ●	● ●	● ●
	Lehrl et al. (2016)	●	EL	-.03	263	97	● ○	K-ABC arithmetics	●	ECERS-E mathematics	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○
●		EL	.15	554	110	● ●	K-ABC arithmetics	●	ECERS-E mathematics	- ●	LGM - S	16	● ●	● ●	● ●	○ ○	● ●	○ ○	● ●	
Contexts and Transition	Abreu-Lima, Leal et al. (2012)	●	EL	.19	215	67	○ ○	Concepts about print	○	ELLCO checklist	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	EL	-.09	215	67	○ ○	Vocabulary	○	ELLCO checklist	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	EL	-.03	215	67	○ ○	Phonological awareness	○	ELLCO checklist	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	EL	.10	215	67	○ ○	Concepts about print	○	ELLCO observation	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	EL	-.05	215	67	○ ○	Vocabulary	○	ELLCO observation	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	EL	-.01	215	67	○ ○	Phonological awareness	○	ELLCO observation	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	EL	-.07	215	67	○ ○	Letters	○	ELLCO observation	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	EL	.00	215	67	○ ○	WPPSI-R Math	●	ELLCO checklist	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	EL	.02	215	67	○ ○	Number identification	●	ELLCO checklist	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	EL	.09	215	67	○ ○	Rote Rote counting	●	ELLCO checklist	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	EL	.10	215	67	○ ○	WPPSI-R Math	●	ELLCO observation	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	EL	-.05	215	67	○ ○	Number identification	●	ELLCO observation	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	EL	-.02	215	67	○ ○	Rote Rote counting	○	ELLCO observation	- ●	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
		●	G	.22	215	67	○ ○	Concepts about print	○	ECERS-R	● -	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	
●	G	.04	215	67	○ ○	Vocabulary	○	ECERS-R	● -	Corr	1	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○			

Table 4 (continued)

Source		Information on effect				Outcome measure			Process measure			Analytical approach										
Project	Reference	Peer-Review? Effect	ES	N	Age	In school?	Growth?	Title	Math?	Title	surrounding? questionnaire?	Type	N of predictors	High N pred.?	Child char.?	Family char.?	HLE?	Structural q.?	Quantity?	Multilevel?	Latent?	
ECCE	Flöter et al. (2015)	○ G	.04	1,331	61	○	●	Sentence comprehension	○	ECERS-R selection	● -	Reg	16	●	●	●	●	●	●	●	○	○
		○ G	.07	931	48	○	●	Receptive Vocabulary	○	ECERS-R selection	● -	Reg	11	●	●	●	●	●	○	○	●	○
		○ G	.06	901	48	○	●	Sentence comprehension	○	ECERS-R selection	● -	Reg	11	●	●	●	●	●	○	○	●	○
		○ G	.37	306	103	●	●	PPVT	○	ECERS	● -	Reg	26	●	●	●	●	●	○	○	○	○
Engagement Study	Pinto et al. (2013)	○ G	.66	173	106	●	●	PPVT	○	CIS	○ -	Reg	26	●	●	●	●	●	○	○	○	○
		○ G	.23	173	106	●	●	PPVT	○	ECERS	● -	Reg	26	●	●	●	●	●	○	○	○	○
		● G	.18	95	69	○	●	Hearing and Speech	○	ECERS-R	● -	Reg	5	○	●	●	●	○	○	○	●	○
EPPNI	Melhuish et al. (2002)	○ EL	.09	678	60	○	●	Early Number Concepts	●	ECERS-E math	- ●	Reg	16	●	●	●	○	○	●	●	○	
		○ G	.08	678	60	○	●	Verbal Skills	○	ECERS-R space	● -	Reg	24	●	●	●	○	○	●	●	○	
	Melhuish et al. (2004)	○ EL	.16	676	80	●	●	Literacy	○	ECERS-R activities	- ●	Reg	21	●	●	●	○	○	○	○	●	○
		○ EL	.20	676	80	●	●	Numeracy	●	ECERS-E science	- ●	Reg	20	●	●	●	○	○	○	○	●	○
		○ G	-.16	676	80	●	●	Literacy	○	ECERS-R care	● -	Reg	21	●	●	●	○	○	○	○	●	○
		○ G	.08	676	80	●	●	Literacy	○	ECERS-R program	● -	Reg	21	●	●	●	○	○	○	○	●	○
		○ G	-.14 ^a	676	80	●	●	Numeracy	●	CIS puntitiveness	○ -	Reg	20	●	●	●	○	○	○	○	●	○
		○ G	-.16	676	80	●	●	Numeracy	●	CIS positive relations	○ -	Reg	20	●	●	●	○	○	○	○	●	○
	Quinn et al. (2003)	○ EL	-.12	676	69	●	●	Word reading	○	ECERS-E math	●	Reg	22	●	●	●	○	○	●	●	○	
		G	.09 ^a	676	69	●	●	Pre-reading	○	CIS detachment	○ -	Reg	27	●	●	●	○	●	○	○	●	○
Quinn et al.	○ G	.16	683	96	●	●	Numeracy	●	ECERS-R parents	● -	Reg	26	●	○	●	●	○	○	○	●	○	
EPPSE	Melhuish (2015)	○ G	.43	2,857	78	●	●	Literacy	○	ECERS-R	● -	Reg	47	●	●	●	○	○	○	○	●	○
		○ EL	.025	2,857	120	●	●	Literacy	○	ECERS-E	- ●	Reg	47	●	●	●	○	○	○	○	●	○
		○ EL	.024	2,857	120	●	●	Numeracy	●	ECERS-E	- ●	Reg	47	●	●	●	○	○	○	○	●	○
		○ G	.042	2,857	132	●	●	Literacy	○	ECERS-R	● -	Reg	47	●	●	●	○	○	○	○	●	○
		○ G	.042	2,857	132	●	●	Numeracy	●	ECERS-R	● -	Reg	47	●	●	●	○	○	○	○	●	○
		○ G	.055	2,857	192	●	●	Literacy	○	ECERS-R	● -	Reg	47	●	●	●	○	○	○	○	●	○
		○ G	.035	2,857	192	●	●	Numeracy	●	ECERS-R	● -	Reg	47	●	●	●	○	○	○	○	●	○

Table 4 (continued)

Source		Information on effect					Outcome measure			Process measure			Analytical approach											
Project	Reference	Peer-Review? Effect	ES	N	Age	In school?	Growth?	Title	Math?	Title	surrounding? questionnaire?	Type	N of predictors	High N pred.?	Child char.?	Family char.?	HLE?	Structural q.?	Quantity?	Multilevel?	Latent?			
	Sylva et al. (2006)	●	EL	.038	2,857	60	○	●	Language	○	ECERS-E	-	●	Reg	45	●	●	●	●	○	●	●	○	
		●	EL	.083	2,857	60	○	●	Pre-reading	○	ECERS-E	-	●	Reg	43	●	●	●	●	○	●	●	○	
		●	EL	.0815	2,857	60	○	●	Math concepts	●	ECERS-E	-	●	Reg	42	●	●	●	●	○	●	●	○	
		●	G	.0415	2,857	60	○	●	Language	○	ECERS-R	●	-	Reg	46	●	●	●	●	○	●	●	○	
		●	G	.0425	2,857	60	○	●	Pre-Reading	○	ECERS-R	●	-	Reg	43	●	●	●	●	○	●	●	○	
		●	G	.0435	2,857	60	○	●	Math concepts	●	ECERS-R	●	-	Reg	42	●	●	●	●	○	●	●	○	
First Steps	Pakarinen et al. (2010)	●	G	-.15	1,268	74	○	○	Pre-Reading skills	○	CLASS Emotional	○	-	Corr	1	○	○	○	○	○	○	○	●	○
		●	G	-.39	1,268	74	○	○	Pre-Reading skills	○	CLASS Organization	○	-	Corr	1	○	○	○	○	○	○	○	●	○
		●	G	-.13	1,268	74	○	○	Pre-Reading skills	○	CLASS Instructional	○	-	Corr	1	○	○	○	○	○	○	○	●	○
		●	G	.13	1,268	80	○	○	Phonological awareness	○	CLASS Emotional	○	-	Corr	1	○	○	○	○	○	○	○	●	○
		●	G	.05	1,268	80	○	○	Phonological awareness	○	CLASS Organization	○	-	Corr	1	○	○	○	○	○	○	○	●	○
		●	G	.23	1,268	80	○	○	Phonological awareness	○	CLASS Instructional	○	-	Corr	1	○	○	○	○	○	○	○	○	●
	Pakarinen et al. (2011)	●	G	.21	1,268	80	○	○	Math skills	●	CLASS Emotional	○	-	Corr	1	○	○	○	○	○	○	○	●	○
		●	G	-.12	1,268	80	○	○	Math skills	●	CLASS Organization	○	-	Corr	1	○	○	○	○	○	○	○	●	○
		●	G	.18	1,268	80	○	○	Math skills	●	CLASS Instructional	○	-	Corr	1	○	○	○	○	○	○	○	●	○
	Pakarinen et al. (2015)	○	G	.10	515	92	●	●	Reading fluency	○	CLASS Global score	○	-	LGM - I	5	○	○	●	○	○	○	○	○	●
		○	G	.08	515	120	●	●	Reading skills	○	CLASS Global score	○	-	SEM	6		●	●	●	○	○	○	○	●
	Pakarinen, Lerkkanen et al. (2011)	●	G	-.32	1,268	74	○	○	Pre-Math skills	●	CLASS Emotional	○	-	Corr	1	○	○	○	○	○	○	○	●	○
		●	G	-.56	1,268	74	○	○	Pre-Math skills	●	CLASS Organization	○	-	Corr	1	○	○	○	○	○	○	○	●	○
		●	G	-.46	1,268	74	○	○	Pre-Math skills	●	CLASS Instruction	○	-	Corr	1	○	○	○	○	○	○	○	●	○
	Silinskas et al. (2015)	○	EL	.11 ^b	514	86	●	○	Reading	○	Reading activities	-	○	Path	3	○	●	●	○	●	○	○	●	○
		○	EL	.10 ^b	514	86	●	●	Reading	○	Reading activities	-	○	Path	3	○	●	●	○	●	○	○	●	○
		○	EL	.10 ^b	514	86	●	●	Reading	○	Reading activities	-	○	Path	3	○	●	●	○	●	○	○	●	○
		○	G	-.08	514	86	●	●	Reading	○	CLASS organization	○	-	Path	3	○	●	●	○	●	○	○	●	○
		○	EL	-.10	514	92	●	●	Reading	○	Reading activities	-	○	Path	3	○	●	●	○	●	○	○	●	○
GUS	Bradshaw et al. (2014)	○	EL	.01	3,658	58	○	●	Vocabulary	○	QI Children's experiences	-	○	Reg	-	-	●	●	○	○	○	○	○	○
		○	EL	.01	3,658	58	○	●	Vocabulary	○	OI Meeting learning needs	-	○	Reg	-	-	●	●	○	○	○	○	○	○
		○	EL	.02	3,658	58	○	●	Vocabulary	○	OI Improvement in perform.	-	○	Reg	-	-	●	●	○	○	○	○	○	○
		○	G	.05	3,658	58	○	●	Vocabulary	○	SCSWIS Care	●	-	Reg	-	-	●	●	○	○	○	○	○	●

Table 4 (continued)

Source		Information on effect					Outcome measure			Process measure			Analytical approach													
Project	Reference	Peer-Review? Effect	ES	N	Age	In school?	Growth?	Title	Math?	Title	surrounding?	questionnaire?	Type	N of predictors	High N pred.?	Child char.?	Family char.?	HLE?	Structural q.?	Quantity?	Multilevel?	Latent?				
KIDZ	Roßbach et al. (2010)	○ G	.02	3,658	58	○	●	Vocabulary	○	SCSWIS Environment	●	-	Reg	-	-	●	●	○	○	○	○	●	○			
		○ EL	.09	191	47	○	○	PPVT	○	ECERS-E	-	●	LGM - I	2	○	●	○	○	○	○	○	○	○	●		
		○ EL	.18	191	47	○	○	K-ABC arithmetics	●	ECERS-E	-	●	LGM - I	2	○	●	○	○	○	○	○	○	○	○	●	
		○ G	.22	191	47	○	○	PPVT	○	ECERS-R	●	-	Corr	2	○	●	○	○	○	○	○	○	○	○	○	
		○ G	.11	191	47	○	○	K-ABC arithmetics	●	ECERS-R	●	-	Corr	2	○	●	○	○	○	○	○	○	○	○	○	
		○ EL	.04	179	60	○	○	PPVT	○	ECERS-E	-	●	Par corr	2	○	●	○	○	○	○	○	○	○	○	○	
		○ EL	.37	179	60	○	○	K-ABC arithmetics	●	ECERS-E	-	●	Par corr	2	○	●	○	○	○	○	○	○	○	○	○	
		○ G	.15	179	60	○	○	PPVT	○	ECERS-R	●	-	Corr	2	○	●	○	○	○	○	○	○	○	○	○	
		○ G	.30	179	60	○	○	K-ABC arithmetics	●	ECERS-R	●	-	Corr	2	○	●	○	○	○	○	○	○	○	○	○	
		○ EL	.14	191	73	○	○	PPVT	○	ECERS-E	-	●	LGM - S	6	●	●	●	●	○	○	○	○	○	○	○	●
		○ EL	.43	191	73	○	○	K-ABC arithmetics	●	ECERS-E	-	●	LGM - S	6	●	●	●	●	○	○	○	○	○	○	○	●
		○ G	.04	191	73	○	○	PPVT	○	ECERS-R	●	-	LGM - S	6	●	●	●	●	○	○	○	○	○	○	○	●
○ G	.47	191	73	○	○	K-ABC arithmetics	●	ECERS-R	●	-	LGM - S	6	●	●	●	●	○	○	○	○	○	○	○	●		
NEPS	Durda Ulferts (2016)	○ EL	.09	813	60	○	○	Receptive vocabulary	○	Dev. stimulating activities	-	○	Reg	18	●	●	●	●	●	○	○	○	○	○		
		○ EL	.022	2,254	60	○	○	Receptive grammar	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	.025	2,254	60	○	○	Receptive grammar	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	.001 ^b	2,256	71	○	○	Early letter knowledge	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	-.001 ^b	2,256	71	○	○	Onset-reim-synthesis	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	-.003 ^b	2,256	71	○	○	Rhyme	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	.028 ^b	2,256	71	○	○	Phoneme identification	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	.016 ^b	2,256	71	○	○	Mathematical competence	●	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	.034 ^b	2,140	71	○	○	Early letter knowledge	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	.001 ^b	2,796	71	○	○	Onset-reim-synthesis	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	-.01 ^b	2,140	71	○	○	Rhyme	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	.002 ^b	2,140	71	○	○	Phoneme identification	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	.025 ^b	2,140	71	○	○	Mathematical competence	●	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	.007	2,140	71	○	○	Early letter knowledge	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
		○ EL	.007	2,795	71	○	○	Onset-reim-synthesis	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○	
○ EL	-.009	2,140	71	○	○	Rhyme	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	○	○	○	○	○			

Table 4 (continued)

Source		Information on effect					Outcome measure		Process measure			Analytical approach											
Project	Reference	Peer-Review? Effect	ES	N	Age	In school?	Growth?	Title	Math?	Title	surrounding?	questionnaire?	Type	N of predictors	High N pred.?	Child char.?	Family char.?	HLE?	Structural q.?	Quantity?	Multilevel?	Latent?	
		○	EL	.031	2,140	71	○	○	Phoneme identification	○	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	●	○
		○	EL	.022	2,140	71	○	○	Mathematical competence	●	Dev. stimulating activities	-	○	Reg	9	●	●	●	●	○	○	●	○
		○	EL	-.064 ^b	498	85	●	●	Receptive grammar	○	Dev. stimulating activities	-	○	Reg	10	●	●	●	●	○	○	●	○
		○	EL	-.031 ^b	498	85	●	●	Receptive vocabulary	○	Dev. stimulating activities	-	○	Reg	10	●	●	●	●	○	○	●	○
		○	EL	.006 ^b	498	85	●	●	Mathematical competence	●	Dev. stimulating activities	-	○	Reg	10	●	●	●	●	○	○	●	○
		○	EL	-.018 ^b	495	85	●	●	Receptive grammar	○	Dev. stimulating activities	-	○	Reg	10	●	●	●	●	○	○	●	○
		○	EL	-.057 ^b	495	85	●	●	Receptive vocabulary	○	Dev. stimulating activities	-	○	Reg	10	●	●	●	●	○	○	●	○
		○	EL	.033 ^b	495	85	●	●	Mathematical competence	●	Dev. stimulating activities	-	○	Reg	10	●	●	●	●	○	○	●	○
		○	EL	.01	402	85	●	●	Receptive grammar	○	Dev. stimulating activities	-	○	Reg	10	●	●	●	●	○	○	●	○
		○	EL	-.022	495	85	●	●	Receptive vocabulary	○	Dev. stimulating activities	-	○	Reg	10	●	●	●	●	○	○	●	○
		○	EL	.012	495	85	●	●	Mathematical competence	●	Dev. stimulating activities	-	○	Reg	10	●	●	●	●	○	○	●	○
pre-COOL	Slot et al. (2014)	○	G	-.03	850	42	○	●	Vocabulary	○	CLASS Behavioral Supp.	○	-	SEM	10	●	●	●	○	○	○	●	●
		○	G	.07	850	42	○	●	Vocabulary	○	CLASS Emotional Supp.	○	-	SEM	10	●	●	●	○	○	○	●	●
TransKiGs	Fried et al. (2012)	○	EL	.20	149	69	○	○	Narration	○	DO-RESI-E	-	●	Path	4	○	○	○	○	●	○	○	○
		○	EL	.20	149	69	○	○	OTZ	●	DO-RESI-E	-	●	Path	4	○	○	○	○	●	○	○	○
Utrecht	de Haan et al. (2013)	●	EL	.52	43	78	○	●	PIPS-emergent literacy	○	Literacy/math activities	-	●	LGM – S	5	○	○	○	○	●	○	○	●
Mixed		●	EL	.36	43	78	○	○	PIPS-emergent math	○	Literacy/math activities	-	●	LGM – S	7	○	○	○	○	○	○	○	●
Preschool		●	EL	.35	48	60	○	●	PIPS-emergent math	●	Literacy/math activities	-	●	LGM – S	7	●	○	○	○	○	○	○	●

Notes. ●/○ – Yes/No; G = global process quality; EL = early learning promotion; (Par) Corr = (Partial) Correlation, Reg = Regression; LGM = latent growth modelling; I = Intercept; S = Slope; Path = Path Model, SEM = structural equation modelling.

^a inverse ES used, ^b Average effect used if ESs referred to same outcome at the same time point with equally detailed model.

Anhang C – Manuskript zur mathematischen Bildungseffektivität (TP2)

Originalmanuskript: Ulferts, H., Anders, Y., Roßbach, H.-G. & Weinert, S. (2016). Effektivität frühkindlicher Bildung und Betreuung in Deutschland – Diskussion und empirische Untersuchung des Ansatzes am Beispiel des Bildungsbereichs Mathematik. Manuskript in Vorbereitung.

Zusammenfassung

Im Beitrag wird das Konzept der Bildungseffektivität für die frühkindliche institutionelle Bildung und Betreuung in Deutschland am Beispiel des Bildungsbereichs Mathematik diskutiert. Er thematisiert Vorteile und Grenzen des Effektivitätsansatzes für Forschung und Praxis anhand eines Vergleichs zum Qualitätsansatz und verdeutlicht seine komplementäre Funktion. Neben Voraussetzungen und Herangehensweisen zur Bestimmung von Bildungseffektivität werden Ergebnisse eigener Effektivitätsanalysen für 97 Einrichtungen der BIKS-3-10-Studie vorgestellt, die auf den durch den K-ABC-Untertest Rechnen erhobenen mathematischen Kompetenzen von 554 Kindern basieren. In fairen Vergleichen resultieren keine Effektivitätsunterschiede. Die Ergebnisse sprechen eher für Unterschiede in den Zusammensetzungen der untersuchten Gruppen hinsichtlich Ausgangskompetenzen und entwicklungsrelevanter Merkmale (z. B. sozioökonomischer Status, Geschlecht). Kompetenzunterschiede zwischen Gruppen sind durch Unterschiede in den Ausgangskompetenzen sowie entwicklungsrelevanten Merkmalen und nicht durch Effektivitätsunterschiede erklärbar.

Schlagwörter: Frühkindliche Bildung – Effektivität – Qualität – Mathematik

Abstract

This article discusses the concept of educational effectiveness for early childhood education and care in Germany, using early math education as an example. It addresses advantages and limitations of the effectiveness approach in research and practice by comparison to the approach of educational quality, and highlights its complementary function. Along with requirements and approaches to determine educational effectiveness, the article presents effectiveness results for 97 preschools or groups of the BIKS-3-10 study, based on mathematical competence assessments with the K-ABC subscale arithmetics of 554 children. In fair comparisons no differences in effectiveness emerged. Instead the results point to differences in group composition by initial competence and development related characteristics (e.g., parental socioeco-

conomic status, gender). Differences in competence between groups are explained by differences in development related characteristics rather than divergent effectiveness of centers.

Key words: Early education – educational effectiveness – quality - mathematics

1. Einleitung

Nach dem erfolgten Ausbau des frühkindlichen Bildungs- und Betreuungsangebots ist die Sicherstellung und Verbesserung der Güte des Angebots eine der zentralen bildungspolitischen Aufgaben. Hierzu gilt es die Güte bestehender Angebote zu ermitteln und kontinuierlich zu prüfen, unter anderem, ob und wie gut die Kinder in verschiedenen Bildungsbereichen gefördert werden. Die Förderung in Mathematik wurde schon von Fröbel und Montessori als wichtiges Feld früher Bildung anerkannt. Seit Einführung der Rahmen- und Orientierungspläne ist sie als Aufgabe jeder frühkindlichen Einrichtung beschrieben (Royar, 2007), wenn auch erhebliche Unterschiede zwischen den Bildungsplänen verschiedener Bundesländer und in der Art und Weise, wie frühe mathematische Bildung implementiert wird. Bildungspolitisch ergibt sich daraus die Aufgabe, allgemein bzw. flächendeckend eine gute mathematische Bildung im frühkindlichen Bildungs- und Betreuungssystem sicherzustellen. Damit eignet sich der mathematische Bildungsbereich als Beispiel zur Erläuterung von Ansätzen zur Beurteilung elementarer Bildungsangebote.

Etabliert hat sich in diesem Zusammenhang der Qualitätsansatz (Katz, 1996); vor allem der strukturell-prozessuale Ansatz (Kuger & Kluczniok, 2008). In diesem Ansatz wird die Güte von Bildungsangeboten anhand verschiedener, struktureller und prozessbezogener Merkmale des Angebots beurteilt, wie dem Erzieher-Kind-Schlüssel oder dem Interaktionsklima in der Gruppe. Er birgt allerdings wichtige Herausforderungen, u. a. hinsichtlich der ökonomischen und repräsentativen Erfassung von Prozessmerkmalen. Das verstärkt die Diskussion ergänzender Möglichkeiten der Beurteilung. Im Effektivitätsansatz wird anstelle von Merkmalen des Betreuungsangebots der Kompetenzstand der Kinder in den Einrichtungen bzw. ihre Kompetenzentwicklung als Güteindikator verwendet.

Im vorliegende Beitrag wurde die frühe mathematische Bildung als Beispiel zur Erläuterung und Anwendung beider Beurteilungsansätze genutzt. Begonnen wird mit einer Beschreibung des strukturell-prozessualen Qualitätsansatzes sowie seinen Herausforderungen (Abschnitt 2.1). Der Abschnitt 2.2 erläutert das Konzept der Bildungseffektivität und stellt verschiedene Herangehensweisen zur Bestimmung von Bildungseffektivität vor. Anschließend werden im Beitrag eigene Effektivitätsanalysen zur Studie Bildungsprozesse, Kompetenzent-

wicklung und Selektionsentscheidungen im Vor- und Grundschulalter (BiKS; vgl. www.biks-bamberg.de; von Maurice et al., 2007; Abschnitte 3-5) beschrieben. Anhand einer Diskussion der eigenen Ergebnisse reflektiert der Abschnitt 6 Nutzen und die Grenzen des Effektivitätsansatzes für die frühkindliche Bildungsforschung und Praxis.

2. Gegenüberstellung von Qualitäts- und Effektivitätsansatz

2.1 Der strukturell-prozessuale Qualitätsansatz

Qualität ermisst sich im Allgemeinen an dem Grad, in dem ein Angebot bestehenden Anforderungen entspricht (Deutsches Institut für Normung [DIN], 2005). Eine Beurteilung frühkindlicher Bildungsangebote erfolgt anhand von Kriterien, welche sich aus Anforderungen unterschiedlicher Interessengruppen ergeben, z. B. Kinder, Fachkräfte, politische Entscheidungsträger oder Eltern (Katz, 1996). In Wissenschaft und Praxis haben sich Modelle pädagogischer Qualität etabliert, in denen die Entwicklungsförderung und das Wohlbefinden der Kinder im Zentrum (Roux & Tietze, 2007) stehen, so auch im strukturell-prozessualen Ansatz. Qualität ist gemäß diesem Ansatz ein mehrdimensionales Konstrukt (Kluczniok & Roßbach, 2014; Pianta et al., 2005). Die Dimension Strukturqualität umfasst politisch regulierbare Aspekte wie die Größe der Gruppen und Einrichtungen, die räumlich-materielle Ausstattung sowie weitere institutionelle Rahmenbedingungen. Die Prozessqualität bildet die Art der Interaktionen zwischen Fachkräften und Kindern und der Kinder untereinander sowie die Auseinandersetzung des Kindes mit Raum und Materialien ab. Hierunter werden sowohl globale Charakteristika verstanden (*globale Prozessqualität*), z. B. ein dem Entwicklungsstand angemessenes Verhalten und ein warmes Klima (La Paro, Thomason, Lower, Kintner-Duffy & Cassidy, 2012; Pianta et al., 2005; Tietze, Schuster, Grenner & Roßbach, 2005) als auch die Anregung in spezifischen Bildungsbereichen wie Literacy, frühe Mathematik oder Naturwissenschaften (*bereichsspezifische Prozessqualität*; Halle, Vick Whittaker & Anderson, 2010; Kuger & Kluczniok, 2008; Sylva et al., 2006). Aus heutiger Sicht kommt der Prozessqualität eine Schlüsselfunktion zu, weil sie die unmittelbaren Erfahrungen und Anregungsmöglichkeiten des pädagogischen Kontextes charakterisiert und sich somit im Gegensatz zu anderen Qualitätsmerkmalen direkt auf Lernprozesse auswirken kann (Kuger & Kluczniok, 2008). Der Qualitätsansatz ist ein wichtiger Bestandteil des Monitorings und der Steuerung des Systems frühkindlicher Bildung und Betreuung (Roux & Tietze, 2007) und hat wesentlich zu dessen Weiterentwicklung beigetragen (Siraj-Blatchford, Sammons, Taggart, Sylva & Melhuish, 2006).

Der Ansatz hat zudem zahlreiche Forschungsvorhaben angeregt (vgl. Anders, 2013), beispielsweise wurden in Deutschland in der Längsschnittstudie BiKS-3-10 seit 2005 die pädagogische Qualität von Kindergärten erhoben sowie kindliche Kompetenzen, Merkmale des familialen Hintergrunds und der Kinder erfasst. Die Resultate von BiKS zeigten (geringe) Unterschiede in der globalen und bereichsspezifischen Prozessqualität zwischen Einrichtungen (Kuger & Kluczniok, 2008). In latenten Wachstumsmodellen ließen sich Zusammenhänge der Prozessqualität, v. a. der bereichsspezifischen, zur Entwicklung mathematischer Kompetenz im der Kindergartenzeit (Anders et al., 2012) und noch bis zur dritten Klasse nachweisen (Anders, Grosse, Roßbach, Ebert & Weinert, 2013; Lehl, Kluczniok & Roßbach, 2016). Die Effekte der Qualität auf die kindliche Entwicklung waren klein, aber statistisch bedeutsam. Kontrolliert wurden in den Analysen individuelle Merkmale der Kinder (z. B. Alter), des familialen Hintergrunds (z. B. der SES für sozioökonomischer Status) sowie der Qualität der familialen Anregung (HLE für Home Learning Environment), welche sich als bedeutsam für die Entwicklung mathematischer Kompetenz erwiesen haben (Sammons et al., 2002; Schneider, Küspert & Krajewski, 2013; Sylva, Melhuish, Sammons, Siraj-Blatchford & Taggart, 2010). Kompensatorische Effekte, d. h. dass bildungsbenachteiligte Kinder besonders von bereichsspezifischer Prozessqualität profitierten, wurden in BiKS nicht gefunden: Kinder benötigten zumindest eine moderate familiale Anregung zu Hause um zu profitieren (Anders et al., 2012; Lehl et al., 2014). In BiKS gab es keine Evidenz für einen herkunftsbedingten selektiven Zugang zu Kindergärten besonderer bereichsspezifischer Prozessqualität (Lehl et al., 2014): Bildungsbenachteiligte Kinder (u. a. niedriger SES, niedriges mütterliches Bildungsniveau, niedriges HLE und Migrationshintergrund) befinden sich in Kindergärten vergleichbarer bereichsspezifischer Prozessqualität. Aus anderen deutsche Studien wurde allerdings berichtet, dass Kinder in Deutschland Einrichtungen besuchten, in denen die Mehrheit der Kinder ähnlich hinsichtlich entwicklungsrelevanter Merkmale wie elterlichem Bildungsabschluss und Migrationshintergrund sind (Becker & Schober, 2015; Becker, 2010). Die Gruppierung bildungsbenachteiligter Kinder in bestimmten Einrichtungen (hier: Migrationshintergrund) erwies sich in BiKS als ungünstig für die Prozessqualität (Kuger & Kluczniok, 2008; Kuger, Kluczniok, Kaplan & Roßbach, 2015).

In andere Studien wurden ebenfalls Zusammenhänge der Prozessqualität, v. a. der bereichsspezifischen, mit der mathematischen Entwicklung beobachtet (vgl. Anders, 2013). Die englische Studie Effective Pre-School, Primary and Secondary Education (EPPSE) gilt als größte und bedeutsamste europäische Längsschnittstudie, in welcher der Einfluss frühkindli-

cher und schulischer Bildungserfahrungen auf die kindliche Entwicklung untersucht wurde. Auswirkungen der Prozessqualität der frühkindlichen Einrichtungen auf die mathematische Entwicklung konnten bis zu einem Alter von 16 Jahren nachgewiesen werden (Melhuish, 2001; Sammons et al., 2014; Sylva et al., 2010). In der EPPSE-Studie zeigte sich jedoch, anders als in BiKS, dass gerade benachteiligte Kinder von einer hohen Qualität profitierten (u. a. niedriger elterlicher Bildungsabschluss und niedriges HLE; Sammons et al., 2014; Sylva et al., 2010). Für EPPSE wurde überdies ein Zusammenhang zwischen der Zusammensetzung in den Einrichtungen hinsichtlich mütterlichem Bildungsabschluss und mathematischer Kompetenz berichtet (Sammons et al., 2002).

Vor dem Hintergrund der Befunde erscheint das allgemein nur mittelmäßige Niveau der (v. a. bereichsspezifischen) Prozessqualität deutscher Einrichtungen problematisch, wie aus BiKS (Kuger & Kluczniok, 2008) aber auch aus der bundesweite Erhebung Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit (NUBBEK; Tietze et al., 2013) ersichtlich wird. Die durchschnittliche Qualität fiel in EPPSE etwas höher aus, auch zeigte sich eine größere Varianz in der Qualität. Insgesamt befindet sich jedoch in den meisten Ländern die Prozessqualität oftmals nur auf einem mittelmäßigen Durchschnittsniveau (vgl. Clifford et al., 2010). Generell steht der Qualitätsansatz, obschon er sich in Forschung und Praxis etabliert hat, vor wichtigen konzeptuellen und methodologischen Herausforderungen.

2.1.1 Konzeptuelle Herausforderungen.

Eine erste konzeptuelle Herausforderung ergibt sich daraus, dass Qualität ein *normatives Konstrukt* darstellt, weil es die Vorstellungen und Wertüberzeugungen von Interessengruppen widerspiegelt (Roux & Tietze, 2007). Durch die Wertverankerung können Vorstellungen über Qualität *kultur-, system- und kontextspezifisch* unterschiedlich ausfallen (Moss & Pence, 1994; Siraj-Blatchford et al., 2006). Außerdem kommt es mit dem Wandel von Ansprüchen, Rahmenbedingungen und neuen Erkenntnissen über kindliche Entwicklung zu einem *veränderten Verständnis von Qualität über die Zeit*. Beispielsweise wurden in klassischen Modellen Strukturmerkmale betont, in aktuelle Modellen hingegen wird hingegen stärker auf Prozessmerkmale fokussiert. Zeitlich, kulturell und kontextuell unterschiedliche Qualitätsauffassungen können auch zu Einschränkungen in der Vergleichbarkeit von Qualitätsmessungen führen.

Eine weitere konzeptuelle Herausforderung ergibt sich aus dem Erfordernis der *gewachsenen Bedeutung der Merkmale bereichsspezifischer Förderung* Rechnung zu tragen (Hamre, Hatfield, Pianta & Jamil, 2014; Kuger & Kluczniok, 2008), z. B. in Mathematik die Verfügbarkeit von Materialien oder anregende Erzieher-Kind-Interaktionen mit mathematischem Ge-

halt abzubilden. In vielen Messkonzepten sind diese bislang unzureichend repräsentiert und stellen sich als schwer messbar heraus aufgrund der geringeren Strukturierung des frühpädagogischen Alltags im Vergleich zur Schule. So finden Bildungsangebote selten zu festgelegten Zeiten statt, sondern richten sich nach dem pädagogischen Alltag und den kindlichen Interessen.

2.1.2 Methodologische Herausforderungen.

Aus methodologischer Sicht stellte die *aufwändige Erhebungsprozedur, die mit dem Qualitätsansatz verbunden ist*, eine Herausforderung dar. Zahlreiche Kriterien und deren Niveaus müssen auf verschiedenen Ebenen (Fachkräfte, Eltern, Kinder und Leitung) erfasst werden (Katz, 1996). Im Allgemeinen wird sich bei der Erhebung auf zentrale Merkmale beschränkt. Die Einschätzung einiger Merkmale ist vergleichsweise einfach, z. B. können Personal- und Gruppenstruktur oft direkt den Einrichtungskonzeptionen entnommen werden. Die Operationalisierung der Prozessqualität(en) erfordert allerdings mehrstündige Beobachtungen durch geschulte Beobachter und anschließende Interviews mit Fachkräften (Halle et al., 2010; Sylva et al., 2006; Tietze et al., 2005). Neuerdings gibt es Forschungsvorhaben, in denen ökonomischere, fragebogen- und interviewbasierte Varianten erprobt wurden (z. B. Bäumer, Aßmann, von Maurice & Blossfeld, 2013; Slot, Mulder, Verhagen & Leseman, 2014).

Beobachtungen beziehen sich zumeist lediglich auf einen Ausschnitt von wenigen Stunden aus dem pädagogischen Geschehen der Einrichtungen. Hierbei ist in Frage zu stellen, *inwieweit diese Zeitstichprobe als hinreichend repräsentativ für das pädagogische Gesamtgeschehen* angesehen werden kann. Eine Untersuchung schulischer Instruktionsqualität kam zum Schluss, dass Einschätzungen der Dimensionen Klassenführung und unterstützendes Klassenklima mit der Beobachtung einer Stunde, Einschätzungen zur förderrelevanten Dimension kognitive Aktivierung jedoch erst ab neun Unterrichtsstunden verlässlich sind (Praetorius, Pauli, Reusser, Rakoczy & Klieme, 2014). Kane und Staiger (2012) empfehlen auf Grundlage einer Untersuchung zur Unterrichtsqualität von 1 333 Lehrkräften mehr als vier Unterrichtsstunden mittels Beobachtungsverfahren auszuwerten und die Einschätzungen mit Effektivitätsmaßen zu kombinieren. Vermutlich erschwert die vergleichsweise geringe Strukturiertheit der Lehr-Lerngelegenheit des frühkindlichen pädagogischen Alltags die Einschätzung prozessualer Merkmale zusätzlich. Für die globale Prozessqualität in frühkindlichen Einrichtungen zeigte sich eine hohe Stabilität über kurze und eine befriedigende Stabilität über längere Zeiträume innerhalb eines Jahres (Clifford, 2005; Clifford, Reszka & Roßbach, 2010; Pianta & Hamre, 2009; Tietze et al., 2005). Die Stabilität der bereichsspezifischen Prozessqualität fiel in

drei aufeinanderfolgenden Jahren in BiKS allerdings geringer aus als die Stabilität der globalen Prozessqualität (Kuger et al., 2015). Darüber hinaus sind die meisten Qualitätsmessverfahren gruppenbezogen konzipiert. Inwieweit die *Prozessqualität einer Gruppe als repräsentativ für die Qualität der gesamten Einrichtung* angesehen werden kann, ist ebenfalls zu diskutieren. Je nach Fragestellung der Untersuchung kann die unzureichend geklärte Frage der Repräsentativität problematisch sein, z. B. wenn in Studien Beobachtungen in einer Gruppe verwendet werden, um Aussagen über die Qualität der gesamten Einrichtung zu treffen (z. B. EPPSE; Melhuish et al., 2013; Sylva et al., 2006). Ebenfalls wird in vielen Längsschnittstudien, auch EPPSE, in denen Qualitätseffekte auf die kindliche Entwicklung untersucht werden, die Prozessqualität lediglich zu einem Messzeitpunkt erhoben und damit implizit von der Repräsentativität der einmaligen Erhebung für den gesamten Betreuungszeitraum ausgegangen (z. B. Slot et al., 2014; Sylva et al., 2006). Die oftmals geringe Vorhersagekraft, gerade der globalen Prozessqualität, für die kindliche Entwicklung lässt allerdings an der Repräsentativität zweifeln (z. B. Anders et al., 2012; Sylva et al., 2006). So ergaben sich in BiKS und EPPSE nur geringe Zusammenhänge zwischen globaler Prozessqualität und mathematischer Entwicklung (Anders et al., 2012; Sylva et al., 2006). Für die weitere Entwicklung mathematischer Kompetenz wurden in Studien ausschließlich Ergebnisse für bereichsspezifische Maße berichtet (Anders et al., 2013; Lehrl et al., 2016; Sammons et al., 2014).

Neben der *mangelnden Vorhersagekraft ist die eingeschränkte Differenzierungsfähigkeit relevanter Qualitätsmerkmale* eine Herausforderung. So können politische Regulierungen der Strukturqualität zur Vereinheitlichung führen, wenn gewisse Standards vorgeschrieben werden, z. B. für den Erzieher-Kind-Schlüssel (Viernickel & Schwarz, 2009). Die niedrigen Werte und die eingeschränkte Varianz der Prozessqualität, gerade für bereichsspezifische Merkmale (Clifford et al., 2010; Kuger & Kluczniok, 2008; Tietze et al., 2013), erschweren den Nachweis konsistenter und langanhaltender Zusammenhänge zur kindlichen Entwicklung. Um den Ansprüchen empirischer Bildungsforschung zu genügen, werden Nachweise über Zusammenhänge der Prozessqualität zur kindlichen Entwicklung und Nachweise der für das jeweilige Prozessmaß angenommenen dimensional Struktur der Prozessqualität benötigt. So sollten Annahmen über die Gruppierung der einzelnen Qualitätsitems zu Skalen und ggfs. zu einem Globalwert überprüft werden. Dabei sollten beispielsweise empirisch ermittelte Faktoren und Faktorladungen der Items die Annahmen widerspiegeln. In der Mehrzahl der Studien zu etablierten *Maßen der Prozessqualität weichen empirisch ermittelte und angenommene Struktur*

voneinander ab (Clifford et al., 2010; Gordon, Fujimoto, Kaestner, Korenman & Abner, 2013; Hamre et al., 2014; Pianta & Hamre, 2009).

Mit der gestiegenen Aufmerksamkeit für mögliche Potenziale frühkindlicher Bildungseinrichtungen für die Förderung bereichsspezifischer Kompetenzen steigt das Interesse an ergänzenden Verfahren der Beurteilung von Einrichtungen (vgl. Anders, 2013). In der Schulforschung und -evaluation wird national (Bischof, Hochweber, Hartig & Klieme, 2013) und international (Chapman, Muijs, Reynolds, Sammons & Teddlie, 2015) der Effektivitätsansatz angewendet. EPPSE ist bislang die einzige Studie im frühkindlichen Bereich, in welcher neben Qualitäts- auch der Effektivitätsansatz verwendetet wurde (vgl. Melhuish, 2001), welcher im Folgenden erläutert wird.

2.2 Der Effektivitätsansatz

Im Ansatz der Bildungseffektivität werden Einrichtungen oder Gruppen gemäß Kompetenzstand bzw. -entwicklung der Kinder beurteilt (Sylva et al., 2010). Der Indikator für die Güte des Bildungsangebots ist demnach der „Bildungoutput“, beispielsweise die mittlere Leistung bzw. der Leistungszuwachs an mathematischer Kompetenz (Bischof et al., 2013). Die Förderung früher Fähigkeiten in schulfachbezogenen Entwicklungsbereichen, wie frühe mathematische, schriftsprachliche oder naturwissenschaftliche Kompetenzen, gilt in gängigen Qualitätsmodellen als zentral (Roux & Tietze, 2007) und einige nationale (Viernickel & Schwarz, 2009) und internationale Modelle (Organisation for Economic Co-operation and Development [OECD], 2006) führen Bildungsergebnisse explizit als Qualitätsmerkmal an. Daraus ergibt sich für Forschung und Praxis, dass das Ausmaß der Förderung kindlicher Kompetenzen als Qualitätsmerkmal von Einrichtungen herangezogen, erfasst und verbessert werden muss. So sollten Forschungsprojekte oder Qualitätsmonitorings Erhebungen kindlicher Kompetenzen beinhalten. Die Ergebnisse aus Effektivitätsanalysen lassen ausschließlich Aussagen zur Güte der Förderung in den Bildungsbereichen zu, für die der Bildungoutput ausgewertet wird (z. B. mathematische oder sprachliche Kompetenzen). Merkmale der Einrichtungen, die aus ethischen oder normativen Gesichtspunkten relevant sind, bilden Effektivitäts- anders als Qualitätswerte nicht ab.

Die EPPSE-Befunde zeigten jedoch, dass der Effektivitätsansatz einen aussagekräftigen, zusätzlichen Indikator liefert, um die Güte frühkindlicher Betreuungserfahrung zu erfassen. Es ergaben sich signifikante Unterschiede zwischen Einrichtungen in verschiedenen Effektivitätsanalysen (siehe Kapitel 2.2) und für verschiedene Bildungsbereiche, z. B. Mathematik. Kinder effektiverer Einrichtungen wiesen höhere mathematische Kompetenzen in der Grundschule

auf, wobei benachteiligte Kinder (niedriges HLE) besonders von hoher Effektivität profitierten (Sylva et al., 2010). Die Vorteile ließen sich auch noch mit 16 Jahren nachweisen (Sammons et al., 2014).

Der Effektivitätsansatz ist lernzielorientiert (Creemers, Kyriakidēs & Sammons, 2010) und berücksichtigt nicht, unter welchen Bedingungen die Kinder lernen und wie ihre Lernerfahrungen aussehen. Es ist aber möglich, bestimmende Merkmale für Effektivitätsunterschiede zu untersuchen, z. B. mit vergleichenden Fallanalysen von sehr und wenig effektiven Einrichtungen (Racherbäumer, Funke, van Ackeren & Clausen, 2013; Siraj-Blatchford et al., 2006). Dadurch können weitere Einrichtungsmerkmale identifiziert werden, die eine bessere Bildungsarbeit bedingen oder charakterisieren, um so Qualitätsmodelle weiterzuentwickeln. Beispielsweise konnte in einer qualitativen EPPSE-Zusatzstudie Unterschiede in Interaktionsstilen zwischen wenig effektiven und sehr effektiven Einrichtungen festgestellt werden (Siraj-Blatchford et al., 2006). Die Befunde dienten als Ausgangspunkt für die Entwicklung einer neuen Qualitätsskala (Siraj-Blatchford, Kingston & Melhuish, 2015).

Effektivitätsbeurteilungen erfolgen relational, weil Ausgangspunkt von Effektivitätsanalysen immer Vergleiche von Einrichtungen sind. Für eine einzelne Einrichtung bedeutet dies, dass beurteilt wird, inwiefern der Kompetenzstand bzw. -zuwachs höher oder niedriger als in anderen Einrichtungen ausfällt (Chapman et al., 2015; Creemers et al., 2010). Während im Qualitätsansatz normative Standards für Kriterien festgelegt sind, basieren Effektivitätsurteile also auf den Verteilungseigenschaften der Einrichtungsstichprobe. Der Ansatz erlaubt daher keine Aussagen darüber, ob das mittlere, erzielte Kompetenzniveau in den Einrichtungen zufriedenstellend ist.

2.2.1 Voraussetzungen für die Ermittlung von Bildungseffektivität.

Grundlage aller Effektivitätsanalysen stellen Angaben zu kindlichen Kompetenzen dar; vorzugsweise Längsschnittdaten (Creemers et al., 2010; Reynolds et al., 2014). Die Daten sollten die hierarchische Struktur des Systems (Kinder gruppieren sich in Gruppen, die in Einrichtungen organisiert sind) reflektieren. Um reliable und valide Effektivitätsaussagen treffen zu können, werden ausreichend große und repräsentative Stichproben benötigt, also Kinderstichproben pro Einrichtung und Gesamtstichprobe der Einrichtungen. Über die Stichprobengrößen existieren unterschiedliche Aussagen, jedoch scheint für die Aussagekraft zu Effektivitätsunterschieden die Größe der Einrichtungsstichprobe wichtig zu sein (Maas & Hox, 2005). Bei sehr wenigen Kindern einer Einrichtung steigt jedoch die Fehleranfälligkeit des Effektivitätsmaßes für die jeweilige Einrichtung (Ballou & Springer, 2015).

2.2.2 Herangehensweisen zur Bestimmung von Bildungseffektivität.

Es gibt unterschiedliche Möglichkeiten Bildungseffektivität zu ermitteln (Creemers et al., 2010; Sammons et al., 2002), z. B. (a) *Statusvergleiche*. Bei diesen Vergleichen gilt eine Einrichtung als effektiv, in der das Niveau kindlicher Kompetenzen zu einem gegebenen Zeitpunkt höher ausfällt als das durchschnittliche Kompetenzniveau in der Einrichtungsstichprobe. Laut Scheerens (1992) sind solche Vergleiche jedoch „unfair“ (vgl. Racherbäumer et al., 2013), weil Kinder bereits vor Eintritt in die Einrichtungen unterschiedliche Kompetenzniveaus aufweisen und sich unterschiedliche Ausgangsniveaus u. U. nicht gleichmäßig auf Einrichtungen verteilen. Einrichtungen, in denen sich von Beginn an viele Kinder mit hohem Niveau befinden, haben bei unkontrollierten Vergleichen Vorteile.

Für eine „faire“ Beurteilung sind zumindest (b) *Zuwachsvergleiche* notwendig, die unterschiedliche Ausgangsniveaus in die Analysen einbeziehen. Qualitätskriterium ist der Kompetenzzuwachs in Einrichtungen über einen gewissen Zeitraum. Zur Anwendung kommen sogenannte Value-Added-Modelle (VAM; Racherbäumer et al., 2013): Effektiv ist eine Einrichtung, wenn sie Kompetenzzuwächse vorweisen kann, die über dem über dem durchschnittlichen Zuwachs in der Einrichtungsstichprobe liegen.

Kontextualisierte Vergleiche (CAM für Contextualized Attainment Models bzw. CVA für Contextualized Value-Added Models; vgl. Creemers et al., 2010; Reynolds et al., 2014) berücksichtigen, dass sich Kinder in Einrichtungen in Merkmalen unterscheiden, die mit kindlichen Kompetenzen bzw. Kompetenzzuwächsen zusammenhängen (z. B. Alter, Bildungsniveau der Eltern oder SES; Anders, 2013; Anders et al., 2012; Schneider et al., 2013). Kontextualisierte Vergleiche des Status (c) oder des Zuwachses (d) berücksichtigen die unterschiedlichen Zusammensetzungen in Einrichtungen. Eine effektive Einrichtung zeichnet sich durch ein höheres Niveau oder stärkeren Zuwachs kindlicher Kompetenzen aus, als für eine Einrichtung mit dieser Zusammensetzung an entwicklungsrelevanten Merkmale zu erwarten gewesen wäre.

Weitere Effektivitätsanalysen, z. B. die Betrachtung von Veränderungen in der Rangordnungen der Einrichtungen über den zeitlichen Verlauf (Creemers et al., 2010), basieren auf den hier skizzierten Herangehensweisen. Effektivitätsanalysen erfolgen durch Mehrebenenanalysen (Chapman et al., 2015; Creemers et al., 2010). Effektivitätsunterschiede zeigen sich durch Varianzen auf der Einrichtungsebene, die sich in hohen Intraklassenkorrelationen niederschlagen (*ICCs*). Ergeben sich signifikante Unterschiede, bildet das Residuum dieser Ebene für die jeweilige Einrichtung den Effektivitätswert (Sammons et al., 2002; Sylva et al., 2010).

3. Untersuchung der Bildungseffektivität in Mathematik in der BiKS-3-10-Studie

In BiKS-3-10 (von Maurice et al., 2007) wurden kindliche Kompetenzen in verschiedenen Bereichen in jährlichen Abständen erhoben. Es liegen Daten zu Teilbereichen mathematischer Kompetenz vor, die als besonders prädiktiv für die weitere mathematische Entwicklung gelten (Jordan, Kaplan, Ramineni & Locuniak, 2009; Schneider et al., 2013), sowie zu Merkmalen, die sich als relevant für die mathematische Entwicklung erwiesen haben und daher bei der Untersuchung institutioneller Einflüsse zu berücksichtigen sind (z. B. Anders et al., 2012; 2013; Lehl et al. 2014; 2016; Sylva et al., 2010). Das Studiendesign ist somit geeignet, um Bildungseffektivität mit den vier vorgestellten Herangehensweisen a-d zu bestimmen und ermöglicht einen Vergleich zu bereits publizierten Befunden des Qualitätsansatzes. Die Analysen bauen direkt auf Vorgängeranalysen auf, insbesondere auf den Arbeiten von Anders und Kollegen (2012), die Evidenz zu den Qualitätsmerkmalen der BiKS-Einrichtungen lieferten. Ergänzt werden nun Analysen zur mathematischen Bildungseffektivität. Dadurch liegt der Fokus der Auswertung erstmalig direkt auf der Mehrebenenstruktur von BiKS, die bisher lediglich indirekt in Analysen berücksichtigt wurde.

Forschungsschwerpunkt des Beitrags sind potentielle Kompetenzunterschiede auf Einrichtungsebene, die auf Effektivitätsunterschiede hindeuten (Fragestellung 1). Die mehrebenenanalytischen Ergebnisse zur Bedeutung einzelner entwicklungsrelevanter Merkmale (Fragestellung 2) können wegweisend für die Auswahl von Merkmalen in zukünftigen Effektivitätsanalysen anderer Studien sein, in denen diese Merkmale ggfs. nur teilweise oder in anderer Form enthalten sind. Zusätzlich ermöglicht das Design die Effektivität zu verschiedenen MZP zu vergleichen und Veränderungen in der Effektivität und der Relevanz individueller Merkmale für die Vergleiche zu betrachten (Chapman et al., 2015). In dem Beitrag wurden folgende Fragestellungen zur (1) Bildungseffektivität und zur (2) Bedeutung entwicklungsrelevanter Merkmale in BiKS untersucht:

1. Unterscheiden sich die Einrichtungen (a) im mittleren Leistungsniveau und/oder (b) im mittleren Leistungszuwachs elementarer mathematischer Kompetenzen ihrer Kinder? Unterscheiden sich die Einrichtungen bei (c) kontextualisierten Vergleichen des Status und/oder (d) Zuwachs, d. h. wenn entwicklungsrelevante Merkmale kontrolliert werden?

2. Welche Bedeutung haben die Merkmale der Kinder und des familialen Hintergrundes bei der Ermittlung von Bildungseffektivität(en) in Mathematik, d. h. welche Rolle spielen verschiedene entwicklungsrelevante Merkmale als Prädiktoren auf Kindebene bei kontextualisierten Vergleichen?

4. Methode

4.1 Datengrundlage und Stichprobe

In BiKS-3-10 wurden kindliche Kompetenzen in 97 Kindergärten erhoben (60 bayerische und 37 hessische Einrichtungen). Das Sampling der Kinder erfolgte gruppen- und altersbasiert, denn an der Studie nahm jeweils nur eine Gruppe pro Einrichtung teil. Die Wahl fiel auf die Gruppe mit den meisten Kindern im Zielalter, d. h. Kinder die bei regulärer Einschulung im Schuljahr 2008/2009 schulpflichtig wurden. Die Stichprobenziehung orientierte sich an den bundeslandspezifischen Einschulungsterminen (Geburtsfenster in Bayern: 01.10.2001–31.10.2002 und in Hessen: 01.07.2001–30.06.2002). Die Stichprobe umfasste insgesamt 554 Kinder (48 % Mädchen), wobei im Durchschnitt sechs Kinder pro Gruppe getestet wurden (durchschnittliche Clustergröße: 5.71; Spannweite: 4 - 10)ⁱ. Die Mehrheit der Stichprobe stammte aus einem Elternhaus, in dem beide Elternteile Deutsch als Muttersprache sprachen. 10 % bzw. 12 % der Kinder hatten ein Elternteil bzw. zwei Elternteile, die nicht Deutsch als Muttersprache sprachen. Von den Müttern hatten 34 % die Schule mit der mittleren Reife, 34 % mit Abitur und 25 % ohne Abschluss oder mit einem Hauptschulabschluss abgeschlossen. 7 % der Mütter besaßen einen sonstigen, zumeist im Ausland erworbenen Abschluss.

In die Vergleiche zur Effektivität flossen die Testergebnisse elementarer mathematischer Kompetenzen zu drei MZP ein, in denen die Kinder sich in den Einrichtungen befanden. Die Kinder waren zum ersten MZP drei Jahre ($M = 44.55$, $SD = 5.00$), zum zweiten MZP vier Jahre ($M = 55.73$, $SD = 4.47$) und zum dritten MZP fünf Jahre ($M = 67.43$, $SD = 4.25$) alt. Bei Eintritt in die Einrichtungen waren die Kinder ungefähr drei Jahre alt ($M = 37.54$, $SD = 5.21$). Die MZP fielen somit für die Mehrheit in das erste, zweite und dritte Kindergartenjahr.

4.2 Instrumente

4.2.1 Frühe mathematische Kompetenzen.

Zu allen drei MZP wurden die elementaren mathematischen Kompetenzen mit dem Unter-test Rechnen der deutschen Version der Kaufman-Assessment Battery for Children (K-ABC, Melchers & Preuß, 2003) erfasst. Der Test misst elementare numerische und geometrische Kompetenzen von Kindern: ihre Zählfertigkeiten, ihre Ziffern- und Formkenntnis, ihr grundlegendes Mengenverständnis sowie basale Rechenfertigkeiten. Er erfasst damit solche Fähigkeiten, die als sehr prädiktiv für Mathematikleistungen in der Schule gelten (Jordan et al., 2009; Schneider et al., 2013). Die vorliegenden Analysen verwendeten die Summenwerte für jeden MZP. Über die drei MZP stieg der Mittelwert in der Stichprobe an, d. h. mit zunehmendem Alter erreichten die Kinder höhere Testwerte: zum 1. MZP mit etwa drei Jahren lag der Mit-

telwert bei 4.90 ($SD = 3.38$), zum 2. MZP mit etwa vier Jahren bei 10.26 ($SD = 3.98$) und zum 3. MZP mit etwa fünf Jahren bei 15.03 ($SD = 3.70$).

4.2.2 Entwicklungsrelevante Merkmale.

Die kontextualisierten Vergleiche verwendeten außerdem solche in BiKS erfassten individuellen Kind- und Familienmerkmale, die sich als bedeutsam für die Entwicklung mathematischer Kompetenz erwiesen haben (Anders et al., 2012; 2013; Lehl et al., 2014; vgl. Sammons et al., 2002; Schneider et al., 2013). Die Auswahl und Annahmen zur Bedeutung der Merkmale basierten unmittelbar auf die umfangreichen BiKS-Voranalysen und publizierten Interpretationen der Ergebnisse. Ziel war dabei einerseits möglichst viele entwicklungsrelevante Merkmale zu kontrollieren, um einen hiervon unabhängigen Effekt der Bildungsqualität abzubilden. Gleichzeitig wurde ein möglichst sparsames Modell angestrebt, das die komplexen Wachstumsmodelle ermöglichte. Durch den engen Bezug zu bisherigen Publikationen ermöglicht der vorliegende Beitrag die Anschlussfähigkeit und Vergleichbarkeit mit den Vorgängeranalysen.

Es flossen das Alter zum jeweiligen MZP, das Alter bei Eintritt in die Einrichtung und das Geschlecht ein. Zu den verwendeten Merkmalen des familialen Hintergrundes zählte neben der Muttersprache der Eltern und dem mütterlichen Bildungsabschluss der Highest International Socio-Economic Index (HISEI, $M = 52.29$, $SD = 16.43$) als Indikator des SES der Familie, welcher auf der höchsten Klassifikation elterlicher Berufe basiert (vgl. Ganzeboom, de Graaf & Treiman, 1992). Außerdem berücksichtigten die Analysen die häusliche Lernumwelt (HLE) als Mittelwert über drei MZP, der auf einen Wertebereich von 0 bis 1 standardisiert wurde ($M = .46$, $SD = .13$). Die HLE-Skala spiegelt die elterliche Anregung in schriftsprachlichen und mathematischen Bereichen wieder. Die Messung basiert auf (a) Beobachtung mit der Home Observation for Measurement of the Environment (HOME; Caldwell & Bradley, 1984), (b) Einschätzungen zur halbstandardisierten elterlichen Vorlesesituation mit der Family Rating Scale (FES; Kuger, Pflieger & Roßbach, 2005) sowie (c) Angaben zu elterlichen Aktivitäten aus selbst-konstruierten Fragebogen- bzw. Interviewitems (vgl. Anders et al., 2012).

4.3 Statistische Auswertung

Die Effektivitätsanalysen nutzten die hierarchische Struktur der Daten: Kinder (Ebene 1: Kindebene) gruppieren sich in Einrichtungen bzw. Gruppen (Ebene 2: Einrichtungsebene). Die Vergleiche wurden mit Random-Intercept-Modellen mit Hilfe des Statistikprogramms *Mplus* angestellt (type = twolevel, Cluster = Einrichtung). Alle Analysen verwendeten einen Maximum-Likelihood-Schätzer mit robusten Standardfehlern (MLR), der Teststatistiken und Standardfehler für die Clusterung und fehlende Normalverteilung von Daten korrigiert. Fehlende

Werte wurden mit der Full Information Maximum Likelihood-(FIML)-Methode geschätzt (Lüdtke, Robitzsch, Trautwein & Köller, 2007), die eine modellbasierte Schätzmethode darstellt. Abhängig von den im Modell enthaltenen Variablen kommt es daher zu abweichenden Fallzahlen in Analysen. Mit verschiedenen Random-Intercepts-Modellen wurden die vier beschriebenen Effektivitätsvergleiche umgesetzt: (a) *Statusvergleiche* zu drei MZP, (b) *Zuwachsvvergleiche* zum zweiten und dritten MZP sowie *kontextualisierte Vergleiche* (c und d). Die kontextualisierten Vergleiche berücksichtigen die kategorialen Merkmale Geschlecht, Muttersprache der Eltern und mütterlicher Bildungsabschluss mit Dummyvariablen (Referenzkategorien: männlich, beide Deutsch und kein Abschluss oder Hauptschulabschluss).

5. Ergebnisse

5.1 Effektivitätsunterschiede im Bildungsbereich Mathematik in der BiKS-Studie

Tabelle 1 bietet eine Zusammenfassung der Ergebnisse für die vier Vergleiche (a – d) wie den Anteil der Varianz auf Einrichtungsebene (*ICC*) zu den drei MZP, bei denen die Kinder im Durchschnitt drei, vier und fünf Jahre alt waren (erstes bis drittes Jahr in den Einrichtungen). Signifikante *ICCs* deuten an, dass sich Einrichtungen bzw. Gruppen bei diesem Vergleich unterscheiden, so dass die dazugehörigen Effektivitätsscores zwischen Einrichtungen differenzieren und als Güteindikator mathematischer Bildung berechnet werden können. Der Determinationskoeffizient ($R^2_{innerhalb}$) zeigt an, wie viel Varianz in den individuellen Kompetenzen die verschiedenen Merkmale auf Kindebene insgesamt erklärten.

<< Tabelle 1 hier einfügen >>

Ergebnisse der (a) *Statusvergleiche* deuten auf signifikante Unterschiede im mittleren Leistungsniveau der Einrichtungen zu allen MZP. Der Anteil, den die Zugehörigkeit zu unterschiedlichen Gruppen erklärt, nimmt über die Zeit ab: von 21 % zum ersten auf 13 % zum zweiten auf 8 % zum dritten MZP.

Die *ICCs* der (b) *Zuwachsvvergleiche* sprechen dafür, dass sich die Einrichtungen bzw. Gruppen nicht signifikant im mittleren Zuwachs mathematischer Kompetenz unterscheiden (weniger als 4 % Varianz zwischen den Einrichtungen zu beiden MZP). Die Kompetenzen zum früheren MZP erklärt fast die Hälfte der Varianz in den mathematischen Kompetenzen ($R^2_{innerhalb} = 43\%$ zum zweiten und 52% zum dritten MZP).

Kontextualisierte Vergleiche (c und d), welche entwicklungsrelevante Merkmale berücksichtigten, sprechen für Effektivitätsunterschiede zum ersten MZP ($ICC = .08$), aber gegen Unterschiede zum zweiten und dritten MZP in Status- oder Zuwachsvvergleichen (2 % bzw. 1 % Varianz zwischen Einrichtungen). Die individuellen Merkmale hängen signifikant mit der

mathematischen Kompetenz zusammen, wobei der Zusammenhang über die MZP abnimmt (von 33 % zum ersten auf 24 % zum zweiten auf 18 % zum dritten MZP) und die Merkmale insgesamt weniger Varianz in den elementaren mathematischen Kompetenzen erklären als Kompetenzen zu einem früheren MZP (43 % bzw. 52 %). Insgesamt sprechen die Befunde daher gegen Effektivitätsunterschiede in „fairen“ Vergleichen (b–d).

5.2 Die Bedeutung entwicklungsrelevanter Merkmale bei Effektivitätsvergleichen

Tabelle 2 zeigt die unstandardisierten Regressionsgewichte und dazugehörige Standardfehler für die entwicklungsrelevanten Merkmale in den beiden kontextualisierten Vergleichen (c und d). Die Tabelle gibt somit Auskunft über die Höhe der Zusammenhänge zwischen den einzelnen entwicklungsrelevanten Merkmalen und den mathematischen Kompetenzen zu den drei MZP. Ebenfalls lassen sich die Beziehungen zwischen den mathematischen Kompetenzen und der mathematischen Kompetenz zu früheren MZP ablesen.

<< Tabelle 2 hier einfügen >>

Die (c) *Statusvergleiche* zeigen, dass das Eintrittsalter und Geschlecht nur im ersten Jahr (1. MZP) mit den mathematischen Kompetenzen zusammenhängen. Je später die Kinder in die institutionelle Betreuung kamen, desto schlechter schnitten sie im ersten Jahr ab ($B = -0.06$, $\beta = -.10$). Mädchen erlangten im ersten Jahr fast einen Punkt mehr bei der Testung ($B = 0.85$).

Außerdem sind der mütterliche Bildungsabschluss zum zweiten MZP und der SES zum zweiten und dritten MZP signifikante Prädiktoren. Im zweiten Jahr erreichten Kinder, deren Mütter einen sonstigen Abschluss erworben hatten, über eineinhalb Punkt mehr ($B = 1.65$). Je höher der SES, desto besser schnitten Kinder im Test ab ($B = 0.04$, $\beta = .16$ bzw. $.19$).

Zu allen MZP sind das Alter zum MZP, die Muttersprache der Eltern und das HLE signifikante Prädiktoren. Je älter die Kinder waren ($B = 0.25 - 0.33$, $\beta = .31 - .46$), desto besser waren ihre Leistungen im Test. Kinder aus einem Elternhaus, in dem mindestens ein Elternteil nicht Deutsch als Muttersprache sprach, erreichten ein bis zwei Punkte weniger bei den Testungen ($B = -1.37 - -2.04$). Das HLE zeigte als einziges Merkmal einen zunehmenden Zusammenhang zur Kompetenz über die drei MZP ($B = 4.19 - 5.13$, $\beta = .14 - .18$).

Ergebnisse für (d) *Zuwachsvergleiche* verdeutlichen, dass die mathematischen Kompetenzen zu einem früheren MZP stärkster Prädiktor ist, wie der Vergleich der Determinationskoeffizienten bereits zeigte. Der Zusammenhang fällt zum dritten MZP sogar höher aus als zum zweiten MZP ($B = 0.65$, $\beta = .70$ vs. $B = 0.63$, $\beta = .58$), was auf eine Zunahme der Stabilität von Unterschieden im individuellen Leistungsniveau hinweist. Bei Kontrolle der mathemati-

schen Kompetenz zu einem früheren MZP nimmt die Stärke der Zusammenhänge für alle Prädiktoren deutlich ab.

6. Diskussion

Im vorliegende Beitrag wurde das Konzept der Bildungseffektivität als einen ergänzenden Ansatz zur Beurteilung der Güte frühkindlicher Bildungsangebote am Beispiel des Bildungsbereichs Mathematik diskutiert. Im Effektivitätsansatz werden Status oder Zuwachs kindlicher Kompetenzen in Einrichtungen verglichen und dabei ggfs. Zusammenhänge zu entwicklungsrelevanten Merkmalen berücksichtigt (Sammons et al., 2002; Sylva et al., 2010). So erlaubt der Ansatz v. a. Aussagen zur Güte bereichsspezifischer Förderung. Die eigenen Effektivitätsanalysen untersuchten für die BiKS-Studie, ob sich frühkindliche Einrichtungen in Deutschland hinsichtlich mathematischer Bildungseffektivität unterscheiden.

6.1 Interpretation der Befunde zur Bildungseffektivität in der BiKS-Studie

Insgesamt sprechen die Effektivitätsbefunde gegen signifikante Unterschiede in der Güte mathematischer Förderung, die sich nicht durch entwicklungsrelevante Merkmale bzw. unterschiedliche mathematische Ausgangsniveaus erklären ließen. Denn die Analysen zeigten, dass sich zwar in (a) Statusvergleichen signifikante Unterschiede zu allen MZP ergaben, es in „fairen“ Vergleiche (Scheerens, 1992), welche die unterschiedlichen Ausgangskompetenzen (b) oder/und Zusammensetzungen in den Einrichtungen bzw. Gruppen hinsichtlich entwicklungsrelevanter Merkmale berücksichtigten (c und d), keine Effektivitätsunterschiede nach dem ersten MZP feststellbar waren.

Die Resultate spiegeln einerseits die Befunde zur geringen Varianz der bereichsspezifischen Prozessqualität, v. a. in Mathematik, wider (Kuger & Kluczniok, 2008; Tietze et al., 2013). Auf der anderen Seite fanden sich in unterschiedlichen Arbeiten der BiKS-Studie geringe aber signifikante Effekte der beobachteten Prozessqualität auf die Entwicklung mathematischer Fähigkeiten. Dieser Widerspruch lässt sich zum Einen dadurch erklären, dass die beobachtete Prozessqualität im BiKS-Design vermutlich repräsentativer für die Arbeit in der gesamten Gruppe ist als die ermittelten Effektivitätswerte sind, die lediglich auf den Kompetenzdaten von einem Teil der Kinder aus den Gruppen basieren. Ferner wurde in den Arbeiten zu den Effekten der beobachteten Prozessqualität eine Fokussierung auf die individuellen Entwicklungsverläufe der Kinder gelegt und nicht auf die durchschnittlichen Zuwächse in den Gruppen. Analytisch unterschied sich der Umgang mit der Mehrebenenstruktur.

Eine mögliche Erklärung für die gefundenen Unterschiede in Statusvergleichen und zum ersten MZP sind eher Selektionseffekte als Unterschiede in der Bildungsarbeit. Es ist anzu-

nehmen, dass die Einrichtungen bzw. Gruppen von Beginn an nicht vergleichbar in ihrer Zusammensetzung hinsichtlich mathematischer Kompetenz und entwicklungsrelevanter Merkmale der Kinder waren (vgl. Becker, 2010; Becker & Schober, 2015). Statusvergleiche sind dann „unfair“ (Scheerens, 1992), weil die mathematische Bildung in den Einrichtungen unter ungleichen Ausgangsbedingungen stattfand (Creemers et al., 2010). Das bessere Abschneiden mancher Einrichtungen wäre dann durch „günstigere“ Ausgangsbedingungen und nicht durch eine bessere Bildungsarbeit begründetⁱⁱ. Die unterschiedliche Gruppenzusammensetzung hinsichtlich entwicklungsrelevanter Merkmale erscheint vor dem Hintergrund der Auswirkungen auf die Prozessqualität und die kindliche Entwicklung problematisch (Kuger & Kluczniok, 2008; Kuger et al., 2015; Sammons et al., 2002). Die Befunde sprechen dafür, dass für Effektivitätsvergleiche möglichst Zuwachsvergleiche oder kontextualisierten Vergleiche zur Anwendung kommen sollten, insbesondere kurz nach Beginn der Fremdbetreuung, weil diese Vergleiche die Zusammensetzungen zumindest teilweise berücksichtigen.

6.2 Die Bedeutung der entwicklungsrelevanten Merkmale

Die Ergebnisse aus kontextualisierten Statusvergleichen (Modell c in Tabelle 1) zeigen, dass entwicklungsrelevante Merkmale einen beachtlichen Teil der Varianz in den Kompetenzen erklären und somit teilweise die ungleichen Ausgangsbedingungen in Einrichtungen berücksichtigen (vgl. Sammons et al., 2002). Wenn lediglich Querschnittsdaten vorliegen, jedoch Informationen über entwicklungsrelevante Merkmale vorhanden sind, sollten Effektivitätsanalysen daher zumindest kontextualisiert erfolgen, so z. B. für die querschnittliche Erhebung NUBBEK: Die von Leyendecker, Agache und Madsen (2014) berichteten signifikanten ICCs für Kompetenzen in unterschiedlichen Bereichen fielen in mittlerer Höhe aus, aber berücksichtigten keine Unterschiede in entwicklungsrelevanten Merkmalen. Die vorliegenden Befunde zeigen, dass insbesondere das Alter der Kinder zum Messzeitpunkt und das häusliche Anregungsniveau, aber auch Indikatoren der Benachteiligung wie Migrationshintergrund oder der SES in Effektivitätsanalysen einfließen sollten (vgl. Anders et al., 2012; 2013; Lehl et al., 2014, 2016; Sammons et al., 2002; Schneider et al., 2013).

Die Stärke der Zusammenhänge nahm für fast alle Prädiktoren im Laufe der Zeit ab, was für einen wachsenden Einfluss der Betreuungserfahrungen spricht (Anders et al., 2012). Betrachtet man das Muster der Zusammenhänge zum Zuwachs fällt auf, dass bei Berücksichtigung der Kompetenzen zu einem früheren MZP die meisten Prädiktoren ihre Vorhersagekraft verlieren. Vermutlich manifestierten sich Kompetenzunterschiede, welche mit verschiedenen Merkmalen der Kinder und des familialen Hintergrunds zusammenhängen, bereits vor Eintritt

in die Einrichtungen in unterschiedlichen Ausgangsniveaus (vgl. Anders et al., 2012; Schneider et al., 2013).

6.3 Einschränkungen der vorliegenden Untersuchung und Vergleich zur EPPSE-Studie

In der EPPSE-Studie wurden, im Gegensatz zu den hier berichteten Ergebnissen, signifikante Unterschiede für den konservativsten Vergleich, d. h. kontextualisierte Zuwachsvergleiche gefunden (Kontrolle von entwicklungsrelevanten Merkmalen und Kompetenzen zu früherem MZP; Sammons et al., 2002; Sylva et al., 2010). Mögliche Erklärungen liefern Unterschiede im System frühkindlicher Bildung und Erziehung. Denn das englische System ist stärker lernzielorientiert und auf bereichsspezifische Förderung ausgelegt (OECD, 2006). Eine andere Erklärung bieten Unterschiede im Studiendesign.

So sind die *Stichproben* der BiKS-Studie vergleichsweise klein. Durchschnittlich wurden in BiKS sechs und in EPPSE ungefähr 20 Kinder pro Einrichtung getestet. Die Aussagekraft eines Effektivitätswertes als Güteindikator für eine spezifische Einrichtung hängt von der Anzahl an Kindern ab, die aus dieser Einrichtung getestet wurden (Ballou & Springer, 2015), d. h. die Aussagekraft der hier ermittelten Werte ist eingeschränkter. Für die Untersuchung von Effektivitätsunterschieden ist die Stichprobengröße auf Einrichtungs- bzw. Gruppenebene wichtiger (Maas & Hox, 2005). In BiKS liegen Daten aus 97 in EPPSE aus 141 Einrichtungen vor.

Vermutlich ist die *Einrichtungsstichprobe in BiKS homogener* als in EPPSE auch wegen unterschiedlicher Samplingstrategien. BiKS ist keine national repräsentative Studie (von Maurice et al., 2007). Die Einrichtungen entstammen zwei Bundesländern (Bayern und Hessen). Pro Einrichtung nahm nur eine Gruppe teil (gruppenbasiertes Sampling). Damit bezieht sich der Effektivitätswert, wie viele Qualitätsmaße, auf die Gruppe und nicht auf die Einrichtung. Durch das altersbasierte Sampling beschränkt sich der Wert außerdem auf die Kompetenz der Kinder im Zielalter und nicht auf die Kompetenz aller Kinder einer Gruppe. Das altersbasierte Sampling führte außerdem u. U. zu verzerrten Effektivitätsschätzungen, weil bekannt ist, dass in Deutschland bildungsbenachteiligte Familien und solche mit Migrationshintergrund Kinder später in Einrichtungen geben (Geier & Riedel, 2009; Schober & Spieß, 2012). Dementsprechend sind die Repräsentativität der Stichprobe und ggfs. auch des Effektivitätswerts eingeschränkt. Außerdem waren einige Betreuungsformen nicht repräsentiert (z. B. Spielgruppen oder Kindertagespflege). Die EPPSE-Einrichtungen entstammten fünf Regionen Englands und repräsentierten sehr unterschiedliche Betreuungsangebote. Die größere Heterogenität der Einrichtungsstichprobe könnte maßgeblich zu den gefundenen Effektivitätsunter-

schieden und den größeren Varianzen (wie Mittelwerten) der Prozessqualität in der englischen im Vergleich zur deutschen Stichprobe beigetragen haben (Sylva et al., 2006).

Schließlich wurden in BiKS in jährlichem Abstand mit dem gleichen Verfahren (K-ABC Subskala Rechnen) die mathematischen Kompetenzen erhoben, wohingegen sich der Zuwachsvergleich in EPPSE auf ein Intervall von zwei Jahren bezog und als Ausgangsniveau ein Testverfahren verwendete, welches nicht numerische sondern verbale und nonverbale kognitive Kompetenzen erfasste. Wenn auch prädiktiv weisen unspezifische kindliche Merkmale wie das allgemeine kognitive oder sprachliche Leistungsniveau geringere Vorhersagekraft für die weitere mathematische Entwicklung als fachspezifische Prädiktoren (Schneider et al., 2013). Die BiKS-Effektivitätsanalysen stellen also insgesamt weitaus konservativere Schätzungen dar.

6.4 Implikationen für die Forschung

Der Beitrag verdeutlichte, dass der Effektivitätsansatz auf das (deutsche) System frühkindlicher Bildung und Betreuung anwendbar ist. Effektivitätsvergleiche sind auf andere Bildungsbereiche (z. B. sprachliche oder sozio-emotionale Bereiche) sowie weitere Studien übertragbar (z. B. NUBBEK). Abschließende Aussagen zur Effektivität früher mathematischer Bildung in Deutschland sind jedoch u. a. aufgrund beschriebener methodischer Einschränkungen nicht möglich. Ein vollständigeres Bild könnten in Zukunft Ergebnisse aus Studien liefern, welche die hierarchische Struktur und Zusammensetzung des gesamten Systems frühkindlicher Bildung und Betreuung besser abbilden, beispielsweise Gruppen- und Einrichtungsebene sowie übergeordnete Ebenen (z. B. Träger-, Bundesland- oder Länderebene; vgl. Creemers et al., 2010) sowie verschiedene Betreuungsformen mit ausreichend großen Stichproben. Der vorliegende Beitrag zeigte, welche Bedeutung einem längsschnittlichen Design und der Erfassung entwicklungsrelevanter Merkmalen zukommt.

Die EPPSE-Ergebnisse illustrieren das Potenzial der Kombination von Qualitäts- und Effektivitätsansatz und zeigte, dass Effektivität einen aussagekräftigen, zusätzlichen Indikator darstellt, um den Einfluss frühkindlicher Betreuungserfahrung auf kindliche Entwicklung umfassend abzubilden (Melhuish, 2001; Sylva, et al., 2010; Sammons et al., 2014). Die EPPSE-Befunde verdeutlichten zudem, dass Befunde aus komplexen Studiendesigns, in welchen neben Qualitäts- und Effektivitätsansatz weitere Ansätze und Methoden verwendet werden, zur Weiterentwicklung des Qualitätskonzepts beitragen können (Siraj-Blatchford et al., 2006; 2015). In der Bildungsforschung werden weitere aktuelle Längsschnittstudien frühkindlicher Bildung und Betreuung benötigt, welche Bildungseffektivität für verschiedene Bereiche besser

abbilden, und es ermöglichen, Zusammenhänge der Bildungseffektivität sowohl zu kindlicher Entwicklung als auch zur Qualität und weiteren Einrichtungsmerkmalen zu untersuchen. Der vorliegende Beitrag machte darüber hinaus auf einen bislang wenig erforschten, potentiellen Wirkfaktor aufmerksam, der weiterer Untersuchung bedarf: die soziale Zusammensetzung in Einrichtungen (vgl. Kuger & Kluczniok, 2008; Sammons et al., 2002).

6.5 Implikationen für die Praxis

In einzelnen Qualitätsmodellen werden ausdrücklich Bildungsergebnisse als Kriterium herangezogen (OECD, 2006; Roux & Tietze, 2007; Viernickel & Schwarz, 2009), jedoch ist ein holistisches Verständnis kindlicher Entwicklung in den meisten frühkindlichen Konzepten, Bildungsplänen und der Forschung leitend. Aus der elementarpädagogischen Praxis und Forschung ist daher mit einigem Widerstand gegen einen bereichsspezifischen, stark lernzielorientierten Beurteilungsansatzes zu rechnen, der zu den Umständen des Kompetenzerwerbs keine Aussagen macht. Außerdem muss die Aussagekraft des Effektivitätswerts vor dem Hintergrund verwendeter Testverfahren beurteilt werden. Für den mathematischen Bereich beschränken sich viele Verfahren und Befunde auf den Bereich Mengen, Zahlen und Operationen (Jordan et al., 2009; Schneider et al., 2013), Bildungspläne und Fachdidaktik vertreten jedoch ein umfassenderes Verständnis mathematischer Bildung (Kaufmann, 2011; Royar, 2007). Darüber hinaus leistet Bildungseffektivität nur eine relative Aussage zur Güte der Förderung in einer Einrichtung durch den Vergleich zu einer Referenzstichprobe und macht weder Aussagen zu Ansatzpunkten für Entwicklungsmaßnahmen noch darüber, ob das mittlere Effektivitätsniveau der Stichprobe zufriedenstellend ist. Dennoch bietet der Ansatz Praxisimplikationen. So zeigte sich in EPPSE der Wert einer Effektivitätssteigerung von Einrichtungen, weil eine hohe Effektivität v. a. für bildungsbenachteiligte Kinder förderlich und kompensatorisch wirken kann (Sylva et al., 2010; Sammons et al., 2014).

Trotz der methodischen Einschränkungen deuten die Befunde zur Effektivität und Qualität frühkindlicher Bildung auf wichtigen Optimierungsbedarf für das frühkindliche Bildungssystem in Deutschland hin (Kuger & Kluczniok, 2008; Tietze et al., 2013): Wenn Effektivitätsanalysen auch keine direkten Aussage darüber liefern, ob das erreichte Durchschnittsniveau mathematischer Effektivität für eine zufriedenstellende Bildungsarbeit sprechen, sollte durchschnittlich eine hohe Bildungsqualität gewährleistet sein. Insgesamt sprechen die Effektivitätsergebnisse dafür, dass gängige pädagogische Konzepte, Programme und Initiativen nicht zur Optimierung des Systems beitragen können. Denn dann sollten nicht ausschließlich Unterschiede in der Gruppenzusammensetzung Kompetenzunterschiede zwischen Einrichtungen

erklären. Unterschiede in der pädagogischen Ausrichtung und Arbeit einiger Einrichtungen wie eine mathematische Schwerpunktsetzung in der Konzeption und pädagogischen Arbeit, der Verwendung mathematischer Programme oder die Teilnahme an mathematischen Bildungsinitiativen sollten sich auch bei „fairen“ Vergleichen in Effektivitäts- und Qualitätsunterschieden niederschlagen. Die frühkindliche Praxis benötigt innovative Programme und Initiativen früher mathematischer Bildung, die von hoher Effektivität und Qualität sind und verstärkt die unterschiedlichen Voraussetzungen der Kinder in den Einrichtungen in den Blick nehmen.

Literaturverzeichnis

- Anders, Y. (2013). Stichwort: Auswirkungen frühkindlicher institutioneller Betreuung und Bildung. *Zeitschrift für Erziehungswissenschaft*, 16 (2), 237–275.
- Anders, Y., Grosse, C., Roßbach, H.-G., Ebert, S. & Weinert, S. (2013). Preschool and primary school influences on the development of children's early numeracy skills between the ages of 3 and 7 years in Germany. *School Effectiveness and School Improvement*, 24 (2), 195–211.
- Anders, Y., Roßbach, H.-G., Weinert, S., Ebert, S., Kuger, S., Lehl, S. et al. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly*, 27 (2), 231–244.
- Ballou, D. & Springer, M.G. (2015). Using student test scores to measure teacher performance: Some problems in the design and implementation of evaluation systems. *Educational Researcher*, 44 (2), 77–86.
- Bäumer, T., Aßmann, C., von Maurice, J. & Blossfeld, H.-P. (2013). Möglichkeiten der Analyse von Kontexteffekten im Rahmen des Nationalen Bildungspanels. In R. Becker & A. Schulze (Hrsg.), *Bildungskontexte* (S. 61–83). Wiesbaden: Springer Fachmedien Wiesbaden. Zugriff am 23.9.2015. Verfügbar unter: http://link.springer.com/10.1007/978-3-531-18985-7_3
- Becker, B. (2010). Ethnische Unterschiede bei der Kindergartenselektion: Die Wahl von unterschiedlich stark segregierten Kindergärten in deutschen und türkischen Familien. In B. Becker & D. Reimer (Hrsg.), *Vom Kindergarten bis zur Hochschule. Die Generierung von ethnischen und sozialen Disparitäten in der Bildungsbiographie* (S. 17–47). Wiesbaden:

VS Verlag für Sozialwissenschaften.

- Becker, B. & Schober, P.S. (2015). *Not just any day-care center? Social and ethnic disparities in the choice of high quality early education institutions*. Manuscript in preparation.
- Bischof, L.M., Hochweber, J., Hartig, J. & Klieme, E. (2013). Schulentwicklung im Verlauf eines Jahrzehnts – Erste Ergebnisse des PISA-Schulpanels. In N. Jude & E. Klieme (Hrsg.), *PISA 2009 – Impulse für die Schul- und Unterrichtsforschung* (S. 172–199). Weinheim: Beltz.
- Caldwell, B., & Bradley, R. (1984). *Home observation for measurement of the environment (HOME)*. Little Rock, AR: University of Arkansas at Little Rock.
- Chapman, C., Muijs, D., Reynolds, D., Sammons, P. & Teddlie, C. (Eds.). (2015). *Routledge international handbook of educational effectiveness and improvement*. Abingdon, UK / New York, NY: Routledge.
- Clifford, R. (2005). Structure and stability of the early childhood environment rating scale. In H. Schonfeld, S. O'Brien & T. Walsh (Eds.), *Questions of quality* (pp. 12–21). Dublin, Ireland: The Centre for Early Childhood Development & Education, The Gate Lodge, St. Patrick's College.
- Clifford, R.M., Reszka, S.S. & Roßbach, H.-G. (2010). *Reliability and validity of the early childhood environment rating scale*. Zugriff am 23.9.2015. Verfügbar unter: http://www.researchgate.net/profile/Stephanie_Reszka/publication/265098120_Reliability_and_Validity_of_the_Early_Childhood_Environment_Rating_Scale/links/5527e2ad0cf2779ab78ae5f2.pdf
- Creemers, B.P.M., Kyriakides, L. & Sammons, P. (2010). *Methodological advances in educational effectiveness research*. Milton Park, Abingdon, Oxon; New York, NY: Routledge.
- Dalli, C., White, E.J., Rockel, J. & Duhn, I. (2011). *Quality early childhood education for under-two-year-olds: What should it look like?* Wellington, New Zealand: Ministry of Education.
- Deutsches Institut für Normung [DIN]. (2005). *DIN EN ISO 9000:2005 - Qualitätsmanagementsysteme - Grundlagen und Begriffe*. Berlin: Beuth Verlag.
- Ganzeboom, H. B. G., De Graaf, P. M. & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21 (1), 1–56.
- Geier, B. & Riedel, B. (2009). Ungleichheiten der Inanspruchnahme öffentlicher frühpädagogischer Angebote. Einflussfaktoren und Restriktionen elterlicher Betreuungsentscheidungen. In H.-G. Roßbach & H.-P. Blossfeld (Hrsg.), *Frühpädagogische Förderung in Insti-*

- tutionen (S. 11–28). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gordon, R.A., Fujimoto, K., Kaestner, R., Korenman, S. & Abner, K. (2013). An assessment of the validity of the ECERS-R with implications for measures of child care quality and relations to child development. *Developmental Psychology*, 49 (1), 146–160.
- Halle, T., Vick Whittaker, J.E. & Anderson, R. (2010). *Quality in early childhood care and education settings: A compendium of measures*. Washington, DC: Child Trends.
- Hamre, B., Hatfield, B., Pianta, R. & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: associations with preschool children's development. *Child Development*, 85 (3), 1257–1274.
- Jordan, N.C., Kaplan, D., Ramineni, C. & Locuniak, M.N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, 45 (3), 850–867.
- Kane, T.J. & Staiger, D.O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Seattle, WA: MET Project, Bill & Melinda Gates Foundation.
- Katz, L. (1996). Die Qualität der Früherziehung in Betreuungseinrichtungen - Fünf Perspektiven. In W. Tietze (Hrsg.), *Früherziehung. Trends, internationale Forschungsergebnisse, Praxisorientierungen* (S. 226–239). Neuwied: Luchterhand.
- Kaufmann, S. (2011). *Handbuch für die frühe mathematische Bildung*. Braunschweig: Schroedel.
- Kluczniok, K. & Roßbach, H.-G. (2014). Conceptions of educational quality for kindergartens. *Zeitschrift für Erziehungswissenschaft*, 17, 145–158.
- Kuger, S. & Kluczniok, K. (2008). Prozessqualität im Kindergarten—Konzept, Umsetzung und Befunde. *Zeitschrift für Erziehungswissenschaft*, 10, 159–178.
- Kuger, S., Kluczniok, K., Kaplan, D. & Roßbach, H.-G. (2015). Stability and patterns of classroom quality in German early childhood education and care. *School Effectiveness and School Improvement*, 27 (3), 1–23.
- Kuger, S., Pflieger, K., & Roßbach, H.-G. (2005). *Familieneinschätzungsskala Forschungsversion*. Unveröffentlichtes Instrument, Universität Bamberg.
- La Paro, K.M., Thomason, A.C., Lower, J.K., Kintner-Duffy, V.L. & Cassidy, D.J. (2012). Examining the definition and measurement of quality in early childhood education: A review of studies using the ECERS-R from 2003 to 2010. *Early Childhood Research & Practice*, 14 (1), 1–17.

- Lehrl, S., Kluczniok, K. & Roßbach, H.-G. (2016). Longer-term associations of preschool education: The predictive role of preschool quality for the development of mathematical skills through elementary school. *Early Childhood Research Quarterly*, 36, 475-488.
- Lehrl, S., Kuger, S. & Anders, Y. (2014). Soziale Disparitäten beim Zugang zu Kindergartenqualität und differenzielle Konsequenzen für die vorschulische mathematische Entwicklung. *Unterrichtswissenschaft*, 2, 132–151.
- Leyendecker, B., Agache, A. & Madsen, S. (2014). Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit (NUBBEK) – Design, Methodenüberblick, Datenzugang und das Potenzial zu Mehrebenenanalysen. *Zeitschrift für Familienforschung*, 26 (2), 244–258.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau*, 58 (2), 103–117.
- Maas, C.J. & Hox, J.J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 1 (3), 86–92.
- von Maurice, J., Artelt, C., Blossfeld, H.P., Faust, G., Roßbach, H.-G. & Weinert, S. (2007). *Bildungsprozesse, Kompetenzentwicklung und Formation von Selektionsentscheidungen im Vor- und Grundschulalter: Überblick über die Erhebungen in den Längsschnitten BiKS-3-8 und BiKS-8-12 in den ersten beiden Projektjahren*. Bamberg: Universität Bamberg.
- Melchers, P. & Preuß, U. (2003). *Kaufman-Assessment Batterie for Children (K-ABC), Deutsche Version. Individualtest zur Messung von Intelligenz und Fertigkeiten bei Kindern*. Göttingen: Hogrefe.
- Melhuish, E.C. (2001). The quest for quality in early day care and preschool experience continues. *International Journal of Behavioral Development*, 25 (1), 1–6.
- Melhuish, E., Quinn, L., Sylva, K., Sammons, P., Siraj-Blatchford, I. & Taggart, B. (2013). Preschool affects longer term literacy and numeracy: results from a general population longitudinal study in Northern Ireland. *School Effectiveness and School Improvement*, 24 (2), 234–250.
- Melhuish, E., Sammons, P., Siraj-Blatchford, I., Sylva, K., Quinn, L., McSherry, K. et al. (2002). *Characteristics of pre-school environments in Northern Ireland: An analysis of observational data*. Belfast: Stanmillis University College Belfast.
- Moss, P. & Pence, A. (Eds.). (1994). *Valuing quality in early childhood services: New ap-*

- proaches to defining quality*. London: Sage Publications.
- OECD. (2006). *Starting strong II: early childhood education and care*. Paris: OECD Publishing.
- Pianta, R.C. & Hamre, B.K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38 (2), 109–119.
- Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D. et al. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, 9 (3), 144–159.
- Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K. & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, 31, 2–12.
- Racherbäumer, K., Funke, C., van Ackeren, I. & Clausen, M. (2013). Schuleffektivitätsforschung und die Frage nach guten Schulen in schwierigen Kontexten. In R. Becker & A. Schulze (Hrsg.), *Bildungskontexte* (S. 239–267). Wiesbaden: Springer Fachmedien Wiesbaden.
- Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C. et al. (2014). Educational effectiveness research (EER): a state-of-the-art review. *School Effectiveness and School Improvement*, 25 (2), 197–230.
- Roux, S. & Tietze, W. (2007). Effekte und Sicherung von (Bildungs-)Qualität in Kindertageseinrichtungen. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 27 (4), 367–383.
- Royar, T. (2007). Mathematik im Kindergarten: Kritische Anmerkungen zu den neuen „Bildungsplänen“ für Kindertageseinrichtungen. *mathematica didactica*, 30 (1), 29–48.
- Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B. & Elliot, K. (2002). *Measuring the impact of pre-school on children's cognitive progress over the pre-school period* (Technical Paper 8a). London: Institute of Education, University of London.
- Sammons, P., Sylva, K., Melhuish, E., Siraj, I., Taggart, B., Toth, K. et al. (2014). *Influences on students' GCSE attainment and progress at age 16: Effective Pre-School, Primary & Secondary Education Project (EPPSE)*. London/Birbeck/Oxford: University of Oxford/University of London.
- Scheerens, J. (1992). *Effective schooling: Research, theory and practice*. London: Cassell.
- Schneider, W., Küspert, P. & Krajewski, K. (2013). *Die Entwicklung mathematischer Kompe-*

- tenzen. Paderborn: Ferdinand Schöningh.
- Schober, P.S. & Spieß, K.C. (2012). Frühe Förderung und Betreuung von Kindern: bedeutende Unterschiede bei der Inanspruchnahme besonders in den ersten Lebensjahren. *Wochenbericht des DIW Berlin*, 43, 17–28.
- Siraj-Blatchford, I., Kingston, D. & Melhuish, E. (2015). *Assessing quality in early childhood education and care: Sustained shared thinking and emotional well-being SSTEW scale for 2-5 year-olds provision*. London: Institute of Education Press.
- Siraj-Blatchford, I., Sammons, P., Taggart, B., Sylva, K. & Melhuish, E. (2006). Educational research and evidence-based policy: The mixed-method approach of the EPPE project. *Evaluation & Research in Education*, 19 (2), 63–82.
- Slot, P.L., Mulder, H., Verhagen, J. & Leseman, P.P.M. (2014). *Domain-general and domain-specific quality characteristics of early childhood education and care predict growth of two-year-olds' vocabulary and attention skills over one year*. Manuscript submitted for publication.
- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. & Taggart, B. (Eds.). (2010). *Early Childhood Matters: Evidence from the Effective Pre-school and Primary Education Project*. Oxford: Routledge.
- Sylva, K., Siraj-Blatchford, I., Taggart, B., Sammons, P., Melhuish, E., Elliot, K. et al. (2006). Capturing quality in early childhood through environmental rating scales. *Early Childhood Research Quarterly*, 21 (1), 76–92.
- Tietze, W., Becker-Stoll, F., Bensel, J., Eckhardt, A.G., Haug-Schnabel, G., Kalicki, B. et al. (Hrsg.). (2013). *NUBBEK*. Weimar/Berlin: verlag das netz.
- Tietze, W., Schuster, K.-M., Grenner, K. & Roßbach, H.-G. (2005). *Kindergarten-Skala (KES-R). Feststellung und Unterstützung pädagogischer Qualität in Kindergärten*. (3. Auflage). Berlin: Cornelsen.
- Viernickel, S. & Schwarz, S. (2009). *Schlüssel zu guter Bildung, Erziehung und Betreuung – Wissenschaftliche Parameter zur Bestimmung der pädagogischen Fachkraft-Kind-Relation*. Berlin: GEW.

Fußnoten

ⁱ In jeder Gruppe wurde lediglich ein Teil der Kinder getestet, daher entspricht die hier berichtete mittlere Anzahl getesteter Kinder pro Gruppe nicht der mittleren Gruppengröße in BiKS (vgl. Anders et al., 2012; Anders, Grosse, Roßbach, Ebert & Weinert, 2013).

ⁱⁱ BiKS erhob keine Kompetenzen vorm Eintritt in die Einrichtungen. Signifikante *ICCs* für Statusvergleiche (s. Tabelle 1) zum ersten MZP, d. h. für die meisten Kinder kurz nach Eintritt in die Einrichtung, deuteten jedoch auf unterschiedliche Ausgangsniveaus vor Eintritt hin. Zusätzlich berechnete Nullmodelle für entwicklungsrelevante Merkmale deuten auf Unterschiede in der Zusammensetzung bezüglich Alter zum MZP, SES, Eintrittsalter, Migrationshintergrund, mütterlichem Bildungsabschluss sowie häusliche Lernumwelt (HLE).

Tabelle 1

Durch Effektivitätsunterschiede und durch entwicklungsrelevante Merkmale erklärte Varianzanteile in den elementaren mathematischen Kompetenzen zu den drei Messzeitpunkten (MZP) für die vier Vergleiche

Effektivitätsvergleich	ICC	$R^2_{\text{innerhalb}}$	N	Mittlere Clustergröße
a Status				
1. MZP	.21***		554	5.71
2. MZP	.13**		554	5.71
3. MZP	.08*		554	5.71
b Zuwachs				
1. auf 2. MZP	.03	.43***	551	5.68
2. auf 3. MZP	.01	.52***	523	5.39
c Status kontextualisiert				
1. MZP	.08**	.33***	547	5.64
2. MZP	.02	.24***	547	5.64
3. MZP	.01	.18***	547	5.64
d Zuwachs kontextualisiert				
1. auf 2. MZP	.01	.40***	547	5.64
2. auf 3. MZP	.01	.53***	547	5.64

Anmerkungen. N = 97 Einrichtungen. * $p < .05$; ** $p < .01$; *** $p < .001$.

Die Kinder waren zu den drei MZP etwa 3, 4 und 5 Jahre alt.

Tabelle 2

Unstandardisierte Regressionsgewichte und dazugehörige Standardfehler der entwicklungsrelevanten Merkmale bei kontextualisierten Status- und Zuwachsvergleichen (Modelle c und d) für den 1. bis 3. Messzeitpunkt (MZP)

Prädiktoren auf Kindebene	Statusvergleich						Zuwachsvergleich			
	1. MZP		2. MZP		3. MZP		von 1. auf 2. MZP		von 2. auf 3. MZP	
Kompetenzen zu früherem MZP							0.63***	(0.05)	0.65***	(0.04)
Entwicklungsrelevante Merkmale										
Eintrittsalter	-0.06*	(0.03)	0.00	(0.03)	-0.03	(0.03)	0.04	(0.02)	-0.04	(0.03)
Alter zum MZP	0.33***	(0.02)	0.33***	(0.04)	0.25***	(0.04)	0.12***	(0.03)	0.05	(0.03)
Geschlecht ^a	0.85***	(0.23)	0.21	(0.28)	-0.45	(0.31)	-0.32	(0.26)	-0.52*	(0.24)
SES (HISEI)	0.01	(0.01)	0.04**	(0.01)	0.04***	(0.01)	0.03**	(0.01)	0.02*	(0.01)
Muttersprache der Eltern ^b										
einer Deutsch	-0.75	(0.45)	-2.02***	(0.45)	-1.37**	(0.52)	-1.51***	(0.38)	0.22	(0.47)
keiner Deutsch	-2.04***	(0.45)	-1.77**	(0.59)	-0.62	(0.50)	-0.60	(0.46)	1.08*	(0.42)
Mütterlicher Bildungsabschluss ^c										
mittlere Reife	0.32	(0.30)	0.60	(0.39)	0.47	(0.45)	0.40	(0.36)	0.25	(0.36)
Abitur	0.46	(0.33)	0.85	(0.49)	0.38	(0.55)	0.47	(0.42)	-0.11	(0.43)
sonstiger Abschluss	0.39	(0.57)	1.65*	(0.77)	0.54	(0.76)	1.35*	(0.64)	-0.57	(0.67)
Häusliche Lernumwelt (HLE)	4.19***	(0.99)	4.25**	(1.50)	5.13**	(1.52)	1.63	(1.29)	2.82*	(1.10)

Anmerkungen. $N = 546-547$. * $p < .05$; ** $p < .01$; *** $p < .001$. Die Kinder waren zu den drei MZP ungefähr 3, 4 und 5 Jahre alt. Referenzkategorien für die kategorialen Prädiktoren sind ^a männlich, ^b beide Deutsch, ^c kein Abschluss / Hauptschulabschluss.

Anhang C – Manuskript zur Struktur- und Orientierungsqualität (TP3)

Originalmanuskript: Ulferts, H., Hachfeld, A. & Anders, Y. (2016). Associations of school experiences, teacher qualification, and work experience to preschool teachers' current motivational math beliefs. Manuscript in preparation.

Abstract

Research has gathered evidence regarding the benefits of early math education. Though professional competences are seen as key factors for high quality education, preschool teachers' math beliefs as essential prerequisites for math education remain understudied. The present study examined potential influencing factors on three current math beliefs that are assumed to be central for motivation, i.e., value of mathematics, math self-efficacy, and math self-concept. We investigated school experiences, teacher qualification, and work experience as influencing factors along the school and professional career. The results for the 211 German preschool teachers in this study imply that preschool teachers seldom have new experiences with mathematics after their schooling. Regression analyses showed the lasting impact of positive and negative school experiences on motivational math beliefs, especially on math self-concept. Teacher qualification related exclusively to math self-efficacy. Work experience positively predicted math self-efficacy and beliefs about the value of mathematics. Findings indicate a great need for the revision of vocational and further training programs to allow preschool teachers to acquire new experiences with mathematics to prepare, support and motivate preschool teachers for math education.

Keywords: teacher beliefs, preschool mathematics, school experiences, teacher qualification, work experience

1 Introduction

Mathematical skills are regarded as fundamental prerequisites to successful academic careers for children and their later professional development. There is a broad consensus that, if given the appropriate learning opportunities, preschool-aged children will learn basic concepts and skills in mathematics that facilitate later acquisition of a more complex mathematical understanding (Cross et al., 2009). In recent years, mathematics is acknowledged as a major field of early childhood education in many countries and, as such, is integrated into the curricula of most preschools (Organisation for Economic Co-operation and Development [OECD], 2011). Results from longitudinal studies provide evidence for the beneficial effects of attending preschool on emerging mathematic competences (National Institute of Child Health and Human Development Early Child Care Research Network [NICHD ECCRN], 2005; Sylva et al., 2013). The size and the persistence of the effects, however, depend fundamentally on the quality of education, which in turn depends fundamentally on the competences of preschool teachers (Cross et al., 2009).

Preschool teachers' beliefs are regarded as one major component of professional competence, because they guide and motivate daily pedagogical actions (Fives & Buehl, 2012). Their influence on pedagogical actions links beliefs to quality of preschool education (NICHD ECCRN, 2003). Some beliefs may show stronger links to pedagogical interaction, thus may be more central to quality. Firstly, research and literature suggest that domain-specific measures of beliefs are more predictive of the quality of pedagogical interactions in a certain domain than are global beliefs (Fives & Buehl, 2012; Pajares, 1992; Wigfield & Eccles, 2000). Hence, preschool teachers' math-related beliefs and attitudes are assumed to be highly relevant for math education. Secondly, those math beliefs which are close antecedents of motivation should have a stronger influence on teaching practice than other beliefs, especially self-efficacy beliefs (Bandura, 1986; Fives & Buehl, 2012). Hence, math beliefs which are central to motivation should be highly relevant for math education meaning that they relate to the extent to which the daily preschool routines are enriched with learning opportunities and stimulating interactions with mathematical content.

So far, very little empirical research exists that sheds light on math beliefs of preschool teachers and on how these beliefs develop or change over the teacher's career.

The current study explores three math beliefs which are central to motivation: value of mathematics, math self-efficacy and math self-concept. Drawing on a sample of German preschool teachers we investigate how school experiences with mathematics, later qualifications, and work experience relate to these beliefs. In the following section we will first introduce the

conceptual and theoretical grounding of the studied beliefs. Then, we will summarize the research findings on the structure and influencing factors of math beliefs that motivated the current study.

2 Background

2.1 Ability and Value Beliefs as Key Factors for Math Education in Preschool

According to expectancy-value-theory (Eccles et al., 1983; Wigfield, 1994; Wigfield & Eccles, 2000), two types of beliefs should be close antecedents to behavior: the beliefs about how well one will perform on certain activities (ability beliefs), and the extent to which one values those activities (value beliefs). In reference to their motivating power they are often referred to as motivational beliefs. Value beliefs comprise different value components. They can refer to the personal importance of performing well on a task or to the instrumental value or utility of a task, i.e., facilitating important current or future goals. Self-efficacy and self-concept both address beliefs in one's ability. Both ability beliefs are closely related conceptually, but should be treated as distinct entities in research (Bandura, 1986; Ferla et al., 2009). Pajares (1997) has argued that this differentiation was partly lacking in previous research and investigations, but empirical research across countries has demonstrated that the constructs can and should be separated (Lee, 2009).

Bong and Skaalvik (2003) compared both ability beliefs in a review: Self-efficacy refers to one's cognitive appraisal or expectations of what one can achieve in specific situations and as such, differs from self-concept in several aspects. Self-efficacy is a context-specific, and more future- and goal-oriented construct. It is also more malleable and amenable to training efforts. Self-concept, on the other hand, reflects aggregated judgments or overall impressions of one's competence. These competence evaluations include an affective component, and tend to be more past-oriented, comparably stable over time, and resistant to change. To sum up, self-efficacy is defined as the confidence in one's ability to solve specific tasks, whereas self-concept covers the perception of one's general ability in certain domains.

Although literature asserts a strong link between preschool teachers' math beliefs, particularly motivational math beliefs, and educational practice (Bandura, 1986; Fives & Buehl, 2012; Pajares, 1992; Wigfield & Eccles, 2000) only few empirical studies investigating the assumed link exist and most of them were conducted in the United States. These existing studies show heterogeneous results. In a sample of German preschool teachers, Thiel (2010) found an association between stronger math value beliefs and epistemological beliefs that favor constructivist educational practices, which consist of supporting children's active exploration and problem solving in mathematics (see also Benz, 2012b). Lee (2005) studied elementary teach-

ers' math attitudes, which entailed ability beliefs as one aspect. Attitudes toward teaching mathematics were related to teaching practice, whereas attitudes towards mathematics were not. Swars (2005) found that elementary teachers reporting higher teaching efficacy also reported higher use of manipulatives. Two studies by Brown (2003, 2005), however, were not able to show relationships between observed instructional practice and preschool teachers' value beliefs, or their teaching efficacy. Studies found that preschool teachers' beliefs related to facets of professional knowledge (Anders & Rossbach, 2015; Dunekacke et al., in press). Studying the relationship between preschool teachers' beliefs and professional knowledge is important to understand the development of professional competences and their link to pedagogical practice (Tsamir et al., 2014).

2.2 What do Preschool Teachers' Motivational Beliefs in Mathematics Look Like?

Mathematics is, in many countries, a relatively new area of preschool education. This is especially true in Germany. For a long time, the promotion of socio-emotional skills was regarded as the main pedagogical function of German preschools. With the introduction of official curricular guidelines between 2003 and 2007 the promotion of cognitive and (pre)academic skills received greater attention (Kaufmann, 2011; Royar, 2007). Since 2012 all of the guidelines include mathematics as a field of early education. Many German preschool teachers graduated before early math education was considered a main educational area in preschool.

Thus, implementing math education may represent a challenge for preschools and their teachers. Vocational and further training programs have not focused on this educational area for a long time and preschool teachers who have not been trained recently might not be prepared didactically and motivationally for math education. It is often assumed that preschool teachers have negative attitudes towards mathematics in general and as an area of early education in particular (Baroody & Lai, 2006; Copley & Padrón, 1999; 2004; Cross et al., 2009). Research findings on the different motivational math beliefs of preschool and primary school teachers suggest that the status of teachers' motivational preparedness is not as bad as has been assumed, but that it still needs improvement, especially in the case of ability beliefs.

In terms of value beliefs, surveys found that preschool teachers value mathematics as an important area of development (Benz, 2012b; Thiel, 2010), though mainly for older preschool children (Anders & Rossbach, 2015). Preschool and elementary teachers tend to prioritize efforts in other developmental areas such as language or socio-emotional development (Kowalski et al., 2001; Wilkins, 2010). Literature on preschool teachers' efficacy beliefs con-

cludes that, by and large, preschool teachers do not feel confident in their math abilities (Baroody & Lai, 2006; Copley & Padrón, 1999; 2004; Cross et al., 2009).

2.3 Influences on Preschool Teachers' Math Beliefs Along Their Career Paths

Beliefs are commonly considered to be relatively stable, changing slowly over time (Pajares, 1992). If preschool teachers' motivational math beliefs play an important role in educational quality, then knowledge about potential influencing factors throughout preschool teachers' school and professional career is very relevant for high quality of math education. The following sections will give an overview of potential influencing factors most relevant to different stages in preschool teachers' careers: school experiences, teacher qualification, and work experience.

2.3.1 The role of school experiences. As vocational and further training programs for preschool teachers have not focused on mathematics for a long time, it may be assumed that for the majority of preschool teachers, experience with mathematics is mostly restricted to experiences in school. That leads to the assumption that for preschool teachers in particular, earlier school experiences have a strong and lasting impact on the current math beliefs (Pajares, 1992; Relich, 1996; van Oers, 2002; Wilkins, 2008). Therefore, studying connections to memories of school mathematics seems to be a prerequisite for understanding preschool teachers' current beliefs. Attached to those memories are emotions and evaluations of past experiences. They represent powerful influences on beliefs, on motivational beliefs in particular, because emotions and evaluations regulate the amount of energy individuals devote to activities, and as such operate as motivational forces (Fives & Buehl, 2012; Nespors, 1987). Positive emotions associated with an activity, like enjoyment or enthusiasm, relate to interest and intrinsic motivation, whereas negative emotions, in particular anxiety, relate to a tendency to avoid the activity (Eder et al., 2013; Schunk, 1991).

Among preschool teachers, negative attitudes towards mathematics and math anxiety are assumed to be common phenomena (Baroody, 2004; Bates et al., 2013; Copley et al., 2004; Copley & Padrón, 1999). School experiences seem to be strongly related to the development of math anxiety (Bates et al., 2013; Brady & Bowd, 2005; Kelley & Tomhave, 1985; Tobias, 1987). Anxiety and negative feelings about mathematics are problematic for teachers as they might evoke the motivation and behavioral tendency to avoid mathematics and math teaching (Baroody, 2004; Baroody & Lai, 2006; Copley & Padrón, 1999; Hembree, 1999).

It is striking that comparatively few studies on preschool teachers' attitudes, motivations, and beliefs have examined positive emotions, or approach motivation and interest towards mathematics. Recent studies assessing positive and negative beliefs and experiences

reported that preschool teachers' school experiences and emotional attitudes towards mathematics are not as negative as is often assumed and reported mixed beliefs and experiences in samples of preschool teachers (Anders & Rossbach, 2015; Benz, 2012b; Thiel, 2010). Benz (2012a), for example, observed in a German sample that participants frequently chose adjectives like "useful", "important", "challenging" and "interesting" to refer to mathematics, but also adjectives like "confusing" and "incomprehensible".

Research has gathered some evidence for the aforementioned association of emotions and evaluations attached to mathematics, to school memories, and to motivational math beliefs. For instance, preschool teachers who were open towards mathematics hold stronger value beliefs, i.e., were more likely to see the benefits of math skills for everyday life (Thiel, 2010). Math anxiety was negatively related to math teacher efficacy in samples of preservice elementary teachers (Bursal & Paznokas, 2006; Gresham, 2008; Swars et al., 2006) and to math self-concept in a mixed sample of preschool and preservice elementary teachers (Isiksal et al., 2009). Moreover, preservice elementary teachers with lower degrees of math teaching efficacy reported more negative past experiences with mathematics as a student (Swars, 2005).

2.3.2 The role of teacher qualification. It is commonly emphasized that it is the qualification of preschool teachers that plays a pivotal role in the quality of education (Economist Intelligence Unit, 2012; OECD, 2006). After school graduation, two aspects of formal qualification seem to be subsequent influencing factors of motivational math beliefs: a higher degree and specific qualification in mathematics. The former could be an influencing factor across domains, whereas the latter (e.g., offering math courses as part of vocational or further training) could be an influencing factor for math beliefs specifically.

Studies have shown that preschool teachers holding a university degree might provide learning opportunities associated with greater learning gains of young children in different areas, including mathematics (Brown et al., 2008; Early et al., 2006; Saracho & Spodek, 2007). Whitebook (2003) concluded that a four-year college degree should be the minimum requirement for preschool teachers. Research evidence indicates that this effect may partly be due to some specific beliefs of those teachers with higher degrees (Snider & Fu, 1990; Tirosh et al. 2011). Vartuli (1999), for example, showed that preschool teachers with a degree in early childhood education seem to hold stronger beliefs about developmentally appropriate practices, and their practices reflected their beliefs. With regards to motivational math beliefs, some studies found that preschool teachers with a higher degree had developed stronger efficacy beliefs (Cobanoglu, 2011; Tirosh et al., 2011).

Although there is a strong voice for the academization of the preschool teacher profession in Germany, German preschools are lagging behind in their attempt to staff their centers with personnel holding higher degrees. The majority of preschool teachers have completed a three year-post-secondary vocational training program. The proportion of preschool teachers with a college or university degree increased from 3.2 % in 2006 to 4.9 % in 2013 in Germany (Statistisches Bundesamt, 2013). Only 6.9 % of the preschool teachers with a higher degree graduated in early childhood education, which means that only 0.3 % of all preschool teachers have a college or university degree that qualifies them specifically for work with young children. Accordingly, research regarding the influence of academization on preschool teachers' beliefs is sparse. A study of prospective German preschool teachers found that teachers studying at a college or university hold less instructive teaching beliefs and places more emphasis on the centers' educational function when compared to teachers who were trained with common vocational training (Mischo, Wahl, Hendler et al., 2012).

It can be hypothesized that qualifications in a specific domain may be relevant to domain-specific beliefs. In recent years, various programs and courses have been developed in which teachers are trained specifically for math teaching. Little is known about the effects of participating in such programs and courses on preschool teachers' beliefs. Some evaluations of training programs have studied associations with teachers' beliefs (e.g., Akay & Boz, 2010; Albayrak & Unal, 2011; Brand & Wilkins, 2007; Evans et al., 2013; Moseley & Utley, 2006; Saçkes et al., 2012; Swars, 2005; Wilkins & Brand, 2004), but the majority did not use any control group, so caution is warranted against interpretations as intervention effects.

With regards to influence on motivational math beliefs, research found evidence for a connection of math courses to value beliefs (Sherman & Christian, 1999; Wilkins, 2008). Quinn (1997) evaluated a math methods course that used manipulatives, technology, and cooperative learning in a sample of preservice teachers. He found an improvement in content knowledge and math attitudes, which was comprised of aspects of value and efficacy beliefs, for elementary teachers. An association to preservice teachers' value and efficacy beliefs was also found for number of math courses during preparation (Kalder & Lesik, 2011). Darling-Hammond et al. (2002) found in their comparison of various teacher education trainings, that some successfully enhanced teachers' teaching efficacy. It should be noted that programs differed in their impact on beliefs. Stark variation in the quality of such programs (Cross et al., 2009) could explain differences in impact. One of the few attempts made to change preservice preschool teachers' math self-concept was evaluated in a qualitative study by Palmer (2009). The course reported that emphasizing the investigative process, the interdisciplinary nature,

and the aesthetic aspects of mathematical learning was successful in changing value beliefs and the concept of “being a math person”. The latter seems to be closely related to the construct of math self-concept. Although overall the body of evidence is still sparse, research suggests that domain-specific programs may be useful in influencing ability and value beliefs in mathematics.

2.3.3 The role of work experience. In addition to learning and training in formal settings, preschool teachers’ professional development takes places on the job. From a theoretical point of view, work experience could represent an important factor in shaping preschool teachers’ math beliefs. Within their work environment preschool teachers are constantly confronted with expectations, beliefs, and feedback on their behavior from children, parents, and colleagues. Working in such an environment may therefore shape the way they think about children’s learning and development, about effective ways of teaching, and about the way they think about themselves.

Research findings, however, again draw a heterogeneous picture. With regard to self-efficacy, Tirosh et al. (2011) discovered that practicing, when compared to prospective teachers, were more knowledgeable, and more confident about their ability to solve different math tasks. McMullen (1997) found that preschool teachers with more work experience did not report higher educational efficacy, but they reported more agreement to developmentally appropriate practices. In a sample of Turkish preschool teachers, years of experience had only a weak correlation to efficacy beliefs (Cobanoglu, 2011). Other studies, however, produced contradictory findings (e.g., Guo et al., 2011; Vartuli, 1999; Wilkins, 2008). In addition to its potential role as an influencing factor, work experience seems to play a role in the consistencies between the beliefs and pedagogical practices of preschool teachers, with higher consistencies in more experienced teachers (Wen et al., 2011).

2.4 Aims and Research Questions of the Present Study

The present study sets out to explore motivational beliefs in a sample of German preschool teachers. It focuses on three beliefs: value of mathematics, math self-concept and math self-efficacy. It will take a closer look at potential influences along the preschool teachers’ academic and professional career.

Specifically, the study sets out to answer two sets of research questions:

1. Questions addressing the profile of German preschool teachers’ motivational math beliefs and past experiences with mathematics:

1.1 What do German preschool teachers believe about mathematics with regards to the value of mathematics, math self-concept, and math self-efficacy? How are the three motivational math beliefs interrelated?

1.2 What kinds of experiences (e.g., positive and negative) with mathematics did preschool teachers make in school? What kinds of experiences with mathematics did preschool teachers make after their schooling throughout their professional careers?

2. Questions addressing the association of experiences with mathematics to motivational math beliefs:

2.1 How do school experiences relate to current motivational math beliefs?

2.2 How are teacher qualification and work experience associated with current motivational math beliefs after controlling for school experiences?

3 Method

3.1 Participants and Procedure

The current study used data from the study “Pedagogical beliefs of preschool teachers with regard to preschool mathematics” (Anders & Rossbach, 2015). The sample consisted of 221 preschool teachers, who were recruited from 29 preschool centers of two German federal states (54.8 % Berlin and 45.2 % Bavaria). Participants were almost exclusively female (91.4 %). The mean age of the sample was 40 years ($SD = 11.07$) ranging from 18 to 63 years. Study participation was voluntary. Participating teachers completed questionnaires and a test of knowledge in mathematics. The testing took place in preschool centers under the instruction and supervision of trained research staff.

3.2 Instruments

3.2.1 Scale measures. For the current study we used scales assessing preschool teachers’ current motivational math beliefs and past experiences with mathematics in school. We used scores representing the average agreement to various statements. Participants indicated agreements on 4-point Likert scales, which ranged from 1 (*not at all agreed*), 2 (*only partially agreed*), 3 (*largely agreed*) to 4 (*fully agreed*). The reliabilities of the scales, assessed in terms of internal consistency with Cronbach’s alpha (α), were moderate to high. Table 1 provides an overview of the scale measures (descriptive results, number of items, Cronbach’s α , and sample items). Questionnaires were based on adapted items from the German Programme for International Student Assessment [PISA] in 2003 (PISA-Konsortium Deutschland, 2006), except for the questionnaire on emotional experiences with mathematics in school (see below for details).

3.2.1.1 Motivational math beliefs. The three scales of motivational math beliefs captured preschool teachers' current beliefs about mathematics and their math abilities. All items written in present tense to distinguish these current beliefs from past experiences (see below). The scale *value of mathematics* covered important components of subjective task value as proposed in expectancy-value-theory (Eccles et al., 1983). The scale measured the extent to which preschool teachers perceived mathematics as important and useful for one's personal life and society on the whole ($\alpha = .73$ for the 6-item scale). Participants rated their agreements (see section 3.2.1) to items, such as "Mathematics helps to solve daily tasks and problems." The scale *math self-efficacy* assessed preschool teachers' confidence in their ability to solve specific math tasks by confronting them with the description of various tasks, for example "Calculating how much cheaper a TV would be after a discount of 30 percent." For each of the five tasks ($\alpha = .75$), they indicated how confident they felt about having to solve the task. The scale *math self-concept* measured preschool teachers' perception of their general ability to learn quickly and perform well in mathematics through agreement to ability statements in mathematics ($\alpha = .91$ for the 4-item scale), like "I have always believed that mathematics is one of my particular strengths." Higher values on the three scales would indicate that preschool teachers perceived mathematics as important and useful (value of mathematics), that they felt confident in their ability to solve specific math tasks (math self-efficacy), and that they believed that they generally learn quickly and perform well in mathematics (math self-concept).

3.2.1.2 School experiences with mathematics. One part of the survey began with an instruction to think back to times at school. It included questions about preschool teachers' positive and negative emotional and motivational experiences with mathematics in school. The self-constructed items of the two scales of *emotional school experiences* (see also Anders & Rossbach, 2015) were based on work of Benz (2012b) and Thiel (2010). Participants were asked which emotions they associated with math classes at the end of secondary school. Then they rated their agreements to a set of *positive* (e.g., "enjoyment" and "fun") and *negative emotions* (e.g., "anger" and "anxiety") using the same metric for both sets of emotions (see section 3.2.1). The 3-item scale of positive emotions had an internal consistency of $\alpha = .89$ and the 4-item scale of negative emotions had an internal consistency of $\alpha = .88$. Higher values on the emotional scales would imply that preschool teachers associate positive emotions and/or negative emotions with math classes.

Motivational experiences comprised of preschool teachers' interest in the contents of school math (*math interest*, $\alpha = .66$ for the 3-item scale) and their tendency to avoid math

classes (*math avoidance*, $\alpha = .76$ for the 3-item scale). Motivational experiences were assessed with adapted items of the student questionnaires of the German PISA assessments in 2003 (PISA-Konsortium Deutschland, 2006). Participants were asked what they had thought about classes and contents of mathematics in school and rated their agreement to items like “In mathematics I wanted not to memorize content but to truly understand it.” for math interest and “There were times I did not want to go to school because of mathematics.” for math avoidance. Higher values on these scales would be a sign of a greater interest in the contents of school math (math interest) and of a tendency to avoid math classes in the past (math avoidance). All items were written in past tense to distinguish these past experiences from current beliefs and experiences about mathematics.

3.2.2 Single item measures and dummy variables. For measures representing experiences with mathematics after secondary school, i.e. teacher qualification and work experience, we used three dummy variables and one single item measures (see description below).

3.2.2.1 Teacher qualification and work experience. To explore the relationships between a higher degree and current motivational math beliefs we used preschool teachers’ university qualifying degrees as a proxy for a higher degree. This was primarily because the number of preschool teachers with university or college degrees is still very low in Germany (Statistisches Bundesamt, 2013); in the current sample the proportion was less than 4 %. We used a dummy variable to indicate if participants left school 1 = *with* or 0 = *without a university qualifying degree* (the German Abitur). Three items covered specific qualifications in mathematics. Two dummy variables indicated if participants had participated in courses that prepare for early math education as part of vocational training or further training, i.e., *math courses in vocational training* and *math courses in further training* (coded as 0 = *no participation in any course* and 1 = *participation in at least one course*). With a single item we assessed *work experience* within early education childcare in number of years.

3.3 Data Analysis

We used descriptive and correlational analyses to answer the first set of research questions. To analyse the relations of school experiences, teacher qualification, and work experiences, we computed blockwise linear regression using *Mplus* (Version 7). We entered school experiences in a first step and teacher qualifications and work experience in second step to compare their strength as predictors. To avoid bias of regression results due to missing values on variables we applied the Full Information Maximum Likelihood procedure for the estimation of missing data (FIML; Enders, 2001; Lüdtke et al., 2007). We did not apply FIML estimation for dummy variables, which resulted in a reduction of sample size in our regressions

($N = 203$). Percentages of missing data ranged from 2.7 % (math avoidance) to 10.9 % (importance of mathematics). As we recruited several teachers from the same preschools who probably share similar training experiences and beliefs we used robust standard errors in regressions to account for the hierarchical structure of the data (i.e., teachers nested in preschools). For each continuous variable we calculated the intra-class correlation (ICC), which is a measure of how correlated observations within one preschool are. It therefore reflects how much preschool teachers from the same preschool resemble each other in each of the studied variables. ICCs ranged from .05 to .08 for motivational math beliefs, and from .01 (math interest in school) to .19 (work experience) for the predictors.

4 Results

4.1 Profile of Preschool Teachers' Motivational Math Beliefs and Past Experiences with Mathematics

4.1.1 What do motivational math beliefs look like and how do they interrelate?

Table 1 shows the descriptive statistics for motivational math beliefs and school experiences with mathematics. The pattern of means reveals important differences between motivational math beliefs within the studied sample. The mean for value of mathematics ($M = 2.89$, $SD = .51$) was close to 3 indicating that preschool teachers in the current study on average appreciated mathematics as useful and important. Comparing the two ability beliefs reveals a low mean for general math self-concept ($M = 1.83$, $SD = .76$), while the mean for the more task-specific math self-efficacy was statistically significantly higher ($M = 3.10$, $SD = .67$, $t(205) = 23.12$, $p < .001$). Bivariate associations between the three motivational math beliefs support the idea of interrelated, but separate constructs. Correlations of value of mathematics to math self-efficacy and math self-concept were statistically significant and moderate in size ($r = .29$ and $r = .34$, $p < .001$). The correlation between the two ability beliefs was slightly higher ($r = .42$, $p < .001$).ⁱⁱⁱ

4.1.2 What kind of experiences with mathematics did preschool teachers make throughout their school and professional careers?

Descriptive statistics of *school experiences* with mathematics are also illustrated in Table 1. We observed a similar pattern for motivational experiences, i.e., math interest and math avoidance, as was previously reported for positive and negative emotions in school of the same sample (Anders & Rossbach, 2015). In the current sample, neither emotional nor motivational experiences were as bad as is often assumed for preschool teachers (Baroody, 2004; Baroody & Lai, 2006; Bates et al., 2013; Copley et al., 2004; Copley & Padrón, 1999; Hembree, 1999; see section 2.3.1). Participants did not report negative emotions or math avoidance

($M = 1.92$, $SD = .81$ and $M = 1.87$, $SD = .71$). However, they also did not report positive emotions or math interest either ($M = 1.98$, $SD = .78$ and $M = 2.19$, $SD = .66$). T-tests showed that the means of positive and negative emotions did not differ ($t(196) = 0.76$, $p = n.s.$), but that the mean of math interest was higher than the mean of math avoidance ($t(209) = 4.19$, $p < .001$).

With respect to formal *teacher qualifications*, 31.7 % of the preschool teachers in this sample graduated from school with a university qualifying degree and 3.6 % achieved a college or university degree. It is striking that experiences with mathematics in formal training seems to happen only in few cases. Only 27.1% of our sample reported that they had attended math courses during vocational training and only 65.2% remembered having participated in at least one math course during further training. Preschool teachers in the present sample differed greatly in *work experience* (ranging from 3 months to 43 years, $SD = 10.87$), meaning that they differed greatly in their opportunities for informal experiences with mathematics and math teaching through daily interactions. On average, participants had worked for 15 and a half years within childcare service.

4.2 Association of Experiences with Mathematics with Motivational Math Beliefs

4.2.1 How do school experiences relate to current motivational math beliefs? Table 2 summarizes the standardized regression weights of the blockwise regressions for all three motivational math beliefs separately. Following the order of proposed research questions we will start with the description of the results for school experiences with mathematics (step 1 of the regressions). Not every facet of school experiences was related to the studied beliefs, and the predictive power of the different facets varied across beliefs. Findings for positive school experiences indicate that positive emotions predicted positively and statistically significantly value of mathematics and math self-concept ($\beta = .24$ and $\beta = .35$, $p < .001$). Math interest was statistically significantly related to math self-concept and marginally related to value of mathematics ($\beta = .25$, $p < .01$ and $\beta = .14$, $p < .10$). No facet of positive experience was related to math self-efficacy. Results for negative school experiences reveal that negative emotions did not play a role in predicting any of the studied beliefs when simultaneously entered with the other facets into the regressions. Math avoidance, however, is a significant predictor of both ability beliefs ($\beta = -.35$, $p < .01$ for math self-efficacy and $\beta = -.50$, $p < .001$ for math self-concept).

4.2.2 How are teacher qualification and work experience associated to current motivational math beliefs? Results of the second step of the regressions displayed in Table 2 show what teacher qualification and work experience with mathematics relate to current moti-

vational beliefs after controlling for school experiences. Generally, adding teacher qualification and work experience as predictors resulted only in minor changes to the coefficients of school experiences and the above-mentioned relationships remained statistically significant. Adding this set of predictors led to changes in R-square for value of mathematics and math self-efficacy ($\Delta R^2 = .06$ and $\Delta R^2 = .07$) but almost no change was observed for math self-concept ($\Delta R^2 = .01$).

Teacher qualification did not relate to either value of mathematics or to math self-concept, once school experiences were controlled for. Teacher qualification did relate to math self-efficacy. Preschool teachers with a higher degree ($\beta = .26, p < .001$) and preschool teachers who reported having attended math courses during vocational training ($\beta = .11, p < .10$) indicated a higher math self-efficacy.^{iv}

Finally, work experience was a positive predictor of the value of mathematics ($\beta = .20, p < .01$) and math self-efficacy ($\beta = .16, p < .05$), indicating that opportunities for informal experiences with mathematics and math teaching through daily work might result in a higher appreciation of mathematics and a higher confidence in one's task-specific abilities.

5 Discussion

The present study reported descriptive overview and examined the associations of school experiences, teacher qualification, and work experience to preschool teachers' current motivational math beliefs in a sample of 211 German preschool teachers.

Descriptive findings for beliefs indicate that participating preschool teachers valued mathematics and felt relatively confident about specific math abilities, though their math self-concept seemed to be comparably lower than their math self-efficacy. Findings for both ability beliefs, i.e., the difference in mean scores, the moderate size of correlation, and the observed differences in regression results (further described below), underline the importance of examining both beliefs separately. With regard to preschool teachers' *school experiences*, the present study examined school as a potential source for both positive and negative experiences with mathematics. Research with preschool teachers thus far has largely focused on school experiences as a source for negative math experiences and ignoring its potential to evoke positive emotions and motivation in preschool teachers. We found that participants did not report any great interest or strongly positive emotions with regards to mathematics, but they also did not report high avoidance or strongly negative emotions either. Current findings do not support previous claims that preschool teachers suffer from unpleasant and painful past experience with mathematics in school (Baroody, 2004; Brand & Wilkens, 2007; Copley & Padrón, 1999; Copley, Clements, & Sarama, 2004; Hembree, 1999).

Descriptive findings for *teacher qualification* reveal that the lack of positive math experiences in school is troubling insofar as preschool teachers made very little formal experience with mathematics during vocational or further training. Despite the fact that German preschool teachers are required to include math education in their daily routines, only few preschool teachers in the current sample reported that they participated in math courses as part of vocational or further training. A reason for the low importance of mathematics in the vocational training of participating preschool teachers might be that most German preschool teachers graduated before early math education was considered a main educational area in preschool: the mean years of work experiences of participants was very high. This implies that most participants had worked in preschools for several years before math education was implemented in the curricula of German preschools. Even after its implementation, training programs apparently do not sufficiently allow preschool teachers to acquire new experiences with mathematics after school graduation.

Implementation of mathematics as an educational area in preschools should be complemented by a thoughtful adaptation of the vocational and further training programs to the new requirements and to the actual needs of in- and pre-service preschool teachers. Research should accompany the adaptation process by examining the competences of preschool teachers, influencing factors, as well as effective interventions with regards to early math education.

Regarding questions addressing the *association of experiences with mathematics to motivational math beliefs*, results of the regression analyses showed that all of the studied predictors (i.e., positive and negative school experiences with mathematics, teacher qualification, and work experience) predicted motivational math beliefs (e.g., value of mathematics, math concept, and math self-efficacy) to some extent.

With regard to *positive and negative school experiences with mathematics*, regression results showed significant associations with all three studied motivational math beliefs. Math avoidance, for example, as one aspect of negative experiences was negatively associated with math self-efficacy and math self-concept. This finding parallels results of other studies (Isiksal et al. 2009; Swars 2005; Swars et al. 2006), which found associations between anxiety and former negative experiences with mathematics and elementary teachers' efficacy beliefs. The current study extended the common view on school mathematics as a negative influence on current math beliefs of preschool teachers by showing that positive school experiences (positive emotions and math interest) related to motivational math beliefs. Generally, it was striking that positive and negative school experiences displayed higher regression weights than any of the other predictors (e.g., teacher qualification and work experience). Adding qualification and

work experience as predictors only resulted in minor changes to coefficients of school experiences. The results speak in favor of the lasting impact of experiences that preschool teachers made with mathematics during their time as students (Pajares 1992; van Oers 2002; Wilkins 2008), despite preschool teachers' new experiences with mathematics during vocational and further trainings or during their daily work with children.

With regard to *teacher qualifications*, regression results showed that the different aspects of formal qualifications did explain variance in math self-efficacy only, after controlling for school experiences. A higher degree and math-specific courses during vocational training were positively associated with math self-efficacy. Cobanoglu (2011) and Tirosh et al. (2011) reported similar findings for a higher degree and domain-unspecific efficacy beliefs in samples of preschool teachers. Similarly, Mischo, Wahl, Hendler, et al. (2012) observed that German preschool teachers with higher and lower degrees differed in their beliefs (see also Mischo, Wahl, Strohmer, et al., 2012). Several studies have shown positive effects of math-specific interventions on teachers' efficacy beliefs in mathematics (Quinn, 1997; Rosenfeld, 2010; Saçkes et al., 2012; Swars, 2005), which is in line with the observed positive association of math courses during vocational training. Although teaching specific content and specific skills seems to enhance confidence in successfully solving specific math tasks (math self-efficacy), this confidence does not necessarily generalize to less task-specific, more general beliefs of preschool teachers, including beliefs about general math ability (math self-concept) or the general value of mathematics. Neither math self-concept nor the value of mathematics was associated with math-specific training. Our interpretation of this finding is that the malleability of math self-concept - as a relatively stable construct over time (Pajares, 1997) - is extremely limited once it has been formed over the school career. Yet another explanation could be the nature and quality of experiences with mathematics that vocational and further training programs provide. Information and research on the contents and methods of vocational and further training courses for math education is sparse. There is some evidence that training programs differ widely in quality (Cross et al., 2009). Low quality might explain, to some extent, the lack of prediction of more general beliefs. When programs differ in quality, they differ in their consequences in terms of the sense of preparedness for teaching and efficacy beliefs as Darling-Hammond et al. (2002) found in their comparison of various teacher education programs in New York.

In sum, our findings of the limited predictive power of specific qualifications, especially during further training, strengthen the aforementioned claim for revising training programs to prepare and support preschool teachers better for early math education. Besides allowing for

and requiring new math experiences, the nature of experiences in trainings should be a subject to change. If math courses are perceived as incomprehensible, frustrating, and boring, or contents are too complex, above one's level of ability, and is taught as being static and of little relevancy for personal life and society, participation may have no or even negative effects on preschool teachers' motivational math beliefs. In addition to providing and requiring more math courses, a thorough revision should entail revising content and methods to allow for new positive emotional and motivational experiences with mathematics and for reflecting on and overcoming emotional and motivational barriers of the past (Akay & Boz, 2010; Lake & Kelly, 2014; Tooke & Lindstrom, 1998).

In the current study, preschool teachers' *work experience* was positively associated with the value of mathematics and math self-efficacy. Previous studies have reported differences in self-efficacy in relation to varying work experience (Cobanoglu, 2011; Tirosh et al., 2011). Daily work in early childhood education provides preschool teachers with opportunities for informal experiences with mathematics and math education. Nonetheless, work settings should meet certain criteria to be beneficial for preschool teachers' self-efficacy or other beliefs (see Guo et al., 2011; Buchanan et al., 1998; Upadyaya et al., 2011). The potential of the work context as an influential factor on preschool teachers' math beliefs should be maximized by strengthening the salience of mathematics as an important area of early education. Salience could be enhanced by incorporating math education into the preschool's conception, by implementing regular documentations of children's math learning as standard procedures within the preschool, or by declaring learning processes and educational experiences in mathematics to be integral parts of team meetings and work with parents.

5.1 Limitations

The current study faces a number of limitations which should be addressed in future research. The current study is subject to selection bias, a common issue in non-experimental field studies. The current sample of preschool teachers is selective in at least two ways. Firstly, participants are preschool teachers from two of the sixteen federal states of Germany, i.e., Bavaria and Berlin. Regional variations in important aspects of the childcare systems (e.g., different federal curricula, qualification requirements and programs) challenge attempts to generalize findings from our sample to the whole population of German preschool teachers. Secondly, although we randomly selected preschool centers for the study, preschool teachers were free in their choice to participate. It is possible that mainly preschool teachers with a high motivation and overall positive math experiences and beliefs participated in our study. Nonetheless, range and variances of math beliefs and experiences were considerable and the data on

school experiences with mathematics did not reflect the picture of overall positive and enthusiastic participants. Another threat to the reliability of results pertaining to math beliefs and school experiences is the phenomenon of social desirability. Participants were certainly aware that mathematics forms part of the preschool curriculum and that preschool teachers are expected to show a certain appreciation and interest, as well as confidence and proficiency in mathematics. Furthermore, the retrospective assessment of school experiences was the only possible way to examine the common notion of the lasting impact of (negative) school experiences on preschool teachers' current math beliefs and attitudes towards mathematics. As a consequence, associations between school experiences and current motivational beliefs may be biased and inflated. However, as mentioned earlier, descriptive findings for the current sample do not suggest that findings were biased enormously by social desirable responses and the observed association did not seem to be inflated enormously. In fact, associations of school experiences varied substantially across studied beliefs.

We could not directly study the relationship between university or college degrees and beliefs as there were too few participants with a university or college degree in our sample. That is the reason why we could explore benefits of an academization of the profession for motivational math beliefs only by using the university qualifying degree as a proxy for a higher degree.

5.2 Conclusion

The current study contributes to the understanding of the understudied formation of preschool teachers' math beliefs by exploring potential influencing factors along their school and professional careers. Positive and negative school experiences with mathematics, later formal qualification, and work experience predicted math beliefs to some extent, but findings varied across studied beliefs. Current findings indicate that preschool teachers made few experience with mathematics after school graduation during teacher trainings or work and that the motivating potential of experiences during formal training and daily work should be enhanced. As the current study focused on three beliefs which are assumed to be central to preschool teachers' motivation, as a next step, research should examine consequences of motivational math beliefs for preschool teachers' practice, the quality of math education, and children's learning trajectories.

References

- Akay, H., & Boz, N. (2010). The effect of problem posing oriented analyses-II course on the attitudes toward mathematics and mathematics self-efficacy of elementary prospective mathematics teachers. *Australian Journal of Teacher Education*, 35(1).
- Albayrak, M., & Unal, Z. A. (2011). The effect of methods of teaching mathematics course on mathematics teaching efficacy beliefs of elementary pre-service mathematics teachers. *International Journal of Humanities and Social Science*, 1, 183–190.
- Anders, Y., & Rossbach, H.-G. (2015). Preschool teachers' sensitivity to mathematics in children's Play: The influence of math-related school experiences, emotional attitudes, and pedagogical beliefs. *Journal of Research in Childhood Education*, 29, 305–322.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, N.J.: Prentice Hall.
- Baroody, A. J. (2004). The role of psychological research in the development of early childhood mathematics standards. In D. H. Clements & J. Samara (Eds.), *Engaging young children in mathematics: standards for early childhood mathematics education* (pp. 149–172). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Baroody, A. J., & Lai, M. (2006). The development of young children's early number and operation sense and its implications for early childhood education. In B. Spedek & O. N. Saracho (Eds.), *Handbook of research on the education of young children* (2nd ed., pp. 187–221). Mahwah, N.J. / London: Lawrence Erlbaum Associates.
- Bates, A. B., Latham, N. I., & Kim, J. (2013). Do I have to teach math? Early childhood pre-service teachers' fears of teaching mathematics. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, 5, 1-10. Retrieved from <http://www.k-12prep.math.ttu.edu/journal/5.attributes/bates01/article.pdf>
- Benz, C. (2012a). Attitudes of kindergarten educators about math. *Journal Für Mathematik-Didaktik*, 33(2), 203–232.
- Benz, C. (2012b). Maths is not dangerous – attitudes of people working in German kindergarten about mathematics in kindergarten. *European Early Childhood Education Research Journal*, 20(2), 249–261.
- Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, 15(1), 1–40.
- Brady, P., & Bowd, A. (2005). Mathematics anxiety, prior experience and confidence to teach mathematics among pre-service education students. *Teachers and Teaching: Theory and Practice*, 11(1), 37–46.

- Brand, B. R., & Wilkins, J. L. M. (2007). Using self-efficacy as a construct for evaluating science and mathematics methods courses. *Journal of Science Teacher Education*, 18(2), 297–317.
- Buchanan, T. K., Burts, D. C., Bidner, J., White, V. F., & Charlesworth, R. (1998). Predictors of the developmental appropriateness of the beliefs and practices of first, second, and third grade teachers. *Early Childhood Research Quarterly*, 13(3), 459–483.
- Bursal, M., & Paznokas, L. (2006). Mathematics anxiety and preservice elementary teachers' confidence to teach mathematics and science. *School Science and Mathematics*, 106(4), 173–180.
- Cobanoglu, R. (2011). *Teacher self-efficacy and teaching beliefs as predictors of curriculum implementation in early childhood education*. Middle East Technical University, Ankara. Retrieved from <https://etd.lib.metu.edu.tr/upload/12613492/index.pdf>
- Copley, J. V., & Padrón, Y. (1998). *Preparing teachers of young learners: Professional development of early childhood teachers in mathematics and science* (Project No. 2061). Retrieved from the American Association for the Advancement of Science (AAAS) website: <https://www.project2061.org/publications/earlychild/online/fostering/copleyp.htm>.
- Copley, J. V., Clements, D. H., & Sarama, J. (2004). The early childhood collaborative: A professional development model to communicate and implement the standards. In D. H. Clements & J. Sarama (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 401–414). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Cross, C. T., Woods, T. A., & Schweingruber, H. A. (Eds.). (2009). *Mathematics learning in early childhood: paths toward excellence and equity*. Washington, DC: National Academies Press.
- Darling-Hammond, L., Chung, R., & Frelow, F. (2002). Variation in teacher preparation: How well do different pathways prepare teachers to teach? *Journal of Teacher Education*, 53(4), 286–302.
- Dunekacke, S., Jenßen, L., & Blömeke, S. (in press). Relationship of prospective pre-school teachers' knowledge, beliefs and motivation to their perception and planning abilities in the field of mathematics. *ZDM – The International Journal of Mathematics Education*.
- Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S., ... Barbarin, O. (2006). Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly*, 21(2), 174–195.

- Eccles, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–146). San Francisco: W. H. Freeman.
- Economist Intelligence Unit. (2012). Starting well. Benchmarking early education across the world. Retrieved from http://www.lienfoundation.org/pdf/publications/sw_report.pdf
- Eder, A. B., Elliot, A. J., & Harmon-Jones, E. (2013). Approach and avoidance motivation: Issues and advances. *Emotion Review*, 5(3), 227–229.
- Enders, C. K. (2001). The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educational and Psychological Measurement*, 61(5), 713–740.
- Evans, B. R., Leonard, J., Krier, K., & Ryan, S. (2013). The influence of a reform-based mathematics methods course on preservice teachers' beliefs. *Journal of Educational Research and Practice*, 3(1), 79–92.
- Ferla, J., Valcke, M., & Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences*, 19(4), 499–505.
- Fives, H., & Buehl, M. M. (2012). Spring cleaning for the “messy” construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, & M. Zeidner (Eds.), *APA educational psychology handbook* (Vol. 2, pp. 471–499). Washington: American Psychological Association.
- Fox, J. (2015). *Applied regression analysis and generalized linear models* (3 ed.). Thousand Oaks, CA: Sage Publications.
- Gresham, G. (2008). Mathematics anxiety and mathematics teacher efficacy in elementary pre-service teachers. *Teaching Education*, 19(3), 171–184.
- Guo, Y., Justice, L. M., Sawyer, B., & Tompkins, V. (2011). Exploring factors related to pre-school teachers' self-efficacy. *Teaching and Teacher Education*, 27(5), 961–968.
- Hembree, R. (1999). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, 21(1), 33–46.
- Isiksal, M., Curran, J. M., Koc, Y., & Askun, C. S. (2009). Mathematics anxiety and mathematical self-concept: Considerations in preparing elementary-school teachers. *Social Behavior and Personality: An International Journal*, 37(5), 631–643.
- Kalder, R. S., & Lesik, S. A. (2011). A classification of attitudes and beliefs towards mathematics for secondary mathematics pre-service teachers and elementary pre-service teach-

- ers: An exploratory study using latent class analysis. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, 5. Retrieved from <http://www.k-12prep.math.ttu.edu/journal/5.attributes/kalder01/article.pdf>
- Kaufmann, S. (2011). *Handbuch für die frühe mathematische Bildung* [Handbook of early math education]. Braunschweig: Schroedel.
- Kelley, W. P., & Tomhave, W. K. (1985). A study of math anxiety/math avoidance in preservice elementary teachers. *The Arithmetic Teacher*, 32(5), 51-53.
- Kowalski, K., Pretti-Frontczak, K., & Johnson, L. (2001). Preschool teachers' beliefs concerning the importance of various developmental skills and abilities. *Journal of Research in Childhood Education*, 16(1), 5–14.
- Lake, V. E., & Kelly, L. (2014). Female preservice teachers and mathematics: Anxiety, beliefs, and stereotypes. *Journal of Early Childhood Teacher Education*, 35(3), 262–275.
- Lee, J. (2005). Correlations between kindergarten teachers' attitudes toward mathematics and teaching practice. *Journal of Early Childhood Teacher Education*, 25(2), 173–184.
- Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual Differences*, 19(3), 355–365.
- Lüdtke, O., Robitzsch, A., Trautwein, U., & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung [Handling missing data in psychological research]. *Psychologische Rundschau*, 58(2), 103–117.
- McMullen, M. B. (1997). The effects of early childhood academic and professional experience on self perceptions and beliefs about developmentally appropriate practices. *Journal of Early Childhood Teacher Education*, 18(3), 55–68.
- Mischo, C., Wahl, S., Hendler, J., & Strohmer, J. (2012). Pädagogische Orientierungen angehender frühpädagogischer Fachkräfte an Fachschulen und Hochschulen [Pedagogical orientations of pre-service early childhood educators at vocational schools and universities]. *Frühe Bildung*, 1(1), 34–44.
- Mischo, C., Wahl, S., Strohmer, J., & Hendler, J. (2012). Knowledge orientations of prospective early childhood teachers: A study of students' scientific versus subjective orientations in teacher education courses in Germany. *Journal of Early Childhood Teacher Education*, 33(2), 144–162.
- Moseley, C., & Utley, J. (2006). The effect of an integrated science and mathematics content-based course on science and mathematics teaching efficacy of preservice elementary teachers. *Journal of Elementary Science Education*, 18(2), 1–12.

- Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies*, 19(4), 317–328.
- National Institute of Child Health and Human Development Early Child Care Research Network [NICHD ECCRN] (2003). Does quality of child care affect child outcome at age 4 1/2? *Developmental Psychology*, 39(3), 451–469.
- National Institute of Child Health and Human Development Early Child Care Research Network [NICHD ECCRN] (2005). Early child care and children's development in the primary grades: Follow-up results from the NICHD study of early child care. *American Educational Research Journal*, 537–570.
- Organisation for Economic Co-operation and Development [OECD] (2011). *Starting Strong III: A quality toolbox for early childhood education and care*. OECD Publishing. Retrieved from <http://dx.doi.org/10.1787/9789264123564-5-en>
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307–332.
- Pajares, M. F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 1-49). Greenwich, CT: Jai Press.
- Palmer, A. (2009). "I'm not a 'maths-person'!" Reconstituting mathematical subjectivities in aesthetic teaching practices. *Gender and Education*, 21(4), 387–404.
- PISA-Konsortium Deutschland (Eds.). (2006). *PISA 2003: Dokumentation der Erhebungsinstrumente* [Documentation of survey measures]. Münster: Waxman-Verlag.
- Quinn, R. J. (1997). Effects of mathematics methods courses on the mathematical attitudes and content knowledge of preservice teachers. *The Journal of Educational Research*, 91(2), 108–114.
- Relich, J. (1996). Gender, self-concept and teachers of mathematics: Effects on attitudes to teaching and learning. *Educational Studies in Mathematics*, 30(2), 179–195.
- Rosenfeld, D. (2010). Increasing perceived efficacy for teaching mathematics: An exploratory study. *Journal of Mathematics Education at Teachers College*, 1, 25–35.
- Royar. (2007). Mathematik im Kindergarten: Kritische Anmerkungen zu den neuen „Bildungsplänen“ für Kindertageseinrichtungen [Mathematics in preschool: Critical comments about the new curricular guidelines for preschool]. *Mathematica Didactica*, 30(1), 29–48.
- Saçkes, M., Flevares, L. M., Gonya, J., & Trundle, K. C. (2012). Preservice early childhood teachers' sense of efficacy for integrating mathematics and science: Impact of a methods course. *Journal of Early Childhood Teacher Education*, 33(4), 349–364.

- Saracho, O. N., & Spodek, B. (2007). Early childhood teachers' preparation and the quality of program outcomes. *Early Child Development and Care*, 177(1), 71–91.
- Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, 26(3-4), 207–231.
- Sherman, H. J., & Christian, M. (1999). Mathematics attitudes and global self-concept: An investigation of the relationship. *College Student Journal*, 33(1), 331–346.
- Snider, M. H., & Fu, V. R. (1990). The effects of specialized education and job experience on early childhood teachers' knowledge of developmentally appropriate practice. *Early Childhood Research Quarterly*, 5(1), 69–78.
- Statistisches Bundesamt. (2013). *Kinder und tätige Personen in Tageseinrichtungen und in öffentlich geförderter Kindertagespflege am 01.03.2013* [Children and the workforce in daycare centers and public founded daycare at the 1/03/13] (Statistiken der Kinder- und Jugendhilfe No. 5225402137004). Wiesbaden: Statistisches Bundesamt.
- Swars, S. L. (2005). Examining perceptions of mathematics teaching effectiveness among elementary preservice teachers with differing levels of mathematics teacher efficacy. *Journal of Instructional Psychology*, 32(2), 139–147.
- Swars, S. L., Daane, C. J., & Giesen, J. (2006). Mathematics anxiety and mathematics teacher efficacy: What is the relationship in elementary preservice teachers? *School Science and Mathematics*, 106(7), 306–315.
- Sylva, K., Sammons, P., Chan, L. L. S., Melhuish, E., Siraj-Blatchford, I., & Taggart, B. (2013). The effects of early experiences at home and pre-school on gains in English and mathematics in primary school: a multilevel study in England. *Zeitschrift Für Erziehungswissenschaft*, 16(2), 277–301.
- Tsamir, P., Tirosh, D., Levenson, E., Tabach, M., & Barkai, R. (2014). Employing the CAMTE framework: Focusing on preschool teachers' knowledge and self-efficacy related to students' conceptions. In U. Kortenkamp, B. Brandt, C. Benz, G. Krummheuer, S. Ladel, & R. Vogel (Eds.), *Early Mathematics Learning* (pp. 291–306).
- Thiel, O. (2010). Teachers' attitudes towards mathematics in early childhood education. *European Early Childhood Education Research Journal*, 18(1), 105–115.
- Tirosh, D., Tsamir, P., Levenson, E., Tabach, M., Barkai, R., Roesken, B., & Casper, M. (2011). *Prospective and practicing preschool teachers' mathematics knowledge and self-efficacy: Identifying two and three dimensional figures*. Retrieved from http://www.ruhr-uni-bochum.de/imperia/md/content/mathematik/Roesken/paper_tirosh_mavi.pdf
- Tobias, S. (1987). *Overcoming math anxiety*. Boston: Houghton Mifflin Company.

- Brown, E. T. (2003). *The influence of teachers' efficacy and beliefs on mathematics instruction in the early childhood classroom*. University of Louisville, Louisville, Ky.
- Brown, E. T. (2005). The influence of teachers' efficacy and beliefs regarding mathematics instruction in the early childhood classroom. *Journal of Early Childhood Teacher Education*, 26(3), 239–257.
- Brown, E. T., Molfese, V. J., & Molfese, P. (2008). Preschool student learning in Literacy and mathematics: Impact of teacher experience, qualifications, and beliefs on an at-risk sample. *Journal of Education for Students Placed at Risk*, 13(1), 106–126.
- Tooke, D. J., & Lindstrom, L. C. (1998). Effectiveness of a mathematics methods course in reducing math anxiety of preservice elementary teachers. *School Science and Mathematics*, 98(3), 136–139.
- Upadyaya, K., Viljaranta, J., Lerkkanen, M.-K., Poikkeus, A.-M., & Nurmi, J.-E. (2011). Cross-lagged relations between kindergarten teachers' causal attributions, and children's interest value and performance in mathematics. *Social Psychology of Education*, 15(2), 181–206.
- Van Oers, B. (2002). Teachers' epistemology and the monitoring of mathematical thinking in early years classrooms. *European Early Childhood Education Research Journal*, 10(2), 19–30.
- Vartuli, S. (1999). How early childhood teacher beliefs vary across grade level. *Early Childhood Research Quarterly*, 14(4), 489–514.
- Wen, X., Elicker, J. G., & McMullen, M. B. (2011). Early childhood teachers' curriculum beliefs: Are they consistent with observed classroom practices? *Early Education & Development*, 22(6), 945–969.
- Whitebook, M. (2003). *Early education quality: Higher teacher qualifications for better learning environments - a review of the literature*. Retrieved from http://www.irle.berkeley.edu/cscce/wp-content/uploads/2003/01/Early_Ed_Quality.pdf.
- Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review*, 6(1), 49–78.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81. doi:10.1006/ceps.1999.1015
- Wilkins, J. L. M. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education*, 11(2), 139–164.

Wilkins, J. L. M. (2009). *Elementary school teachers' attitudes toward different subjects*. *The Teacher Educator*, 45(1).

Wilkins, J. L. M., & Brand, B. R. (2004). Change in preservice teachers' beliefs: An evaluation of a mathematics methods course. *School Science and Mathematics*, 11(2), 226–232.

Footnotes

ⁱⁱⁱWe additionally tested if motivational math beliefs represented separate constructs by conducting a confirmatory factor analysis. The fit of the model with three factors achieved better fit results ($\chi^2(87) = 179.16, p < .01, CFI = .921, RMSEA = .07$) than a single factor model ($\chi^2(90) = 441.64, p < .001, CFI = .699, RMSEA = .13$; chi-square difference test: $\Delta\chi^2(3) = 262.48, p < .001$).

^{iv}For the dummy predictors, i.e. higher degree and math courses in vocational or further training, reported standardized regression weights should not be interpreted as the differences in means of the motivational math beliefs between subsamples of preschool teachers, for example teachers with higher degree versus no higher degree (Fox 2015).

Table 1

Descriptive Statistics of Likert-Type Scale Measures Ranging from 1 to 4

Scale measure	<i>M</i>	<i>SD</i>	<i>N</i>	<i>α</i>	Items	Example item
Motivational math beliefs						
Value of mathematics	2.89	0.51	197	.73	6	“Mathematics helps to solve daily tasks and problems.”
Math self-efficacy	3.10	0.67	211	.75	5	“How confident do you feel about having to solve the following mathematics problems? Calculating how much cheaper a TV would be after a discount of 30 percent.”
Math self-concept	1.83	0.76	211	.91	4	“I have always believed that mathematics is one of my special/particular/exceptional qualities / strengths.”
School experiences						
Positive emotions	1.98	0.78	203	.89	3	“Which emotions do you associate with your former mathematics classes?” (e.g., “fun”)
Math interest	2.19	0.66	215	.66	3	“In mathematics I wanted not to memorize content but to truly understand it.”
Negative emotions	1.92	0.82	203	.88	4	“Which emotions do you associate with your former mathematics classes?” (e.g., “anxiety”)
Math avoidance	1.87	0.71	212	.76	3	“There were times I did not want to go to school because of mathematics.”

Table 2

Unstandardized Regression Coefficients and Standard Errors for the Prediction of Motivational Math Beliefs by School Experiences, Teacher Qualification, and Work Experience

Predictors	Value of mathematics		Math self-efficacy		Math self-concept	
	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2
School experiences						
Positive emotions	0.15*** (0.04)	0.18*** (0.04)	-0.01 (0.08)	0.03 (0.08)	.29*** (0.06)	.29*** (0.07)
Math interest	0.11# (0.05)	0.09# (0.05)	0.09 (0.07)	0.06 (0.07)	.24** (0.07)	.25*** (0.06)
Negative emotions	0.05 (0.05)	0.05 (0.06)	-0.10 (0.08)	-0.10 (0.06)	-.03 (0.05)	-.03 (0.06)
Math avoidance	-0.07 (0.06)	-0.04 (0.07)	-0.32** (0.11)	-0.26** (0.10)	-.46*** (0.06)	-.45*** (0.06)
Teacher qualification						
Higher degree ^a		-0.03 (0.06)		0.34*** (0.09)		-0.02 (0.07)
Vocational courses ^b		0.09 (0.07)		0.16# (0.09)		-0.02 (0.07)
Further training courses ^b		0.09 (0.08)		-0.06 (0.10)		-0.06 (0.08)
Work experience		0.01** (0.00)		0.01* (0.00)		0.00 (0.00)
R^2	.10**	.16***	.15*	.22**	.44***	.43***
ΔR^2		.06		.07		.01

Note. $N = 203-218$. # $< .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed). Reference groups of binary variables are ^ano university qualifying degree and ^bno course. For the binary variables regression weights should be interpreted as effect sizes (see Footnote 1).

Anhang E - Lebenslauf

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

Anhang F – Erklärung

Hiermit versichere ich, dass ich die vorgelegte Arbeit mit dem Titel „Komponenten und Auswirkung der Qualität mathematischer Bildung in der frühkindlichen Bildung und Betreuung“ selbständig verfasst habe. Andere als die angegebenen Hilfsmittel habe ich nicht verwendet. Die Arbeit ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Datum

Unterschrift