

Structural elements of quality early years provision: A review of the evidence

**Sara Bonetti and
Kristen Brown**

August 2018

EDUCATION
POLICY
INSTITUTE

Research Area:
Early Years
Development



About the authors

Sara Bonetti is Associate Director of Early Years at the Education Policy Institute. Sara is the principle author of the publication, 'The early years workforce: A fragmented picture'. Prior to joining EPI, Sara spent ten years working in the early years field in the United States. She led data collection efforts on topics such as funding and workforce professional development and conducted analyses on areas such as educational leadership and systems integration. Sara's background also includes almost ten years in the field of international development as a project officer and researcher. Sara has a doctorate in Educational Leadership with a focus on early childhood education from Mills College, in Oakland, California.

Kristen Brown is a student in Psychology and Sociology at Yale University. Kristen has an interest in research related to mental health, bias and education and has conducted lab research on topics such as discrimination, prejudice and psychopathology. Kristen joined EPI for the summer 2018 through the Yale International Internship Programme.

Acknowledgements

Jo Hutchinson is Director for Social Mobility and Vulnerable Learners

Natalie Perera is Executive Director and Head of Research at the Education Policy Institute.

About the Education Policy Institute

The Education Policy Institute is an independent, impartial, and evidence-based research institute that promotes high quality education outcomes, regardless of social background. We achieve this through data-led analysis, innovative research and high-profile events.

Education can have a transformative effect on the life chances of young people, enabling them to fulfil their potential, have successful careers, and grasp opportunities. As well as having a positive impact on the individual, good quality education and child wellbeing also promotes economic productivity and a cohesive society.

Through our research, we provide insight, commentary, and a constructive critique of education policy in England – shedding light on what is working and where further progress needs to be made. Our research and analysis spans a young person's journey from the early years through to entry to the labour market.

Our core research areas include:

- Benchmarking English Education
- School Performance, Admissions, and Capacity
- Early Years Development
- Vulnerable Learners and Social Mobility
- Accountability, Assessment, and Inspection
- Curriculum and Qualifications
- Teacher Supply and Quality
- Education Funding
- Higher Education, Further Education, and Skills

Our experienced and dedicated team works closely with academics, think tanks, and other research foundations and charities to shape the policy agenda.

ISBN: 978-1-909274-56-3

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. For more information, visit: creativecommons.org

Contents

Contents	4
Executive summary	5
Introduction	8
What is 'structural quality'	8
Methodology	9
Research questions and objectives	10
Scope of the literature and exclusion/inclusion criteria	10
Data collection and analysis	11
Analysis of the research	12
Workforce training and professional development	14
Message from research	15
Conclusions	20
Child to staff ratio	22
Message from research	22
Conclusions	26
Classroom and/or group size	27
Message from research	28
Conclusions	31
Discussion and recommendations	32
Evidence gaps	32
Policy recommendations/Future research needed	34
Conclusions	35
References	36
Appendices	40
A1. Search terms	40
A2. Analytic review template	41
A3. Coded table and list of studies analysed	45
A4. Brief description of reviews and meta-analyses	49

Executive summary

There is a strong case for investment and intervention in the early years. EPI analysis finds that by the time disadvantaged young people sit their GCSEs at age 16 they are, on average, 18.4 months behind their peers (Hutchinson, Robinson, Carr, Hunt, Crenna-Jennings, & Akhal, 2018) and around 40 per cent of the gap at age 16 has already emerged by age five (Hutchinson, Dunford, & Treadaway, 2016).

The evidence is clear that early years education can play an important role in tackling social disadvantage as well as boosting children's socio-emotional and cognitive development. Even after accounting for the home learning environment and other socio-economic factors such as family income and parent's education level, the EPPE 3-11 study found that pre-school had almost as much impact on children's educational achievement at age 11 as primary school did. The same study also showed that the benefit of pre-school is especially greater for children who are disadvantaged and 'at risk'. While one in three children were 'at risk' of developing learning difficulties at the start of the pre-school, this proportion fell to one in five by the time they started primary school (Sammons, Sylva, Melhuish, Siraj-Blatchford, Taggart, Grabbe, & Barreau, 2007).

The evidence is also clear that for early years provision to deliver its promise, it needs to be of high quality. Yet, what 'high-quality' means is still debated. The existing evidence on what is meant by high quality is dense, poorly understood, and inaccessible to practitioners, commissioners and policy-makers. The research on quality can be distinguished according to two areas: process - what is done in classrooms/settings within the existing resource envelope – and structures – ratios, wages, qualifications, leadership structures, etc. The distinction between process and structural elements drawn by the literature does not imply that they are strictly separate and act in isolation; quite the opposite. Structural elements provide the framework for the elements of process quality to operate and to have the fullest impact on children's outcomes.

Overall, a lack of strong evidence about what works best, and failure to embed this knowledge in the early years system, leads to a lack of coherent structure to problem solving and a sub-optimal use of resources. This report consolidates, summarises and analyses existing research with a focus on the structural elements of early years provision. We have focused on structural elements because they are the most directly observed, measured and regulated ones. In particular, we looked at the so-called 'iron triangle', namely: workforce training and professional development, child to staff ratios and group/classroom size.

The goal of this review is to look at theory(ies) and practice(s) related to structural elements of the early years provision to produce a clear and accessible summary of current research, and to identify areas where the evidence base is already strong and where significant gaps exist.

The texts analysed varied significantly with respect to methodology. While trying to focus on quantitative studies or meta-analyses of quantitative studies, the search found very few carried out through randomised controlled trials or quasi-experiments. The research also varied in terms of the structural factors analysed, with 34 studies looking at teacher training and/or professional development, 13 looking at ratios, and 12 looking at group size. Overall, the evidence analysed in this report demonstrates the importance of structural elements of quality on children's outcomes

across several domains, both socio-emotional as well as cognitive. However, very few studies isolated the impact of each element of the iron triangle. Key findings include:

- A formal degree with at least some specialised training in early childhood education or child development is useful in delivering the skills and knowledge that support optimal teacher behavior. However, a clear strategy to deliver systematic, sustainable and transformative continuing professional development to staff working in different roles is also necessary for training to make a difference for quality and children's outcomes;
- The evidence on child to staff ratios is fairly conclusive: having fewer children per staff leads to better children's outcomes as it provides the opportunity for more individualised attention and leads to better teacher and child behaviour. There is not a golden rule to determine exact ratios for each age group, but research and practice agree on some general guidelines. While early years settings in England are required to adhere to ratios that are in line with these guidelines, ratios in Reception Year classes are generally much higher than what is usually recommended to maximise the impact on children's outcomes;
- International evidence strongly indicates that smaller class sizes for the entire school day are associated with improved children's outcomes, greater educational effectiveness and other benefits at classroom level. Classroom size for children aged from birth to four is not regulated in England but practice seems to be in line with what the international evidence establishes as best practice. What stands in starker contrast is that for children in Reception Year, 30 pupils per class is the norm despite international evidence that clearly points to a maximum average size of 20 children per class for this age group;
- The elements of the 'iron triangle' are tightly intertwined and the effects might get confounded. Each element may be a necessary, but often not sufficient, condition to improve the quality of provision. Relaxing regulations on any one of the three elements might lead to disappointing outcomes.

Most of the research analysed in this report took place outside of England and/or in very specific conditions. At times, results were also potentially different and/or not comparable depending on how 'children's outcomes' was defined (cognitive, non-cognitive, different subcategories of each, etc.) and measured in the study. However, based on the evidence gaps detailed in the report, the following areas were identified as having strong potential for further investigation and trials:

- More rigorous research is needed around the topic of pre-service and in-service training both to understand the value-added of a degree with a specialisation in early childhood as well as to assess whether CPD can be a substitute for pre-service training and, if so, to what extent;
- Additional testing should be carried out on the effectiveness of different staffing structures in terms of levels and types of qualifications, training routes into the profession, and years of experience. A clear focus should also be taken as to understanding how preparation and working conditions of staff working with very young children (aged from birth to two) differ from staff working with three- and four-year-old and Reception Year children;
- Coaching has shown strong promise as an effective model of in-service training delivery. Other models should also be tested, such as supervised teaching or mentoring;
- Very little rigorous research has been carried out to isolate the impact of smaller ratios and class size in nursery and Reception Year classes, e.g. without changing teachers' qualification requirements;

- Specific testing in Reception Year should be carried out, above all to assess the impact of smaller classroom sizes in Reception Year. Recognising the cost implications of such policy, we suggest that different staffing structures in Reception Year could allow for bigger classes to be organised into small groups to achieve smaller ratios at least during certain activities (e.g. those which are more academically focused).

The structural elements considered in this report do not act in isolation. They interact with each other, with elements of process quality and with other factors, such as financing, policies and regulations, entitlement and participation rates. All these elements contribute to the ultimate functioning of the early years system. Nevertheless, this report seeks to provide some clarity on the impact of the three most important elements of structural quality and, whenever possible, to draw significant conclusions from the existing evidence. Ultimately, it strives to provide actionable recommendations for how to prioritise research and investments in the early years.

Introduction

What is 'structural quality'

It is widely accepted that for early years provision to improve children's outcomes it needs to be of high-quality (Cunha, Heckman, Lochner, & Masterov, 2006; Felfe & Lalive, 2014; OECD, 2015; Siraj-Blatchford, Taggart, Sylva, Sammons, & Melhuish, 2008; Smith, Grima, Gaffney, Powell, Masses, & Barnett, 2000; Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2011). However, the definition and features of quality are still debated - and greater clarity is needed to improve both policy-making and spending decisions. Different stakeholders might use a different definition of, or place varying weights on, the elements that contribute to a high-quality provision. Smith et al. (2000) point out that 'many viewpoints on quality can be identified even within one cultural setting - for example the child development, the government or regulatory perspective, the welfare or social services perspective, the parent perspective, the cultural perspective, the social policy and funding perspective, and the child's perspective. How quality is defined depends on the concern of the stakeholder' (pp. 48-49).

In the last few decades, research has made big steps towards the identification of some more universal elements of quality. Initially, early childhood education research was mainly about interventions directed at improving children's cognitive development and school achievement, and at reducing delinquency. It also looked into whether early child care could have a harmful effect on children. A second wave of research, recognising the shortcomings of the previous one, moved onto the topic of quality of the early years settings, from which a number of assessment tools emerged, such as the Early Childhood Environment Rating Scale (ECERS) and the Infant/Toddler Environmental Rating Scale (ITERS). Fenech (2011) comments that 'reviews of this body of research highlight generally accepted understandings (at least in the Western world) as to what constitutes key elements of quality' (p. 103). The third wave of research went further by acknowledging that there are other factors contributing to children's outcomes beyond the early years setting, such as the family environment and the child characteristics (Fenech, 2011). Looking at the complex intersection of all these elements, this strand of research has focused on the nature of quality, in an attempt to identify variations across settings and their impact on children's outcomes (Smith et al., 2000). In this respect, Smith et al. (2000) define quality as 'the essential components of early childhood environments which are valued in our society, and which support the well-being, development and rights of children, and support effective family functioning' (p. 48).

It is this more recent strand of the literature that brought forward the distinction between process quality and structural quality. Process quality relates to the more proximal features of early years provision, such as the presence of a stimulating, developmentally appropriate environment; positive interactions between staff and children and among peers; and parental involvement. More specifically, process quality involves 'social, emotional, physical, and instructional elements of interactions with young children, elements that are reflected at several levels of the classroom environment: moment-to-moment displays of discrete behaviors as well as global characterizations of the overall setting' (Pianta et al., 2005, p. 145). Structural quality relates to inputs that are more easily observed, measured and regulated (Slot, Leseman, Verhagen, & Mulder, 2015; Smith et al., 2000).

Elements of structural quality generally agreed upon in the literature are:

- Staff to child ratios;
- Group size;
- Staff training, education and experience;
- Staff wages and working conditions; and
- Staff stability.

Other studies include additional features, such as: adequate indoor and outdoor space, health and safety provisions (Fenech, 2011), and management practices (Munton et al., 2002).

While process and structural quality are not strictly separable, this literature review and the subsequent policy recommendations focus solely on structural quality. Structural quality enables the elements of process quality to operate and have the fullest impact on children's outcomes (Fenech, 2011). In fact, some research stemmed from the premise that there exists a mediated pathway that links structural elements to process elements to developmental outcomes for children (NICHD Early Child Care Research Network, 2002). Our findings will reflect this as they show that it is often hard to disentangle the elements of process quality from those of structural quality.

Even within the realm of structural quality, it is often difficult to isolate the impact of each element. For example, ratios, group size and training - usually called 'the iron triangle' of quality because many other elements of quality depend on those - are often intertwined both in government regulations as well as in their impact (Slot et al., 2015; Smith et al., 2000). Research shows that these three variables have a direct impact on the ability of staff to provide sensitive, responsive care for children. They also happen to be the elements that are more easily regulated through government legislation. Staff salaries and management practices also affect the quality of provision and children's well-being through their impact on staff turnover. In settings with high turnover rates it is less likely that children receive consistent care, another element strictly connected with quality provision (Mims, Scott-Little, Lower, Cassidy, & Hestenes, 2008; Munton et al., 2002). Finally, salaries and management practices also have a direct impact on staff well-being and job satisfaction (Munton et al., 2002). Howes, James and Ritchie (cited in Whitebook, 2003, p. 16) report that within a group of primarily African American and Latino teachers with less than four-year degrees working in high-quality programs serving children of low-income families, teacher responsiveness was predicted, after controlling for formal education, by such factors as staying in the field for the sake of benefiting one's community, being mentored early in their careers, and receiving ongoing supervision.

Many studies do not isolate the effect of each element, in part because it is difficult, in part to acknowledge that they are tightly interconnected. Nevertheless, in the interest of making our policy recommendations actionable, we have tried to distinguish the impact of each element of structural quality considered in our review.

Methodology

To find and synthesise key sources related to structural quality in early years provision we carried out a literature review of a variety of publications, from peer-reviewed journal articles to books, from reports to policy briefs. One of the main purposes of a literature review is to discover what research has already been conducted on a topic and to identify gaps in the current knowledge base;

therefore, it fitted with the goal of our work (Greenfield, 2002). While resource constraints prevented us from undertaking a systematic review, we strove to give our work some structure by formalising our searches and analyses in a way that is closer to that of a systematic review. Below, we will detail our methodology.

Research questions and objectives

The following research questions guided our review:

- How has structural quality been addressed and treated in the literature?
- What elements of structural quality have (the potential of) a positive impact on children's outcomes?

For our conclusions and recommendations to be appropriate for the English context, our analysis looked at the literature through the lenses of these additional questions:

- How is this literature relevant to the English context?
- How does the structure of early years provision in England compare to the countries mentioned in the literature?
- Have cross-country differences contributed, or the potential to contribute, to different outcomes? If so, how?
- Is there similar research done in England? If so, are conclusions the same? If not, what evidence gaps need to be filled?

The goal of this work was not to critique these studies, their methodologies and/or their conclusions, but to identify which topics have already been thoroughly investigated, which hypotheses have only partially been studied and would benefit from further testing, and, finally, which ones are still virtually absent in the realm of empirical research.

The specific objectives of our work were:

- to synthesise existing literature and identify areas where the evidence base is already strong, and practice should be embedded in/by the sector (what we know and how we know it);
- to identify areas with a sound theoretical basis where practice is more developed but further, rigorous evaluation is needed or would be ideal (further evidence needed);
- to identify areas underpinned by a sound or emerging theoretical basis of child development which are not yet the subject of well-evaluated practice but should be proactively developed (evidence gap).

Scope of the literature and exclusion/inclusion criteria

In order to make the best use of the time available for this review and provide adequate depth, we focussed on the elements of 'the iron triangle': workforce training and professional development, child to staff ratios and group/classroom size. We recognise that other components of structural quality are also important.

In addition, we looked at international literature published in English. We acknowledge that the focus on English as language limited the breadth of our conclusions, but we tried to compensate by

gathering and analysing literature with an international breadth. For each item, we considered if, and how, the study was relevant for the English context.

Finally, the early years span from age zero to five in England but some strands of the international literature span until age seven or eight. Therefore, we kept our search open to children aged zero to eight. When appropriate, we looked at whether results differed depending on socio economic status, race and ethnicity, urban versus rural living, education needs and disabilities, and age ranges within the zero to eight continuum (e.g. zero to two, three- and four-year-olds, Reception Years, infant school years). We looked into whether structural elements of early years work in a significantly different way across these variables. However, we did not target our search to any one of these characteristics.

Data collection and analysis

Our work consisted of a three-layered analysis. First, a search and filtering process was created to identify the most relevant literature for our study within the scope described in the previous section. Second, a template was developed to analyse each text in terms of how structural quality was conceptualized, whether the study looked at the impact on children's outcomes directly or indirectly through, for example, teaching quality, process quality or classroom quality, and how, and to what extent elements of structural quality were considered in the context of the wider early years and education system. Finally, a coding system was developed to describe the key features of the papers included in the review.

We carried out our searches through Google Scholar. While limiting compared to the use of multiple search engines and databases, Google Scholar allowed our search to include not just peer-reviewed articles but also grey literature (that could help us contextualise the studies) and meta-analysis and/or literature reviews (that allowed us to access – albeit in a summative form – the results of many more studies than we would have had the capacity of explore ourselves).

We searched title and full text. Initially, the title search usually resulted in no or too few items, while the full text search resulted in too wide of a results set and in a compilation of texts that were not focused enough. For example, searching Google Scholar for *“early years” OR ECEC OR “early childhood education” OR ECE OR “early care and education” OR “early education” structures outcomes* returned 0 items through title search and 17,900 items through full text search; the search *“early years” OR ECEC OR “early childhood education” OR ECE OR “early care and education” OR “early education” “structural quality” outcomes* gave us 0 (title) and 1,650 (full text) results; finally, searching for *“early years” OR ECEC OR “early childhood education” OR ECE OR “early care and education” OR “early education” “structural characteristics” outcomes* returned 0 and 6,140 results respectively. Hence, we decided to use more specific search words for the different elements of structural quality (see Appendix A1 for the full list of search terms), while running a higher number of searches. Moreover, we limited the search to items published from 2000 onwards and we excluded patents and citations, as well as texts not written in English.

A second level of filtering was applied by eliminating texts that did not mention *any* of the following terms, or variations thereof, in the title and/or the abstract: (a) early years, early childhood education, or early care and education; (b) structural quality, ratios, staff/teachers/workforce training, staff/teachers/workforce education, or group size; (c) outcomes, children's outcomes,

socio-emotional outcomes, cognitive outcomes, process quality, teaching quality, child-teacher interaction, or classroom quality.

The third level of filtering was completed by reviewing abstracts and/or introductions and aimed at determining the general focus of the text. Texts that did not focus on, or mention, any of the key terms considered in the second level of filtering were eliminated, as well as texts that did not explicitly aim at analysing structural quality in any depth. Moreover, our search was for texts focused on children aged zero to eight. Texts related exclusively to other stages of the education system only were excluded. Finally, any duplicate was removed at this point. This last filter left us with 69 results, of which 17 we could not gain access to. Hence, the search through Google Scholar gave us a total of 52 texts to analyse. We also re-ran some of the searches specifically targeting studies that were randomised controlled trials and/or quasi-experiments. This search gave us an extra 20 texts, of which there were four we could not gain access to. Therefore, we were left with 16 texts to analyse. Finally, we also included a small number of items that were suggested as key readings by experts in the sector. Ultimately, our analysis focused on 92 resources: 61 peer-reviewed articles, 18 reports, eight policy briefs, three books, one discussion paper and one non peer-reviewed article.

Analysis of the research

After the texts were selected, the analytic review template was created as a guide to analyse the texts both qualitatively and quantitatively (see Appendix A2). The dimensions of our analysis included: the type of resource (book/book chapter; peer-reviewed article; non peer-reviewed article; thesis or dissertation; report); the type of research (literature review; empirical study; theoretical/position paper; practical perspective; etc.); the methodology used (qualitative; quantitative; case study; ethnography; discourse analysis; other); the structural elements considered in the text and their impact on children, classroom, teaching and process quality; and whether, and to what extent, the study was relevant to the English context. Finally, we looked at whether, and to what extent, other types of ‘outcomes’ were considered (adult, family, society, school, school system outcomes), and whether, and to what extent, macrocontextual issues/factors are addressed. The structural elements investigated with this literature review are affected by macro elements that are specific to each education system. Examples of these extra elements we considered are:

- Government (policy orientation, priorities and approaches, e.g. child-care versus education divide; market versus public provision; types of ECEC settings that exist in a country);
- Financing and costs;
- Curriculum;
- Initial teaching training (ITT) Systems;
- Policies and procedures (e.g. benefits system, parental family leave policies, entitlement to ECEC – targeted versus universal delivery).

While it was beyond the scope of this literature review to tackle these topics in detail, we deemed it important to keep track of whether a text addressed them and to what extent their impact on other structural elements of quality and, ultimately, outcomes were considered.

The information collected with the analytic review template was used to create a database where the results of our analysis were stored and coded (Appendix A3). This allowed us to identify any patterns across texts in terms of type of research, methodology used and research findings, as well as to identify potential evidence gaps.

The 34 texts analysed through the analytic template varied significantly with respect to methodology. We tried to focus on quantitative studies or meta-analyses of quantitative studies, but we found very few carried out through randomised controlled trial or a quasi-experiment. Overall, 12 studies were quantitative, 12 were qualitative, six were mixed methods and four were experimental analyses. We focused on studies carried out since the turn of the century, with 18 studies published from 2000-2009 and 16 studies published from 2010-2018. Finally, our research varied in the structural factors analysed, with 34 studies looking at teacher training and/or professional development, 13 looking at ratios, and 12 looking at group size. However, many of these studies examined more than one structural factor. Seven studies considered all three of these factors, 12 considered training and professional development and ratios, and eight considered teacher training and professional development and group size.

The literature we examined documented two types of associations: the one between structural and process features of child care quality and that between structural features and children's outcomes. Our original goal was to focus on the latter, but we decided to expand to different types of associations because of the otherwise limited literature. At times results were potentially different and/or not comparable depending on how 'children's outcomes' was defined (cognitive, non-cognitive, different subcategories of each, etc.) and measured in each study. We kept track of the differences in our template and made each study's approach clear in the full discussion of the literature.

In the next three chapters we will provide a detailed description of the literature about the three elements of the 'iron triangle': workforce training and professional development; child to staff ratios and group/classroom size. To the extent that this was possible, we organised the studies according to whether they looked at the impact of structural elements on children's outcomes or on various elements of process quality. In the last chapter, we will focus on the general findings and on where evidence gaps are manifest and provide a clear direction for concrete policy recommendations.

Workforce training and professional development

The topic of workforce training and professional development is perhaps the most complex and controversial to address among the structural elements of quality. It is generally recognised that a qualified early years workforce is key for high quality provision (Minervino, 2014; Sylva et al., 2011; Whitebook, 2003; Workman, Guernsey, & Mead, 2018). What is not always clear and agreed upon is what 'qualified' means in practice. Research is still trying to answer questions, such as: What are the optimal teacher qualifications? What is the 'minimum level' that guarantees good provision? If a country is not at that level yet, what is needed to achieve it?

Many long-standing experts support the idea that optimal teacher behaviour and the skills and knowledge it rests upon are best achieved through a formal degree, which includes at least some specialised content in early childhood education or child development (Pianta, et al., 2005; Whitebook, 2003; Workman et al., 2018). While the evidence on the validity of these claims might be clear in principle, the practicalities of early years staff obtaining a degree, particularly under constrained resources, are not. Little is known about different thresholds of education levels and possible combinations within a teaching team, or about the value added of an advanced degree, or the role of the work environment in building upon formal training. We will address some of these topics below.

In conducting the literature review and extrapolating evidence gaps and policy recommendations, we had to keep in mind several important factors.

First, researchers have found it difficult to collect reliable, consistent and/or comparable information about training that could help understand how the amount, intensity, content and quality of instruction impact its effectiveness. This might be due to the nature of the data. For example, teachers and providers themselves are often unable to recall their educational and training histories in detail, as they might have engaged in a wide range of professional development activities for many years. Other times, poor quality data derive from differences across studies, where some look at the number of courses a teacher has taken over the course of a career and others over a determined amount of time. Finally, few studies focus on the actual content of training (Whitebook, 2003).

Secondly, empirical studies are usually very contextual, with the wide majority taking place in the USA. Munton et al. (2002) comment that 'the types and levels of qualification found among early years workers in different countries vary considerably. Differences are related to a wide range of issues, in particular how the workforce and the services themselves are structured; and how early childhood work, and therefore the role of early childhood worker, is understood' (p. 8).

Finally, when discussing the relationship between teachers' preparedness and quality of provision and/or children's outcomes, it is sometimes erroneously put forward that the evidence is inconclusive. The claim comes from the fact that a small subset of studies showed mixed results (Early et al., 2007; Pianta et al., 2005). Yet, none of the authors of these studies came to the conclusion that a qualified workforce was not needed. Instead, they acknowledged that results differed depending on the specific elements of qualifications that were considered in their study (as well as their interaction with other structural elements in place in that country, programme or classroom) and/or different definitions of children's outcomes used. Instead, they suggest that 'a

new era of research is needed to address the complexity of teacher quality' (Early, et al., 2007, p. 576) and that to more fully understand factors associated with classroom quality, 'the field must move beyond making simple associations between teachers' education level and the quality of classrooms. Classrooms and teachers are one component within a larger education system' (Mims et al., 2008, p. 11).

As we mentioned in the introduction, studies can be distinguished according to whether they look at the impact of different elements of teachers' qualifications and training on children's outcomes on process quality. Only one item (NICHD Early Child Care Research Network, 2002) looked into the mediated path from structure to process to children's outcomes. Below, we present the evidence that emerged from our review of the literature.

Message from research

Impact on children's outcomes

The vast majority of the texts we analysed found that staff training and professional development - intended as both pre-service, or initial teaching training (ITT), as well as in-service, or continuing professional development (CPD) - are associated with increased positive outcomes for children. Whitebook's (2003) review of the literature on the relationship between teacher preparation and child outcomes includes most of the major, large-scale investigations of child care settings conducted until then. Despite the limitations presented in terms of sample, measures and analytic methods of the reviewed studies, the author concluded that the results 'underscore the importance of more higher education and specialized training, and identify the particular role of the bachelor's degree, most often in early childhood education, in producing teacher behaviors consistent with high-quality programming, which in turn supports better developmental outcomes for children' (p. 16). However, the author also introduced some complications that arise when translating these results to practice, such as the ability to increase compensation relative to qualifications, the capacity of higher education institutions to handle an influx of pupils seeking early years qualifications, the potential erasure of linguistic and cultural diversity from the early years workforce, and the costs involved in requiring higher level of education for every teacher.

Among more recent reviews, Werner, Linting, Vermeer and Van IJendoorn's (2016) meta-analysis of randomised controlled trials found that children's outcomes, such as social-emotional development and communication skills, were improved by targeted interventions to increase childcare quality, where these targeted interventions were specific, short-term caregiver/teacher training. Scobie and Scott's (2017) literature review also concluded that across a number of U.S. early years programmes, higher qualified teachers and staff, good working conditions, continuous professional development, fair pay and an age appropriate curriculum were all important factors in improving cognitive and non-cognitive outcomes.

Howes (analysed by Munton et al., 2002, p. 53) studied the relationship between staff to child ratios and the educational background of the classroom's lead teacher (i.e. formal education and early childhood education training) with teacher behaviour, children's activities and outcomes. The main conclusion reached by the author was that the higher the level of education of the lead teachers the more effective their teaching was. Children in classrooms with teachers who had at least a degree in Early Childhood Education or a specialised credential engaged in more complex play with peers and in more creative activities. However, the study also pointed to the fact that while both elements

(low child to staff ratio and high level of qualification) are necessary for quality, each on its own is not sufficient. Teachers with a degree or a specialised credential were not more effective when assigned to more children than teachers who had lower level of education but worked with more favourable ratios. Similarly, Wylie et al. (analysed by Munton, et al., 2002, p. 84) looked at the impact of early years provision on developmental outcomes for 307 four-year olds in New Zealand and found that children in low quality settings had poorer developmental outcomes than those attending high quality one. Quality was again determined by a variety of structural elements, such as staff education level, staff salaries, ratios, group size and type of service.

Research on staff qualifications also investigates how training and continuous professional development can be used to supplement a teacher's formal education. Using a randomised controlled trial, Bierman et al. (2008) studied the impact of the Head Start Research-based, Developmentally Informed (REDI) program, which provides teachers with strategies to promote social emotional competencies and language development in their classrooms. Specifically, it consists of brief lessons, extension activities and specific teaching strategies that have been shown to be linked to social-emotional competences and language and literacy skills. The study explored a number of child outcomes, finding improvements in vocabulary skills, social-emotional skills, learning engagement, phonological awareness and language usage at home, but no significant differences in print awareness, reading involvement at home, aggression at home or attention problems at home.

Studies that examined intensive teacher training programmes often found significant results as well. Rhodes and Hennessy (2000) investigated the impact of a 120-hour preschool training course on teachers' behaviour and children's development and found that teachers who received training made significant gains in positive relationship and decreased in levels of detachment. Children gained significantly both in terms of social as well as cognitive play. The comparison group adults and children showed no significant improvements from pre- to post-test times. However, the small sample size of this study (33 caregivers and 68 children) should caution against generalising the results. The same authors mentioned other studies that found an association between training and children's social development, such as: The National Day Care Study (1979), which found that children in classes with trained caregivers showed more engagement, longer attention to activities, and more cooperation and compliance; Clarke-Stewart et al. (1994), who found that more aggressive children tended to have teachers who had less training in child development; and Kontos, Hsu and Dunn (1994), who found a positive association between the amount of specialized training caregivers received and the amount of children's complex social play.

Additionally, Buysse, Castro and Peisner-Feinberg (2010) conducted a randomised controlled study of a very intensive, ongoing professional development programme targeted at kindergarten teachers working with Latino dual language learners (DLLs) children whose primary language is Spanish. They found that the programme was very successful as it led to measurable improvements for both teachers (increasing overall quality of teachers' language and literacy practices as well as those specific to working with Latino DLLs) and children (providing greater gains in children's phonological awareness skills in their primary language). It is important to remark, though, that the programme was very targeted and intensive and that, in a way, it can be seen more as an element of process quality rather than structural quality.

Finally, Chambers, Cheung, Slavin, Smith and Laurenzano's (2010) systematic review looked into a variety of studies on the effectiveness of early childhood programmes and found similar positive impacts on children's outcomes. Importantly, the authors point out that, aside from the programme's content, what made a difference was the support that teachers received. In fact, 'In most of the studies reported here, teachers received more support for implementation of the programme than teachers typically receive when implementing a new programme. In practice, teachers often receive very little support, perhaps just a teacher's manual with suggested activities. In some of the research studies summarised here, they received extensive initial training and very frequent follow-up coaching by the developer or researchers, which may not be typical when the programme is implemented at scale' (p. 60).

However, not all research has arrived at these conclusions. Burchinal et al (analysed by Munton et al., 2002, p. 83) looked at the relationship between childcare centre quality and cognitive and language development for a sample of 79 African-American one-year-olds and found that, after controlling for child and family characteristics, neither staff education nor group size was significantly related to infant developmental scores. They did find a modest correlation between global measures of provision's quality and infant development, but the sample size is small and variability in staff education levels is low, and likely causing the lack of significant relationship, as well as making generalisation of results very limited.

Blau (cited in Munton et al., 2002, p. 82) looked at the effects of group size, adult:child ratios, staff training and other characteristics of child care on child development using data from the National Longitudinal Survey of Youth (NLSY). The author found mixed results, but there are several issues with this study. First, the data were collected at irregular intervals and, as the author recognised, susceptible to measurement errors. Also, the effects of several structural elements are confounded. Finally, the study used a very rudimentary measure of staff training, represented by a dichotomous variable indicating whether the members of staff had any specialised education and training, such as early childhood education, special education, or childhood psychology.

Finally, Early et al. (2007) reviewed seven major studies of early care and education and looked at whether educational attainment and subject of teachers of four-year-olds predicted classroom quality and children's academic outcomes. They found null or contradictory associations. However, the study presented several limitations, the crucial being that the authors looked only at short-term cognitive outcomes and disregarded socio-emotional and long-term outcomes. Moreover, the authors did not use these findings to draw the conclusion that education levels do not matter. Rather they pointed out that we do not have much information regarding the quality of training programmes as well as that what is needed is a more comprehensive approach to teachers' preparation that includes both high quality pre-service training and continuing professional development.

Impact on process quality

As we mentioned in the introduction, strictly separating process and structural quality is often not possible, nor appropriate. In fact, a high proportion of the literature on quality looks at the connection between the two and tries to test and measure the effect of structural elements on process quality, rather than directly on children's outcomes. Education and training has been shown to have an impact on teachers' behaviour and practice and, therefore, on process and classroom

quality through improved adult-child interactions (Mims et al., 2008; Minervino, 2014; Mitter & Putcha, 2018). Below we summarise additional literature on this topic and recall previously mentioned studies that also looked at the impact on process quality.

The evidence of the relationship between workforce training and professional development and process quality seems to be much clearer and more conclusive than the evidence explored in the previous section. Studies discussed in Friendly, Ferns and Prabhu's (2009) review found that the amount of teachers' formal schooling was correlated with higher quality and better teacher behaviour. Furthermore, specialised education in early childhood or child development was causally related to quality. In a study looking at the impact of structural elements of quality on classroom quality and child-teacher interactions, Pianta et al. (2005) found that the combination of more than 60 per cent of classroom being from homes below the poverty line, teachers lacking formal training (or a degree) in early childhood education and teachers holding less child-centred beliefs was associated with lower quality. In fact, programme and teacher attributes were statistically significant predictors of observed quality. Additionally, Hayakawa and Reynolds (2014) reviewed evidence of key elements and strategies of effective early childhood programs, finding that teacher training improves teacher effectiveness and teacher-child interactions. Teachers with bachelor's degrees demonstrated more warmth and enthusiasm and were less punitive. Finally, Anne Smith (cited in Smith et al., 2000, p. 56) isolated the impact of teachers' training from that of other structural elements. Both the length of the training as well as the number of trained staff with high qualifications in the setting mattered. In fact, the percentage of staff with only three years of training was the strongest negative predictor of children wandering and waiting.

Other studies focussed on in-service training. Werner et al.'s (2016) study of teacher training interventions showed that in early years settings 'the effects of specialised in-service training on process quality are larger than those of pre-service training, particularly on collaborative work, support for play, and support for early literacy, mathematics and science' (p. 12). The authors concluded that these training interventions were moderately effective in improving classroom quality - which included classroom atmosphere, child support and instruction and level of conflict between children - as well as caregiver interaction skills - which included practices, attitudes, beliefs or knowledge about caregiving. Additionally, a randomised controlled trial carried out in Head Start programmes in the United States found that an evidence-based curriculum workshop training called REDI resulted in improved classroom management, and linguistic support and emotional climate, but no difference in instructional support or positive discipline (Dimitrovitch et al., 2009).

Another randomised controlled trial examined which kinds of CPD and teacher training are most effective, comparing traditional training through coursework, coaching and a wait-list control group. While the groups were not significantly different in terms of post-test teacher knowledge, interactional environment, support for learning, or teaching strategies, the coaching group improved the overall literacy environment more than both the traditional training group and the wait-list control group (Neuman & Wright, 2010). Additionally, the authors conducted a series of interviews in which teachers reported their experiences in the study. They found that the literacy demands for the traditional training were high, and many of the concepts were too abstract to translate into everyday practice. Meanwhile, the coaching intervention provided real-time, context-specific, individualised feedback as well as accountability. Milburn, et al. (2015) also studied the impact of coaching through a randomised controlled trial, dividing educators into a coaching and workshops

intervention and a workshops-only comparison group. Though the intervention group did not differ from the comparison group in references to print text, as observed through a videotaped classroom activity, the coaching and workshops intervention group used references to phonological awareness at significantly higher frequency. Similarly, Wasik and Hindman (2014) examined a series of randomised controlled trials comparing workshops and coaching to workshops alone. While the addition of coaching resulted in only nearly significant impacts on children's outcomes, it significantly increased teaching quality, specifically the extent to which teachers used contextualized references to target vocabulary.

Slot et al. (2015) investigated the impact of training and professional development activities of teachers working in Dutch child care centres. They found that teachers' formal pre-service education has a positive, but small, association with emotional process quality, which includes measures of emotional support, facilitation and enrichment of play, and self-regulation. However, this conclusion must be inserted in the specific context of the study, as government regulations of structural elements are quite different from England. For the case of teachers' training, for example, the Dutch government requires all teachers to have completed a minimum of three years vocational training in a relevant subject. Consequently, the variation in education level is restricted, which in turn can explain the lack of stronger effects, as the authors suggested.

Some of the studies on training and professional development also consider staffing structures. Whitebook (2003), for example, while recognising the importance of a specialised degree, also concluded that this does not imply that all staff need a degree; instead 'a standard might be set at one BA-level teacher per classroom, or for a certain number of children (e.g., 20), with this teacher working with one or more assistant teachers' (p. 3). Kagan and Neumann (cited in Smith et al., 2000) compared the finding of three large U.S. studies relating training to outcomes: the National Day Care Study (focused on infants and toddlers), the Cost and Quality study and the National Child Care Staffing Study (focused on in-service training). All three studies concluded that both general education and early childhood training had positive effects on quality and one study showed that in-service training had a positive effect. The authors emphasised that an important aspect of early childhood centre quality was that staff should possess a diverse range of educational and training backgrounds, skills and abilities. Good staff training policies should assess the individual training needs of staff and 'attend to the professional development of teaching staff in ways that are inventive, individualised and effective' (quoted by Smith et al., 2000, p. 57). On the other hand, Marshall et al. (2001) used a variety of data collection methods - from classroom observations to interviews - to measure the quality of care in Massachusetts community-based centres and concluded that among the three elements of the iron triangle, staff education was the most strongly associated with process quality. In addition, they found that 'the use of more teacher hours in the classroom, rather than staffing with assistant teachers, was associated with higher process quality ratings' (Marshall et al., 2001, p. 35).

Finally, we need to differentiate between frontline and leadership staff as they not only might have different education levels and specialisations but, to some extent, they are likely to have different needs in terms of CPD. In addition, when trying to measure the impact of such training, researchers need to keep in mind that the mechanisms through which they exert an impact might be different. For example, Mims et al. (2008) looked at the impact of the education levels of teachers and directors in the North Carolina Rated License Process on classroom quality and at the relationship

between stability of position (i.e., consistently working with the same age group) for teachers and classroom quality scores. Teacher education level and stability were positively correlated to classroom quality scores. In addition, higher education levels for directors and their enrolment in a college course were positively related to programmes' quality scores, therefore pointing to the fact that the level of education of different types of staff within a centre plays a role in driving quality of provision. Mims et al. (2008) also recalled that data from the Cost, Quality and Child Outcomes study 'revealed [that] a director's effectiveness and involvement with a center's curriculum, years of experience, and education level were related to quality scores. In a discriminant analysis from this study, the director's education level was found to be highest in good quality centers, lower in mediocre centers, and lowest in poor-quality centers', therefore showing a clear association between director characteristics and quality (p. 5).

On the mediated path from structure to process to children's outcomes

One of the studies we reviewed was unique in its approach. Instead of looking at the impact of structural elements on either children's outcomes or process quality, the NICHD Early Child Care Research Network's (2002) study tested the mediated path from structural elements through process elements to children's outcomes. Process measures of child-care quality assessed the caregivers' relationship with the children (sensitivity to nondistress, detachment, stimulation of cognitive development and intrusiveness) as well as the classroom (chaos, overcontrol, positive emotional climate and negative emotional climate). Structural quality was evaluated through caregivers' training and child to staff ratios. Cognitive competence was measured as composed by seven measures well established in research: Incomplete Words, Memory for Sentences, Letter Word Identification, Applied Problems, Preschool Language Scale, Auditory Competence, Expressive Language, and a measure of errors derived from lack of sustained attention (p. 201).

The authors tested their hypothesis of a mediated path through four models in order to account for different cognitive competence and caregivers' rating of social competence and found that caregiving training and child to staff ratios had a significant indirect effect (p. 204). In addition, 'the child-care effects, both direct and indirect, do not appear to be due to family selection of child-care quality per se'. In other words, the study showed that while family selection does exist, as the models were better fitted when indirect paths from family variables to child care variables were included, it did not account entirely for the path from child-care structure to process and outcome (NICHD Early Child Care Research Network, 2002).

Conclusions

The literature on workforce training and professional development is straightforward in terms of general conclusions and recommendations: highly qualified staff are key to high-quality provision; a degree with a specialisation in early childhood is important but not sufficient on its own; and in-service training is necessary for both new and experienced staff. Nevertheless, the topic is complex from an implementation point of view. Pre-service training, induction and in-service training are interconnected. Just as researchers agree that a degree in early childhood prepares teachers in a way that is more impactful for children, they also know that a degree should not just be a label and that, in itself and on its own, it does not guarantee quality teaching and caring if it is not accompanied by strong induction and CPD practices. In fact, many studies we found looked at all these components together, making it hard to disentangle the impact of each one of them alone.

In the past few years, more efforts have been devoted to understanding better practices, mechanisms and formats of CPD. Recent studies point to the need for professional development to possess three key characteristics in order to make a difference for quality:

- Systematic: there needs to be a system-wide approach to CPD, which includes different paths for staff at different levels as well as sufficient funding;
- Sustainable: all staff and all types of settings need to be able to access CPD and to afford employing qualified staff;
- Transformative: one-off or short courses and workshops might be effective in tackling specific topics but are not evidenced to have the most impact on changing practices and affecting children's outcomes. On the other hand, CPD that happens over time, is based at the setting where staff is employed, brings in outside expertise, and encourages collaboration between staff seems to be more effective. Coaching, in particular, seems to be a very promising approach (Waters & Payler, 2015).

Finally, this renewed focus on in-service is partly due to the fact that, in many countries, requiring all staff to have a degree is not politically and/or practically feasible. The cost impact is high, and many education systems are not equipped to handle such a transition fast enough to make sure all children currently attending early years provision will benefit. As Whitebook commented (2003), 'there is emerging evidence that alternative pathways to effective teaching exist, and may be important for increasing and diversifying the corps of skilled instructors of young children' (p 16). And yet, the evidence reminds us that just like a degree by itself is not enough, professional development should not be considered a cure-all to shy away from higher investment commitments in the early years.

Child to staff ratio

The literature on child to staff ratios is uncontroversial, with its goal usually being to provide supporting evidence for the generally recognised idea that smaller ratios (i.e. fewer children per staff) are better. The main mechanism through which ratios influence quality is through adult-child interactions. Evidence is consistent in showing that having more staff is more likely to facilitate positive interactions, increase the frequency of adults playing with and talking to children, as well as the complexity of children's play, assuring a more appropriate caregiving style (less restrictive and controlling, more physical contact and smiling) and greater job satisfaction (Camilli, Vargas, Ryan, & Barnett, 2010; Munton et al., 2002; Smith et al., 2000).

Less important has been the focus on identifying a golden rule in terms of what the ideal ratio should be. There are, however, some internationally accepted norms. The younger the children, the smaller the ratios should be. Smaller ratios are also likely to be better for children with disabilities and children from disadvantaged backgrounds (Smith et al., 2000).

There are several caveats to these general statements. These caveats do not negate the general results of the literature but need nevertheless to be kept in mind to manage expectations of what exactly the research is showing, how to upscale effective programmes and/or how to translate these programmes into different contexts.

First, ratios are sometimes defined and/or measured differently. Some studies use the total number of childcare places available divided by the number of full-time equivalent staff employed. Others use the number of places by the number of staff at work at a point in time. Finally, some use the number of staff and children actually observed in the same area over a given period of time (Munton et al., 2002). These differences could bear important consequences in terms of the interpretation of the findings.

Second, as mentioned in the previous section, the influence of ratios on quality is inextricably linked to other elements of the environment, including staff education and training, staff salaries and group size ('the iron triangle'), and it is often difficult to isolate the impact of ratios from the impact of these other factors (Munton et al., 2002; Smith et al., 2000).

Third, changes creating stricter ratios can lead to children being organised into larger groups across activities or larger classrooms, potentially compromising the benefits deriving from smaller ratios. Consequently, some authors suggest that 'regulations specifying adult:child ratios should also address the issue of group sizes' (Munton et al., 2002). We will touch on this in the sections below.

Despite these caveats, the evidence is clear and provides some important conclusions, which we proceed to summarise below.

Message from research

Impact on children's outcomes

Across the analysed studies, the impact of lower child to staff ratios on children's outcomes was mostly positive, though these results were often discussed along with other aspects of structural quality. Using seven different measures of cognitive competence as well as assessments of social competence, the NICHD Early Child Care Research Network's study (2002) found that the

combination of lower ratios and more advanced caregiver training indicated higher structural quality, which was associated with increased cognitive and non-cognitive outcomes through improved process quality. Importantly, this was observed even after accounting for the potential confound of family selection.

These findings are supported by Mitchell, Wylie and Carr (2008), who discussed three main studies on the direct impact of ratios on children's outcomes. Karoly, Kilburn and Cannon (2005) reviewed 20 programmes, identifying three features which had larger or longer-lasting positive outcomes for children: better trained staff, smaller ratios and greater intensity (Mitchell et al., 2008).¹ However, these elements are so interrelated that untangling their individual impacts is difficult. In fact, Frede (1995) explicitly discussed their overlap, suggesting that intensity includes 'the concentration that comes from low ratios, home visiting, and coherent curricula' (cited in Mitchell et al., 2008, p. 123). Additionally, Gilliam and Zigler's (2000) statistical meta-analyses found that a U.S. government-funded pre-school programme, which reported mixed results generally, fared much better in states with higher regulatory standards for staff training, group size and staff-child ratios (Mitchell et al., 2008, p. 21). This suggests that these three factors may be instrumental in the successful implementation of high-quality early years education.

It is also an indication of how hard it is to disentangle the effect of each of the three elements. None of these studies analysed child to staff ratios in isolation. However, research on 'individuation' (or 'individualised' instruction) and learning in small groups clarifies the role that ratios have in the classroom and provides evidence regarding their effectiveness separate from other aspects of structural quality. Camilli et al. (2010) found that 'individuation (...) had a positive impact on cognitive and school outcomes'. They also suggested that this was because 'smaller groups and lower staff ratios provide more opportunity for teachers to match content to children's particular developmental levels so that they are able to learn various academic concepts' (cited in Camilli et al., 2010, p. 604). We will discuss this topic more in detail in the following section.

Some studies have examined the relationship between child to staff ratios and academic achievement in various domains. Broberg et al. (cited in Mitchell et al., 2008, p. 31) found that in Swedish schools, better ratios and smaller group size were associated with enhanced mathematical abilities at age eight. However, not all research on math outcomes has indicated a positive impact. Friendly, Ferns and Prabhu (2009) analysed 12 studies on this topic and found that only two of them reported significant effects, while the rest were nonsignificant. Perlman et al.'s (2017) literature review found that across 12 studies using five different measures of math outcomes, most results were nonsignificant, though one was significant and positive, and one was significant and negative. Both studies also discussed language outcomes, which were mostly nonsignificant, along with other mixed results. Perlman et al. (2017), whose study focused on early childhood education settings meeting local regulations for child to staff ratios, however, acknowledged that 'our findings should not be interpreted as indicating that regulation of ratios can be relaxed in any way. Rather, we emphasize that within the range of what is currently permissible by licensing regulations, better ratios in preschool-aged ECEC classrooms are not associated with better outcomes for children' (p. 63). On the other hand, the NICHD Early Child Care Research Network (cited in Mitchell et al., 2008, p. 37) studied language outcomes and 'found a linear association between the number of

¹ Although no optimal number of attendance hours could be estimated.

recommended standards for quality (teacher training, teacher education, group size, and teacher: child ratios) met and language comprehension scores at 36 months'. While the study found no clear threshold of minimum standards to be met, it clearly showed that not meeting any of the quality standards had a negative impact on average scores at 36 months for language comprehension, while meeting all of them meant above average scores. It is interesting to note that children's outcomes were best predicted by ratios for children aged two and by caregiver training and education for children aged three. This is an important reminder that children's age needs careful consideration when establishing regulations and determining the quality elements that matter. To facilitate this, studies should be very careful in separating the programmes' impacts on children of different ages, particularly with a focus on distinguishing zero- to two-year olds, three- and four-year-old and five- to eight-year olds.

Finally, Burchinal et al.'s (cited in Mitchell et al., 2008, p. 37) looked at a group of 89 African American children attending community-based provision since before their first birthday and found that children in classrooms with the suggested ratios had higher receptive communication scores at 12, 18, 24, and 36 months and higher expressive communication skills at 36 months. The authors suggested this finding is consistent with the fact that scaffolded conversations are especially important for language development during early childhood.

Some of the literature has also explored the association between child to staff ratios and non-cognitive outcomes, though results have been mixed. Friendly, et al. (2009) found that among the studies they analysed, most found no significant impacts on child behaviour. Two other studies found no significant impact of ratios on socio-emotional outcomes but in both cases, results had to be interpreted with caution. In the first study, the Effective Provision of Pre-School Education (EPPE) Project reviewed, among others, by Mitchell et al. (2008, p. 50), one element that might account for this finding is that centres with more favourable ratios also tended to have less-qualified staff. In other words, the effects of ratios were confounded with qualifications. On the other hand, Slot et al. (2015) presented results that are very contextual to the Dutch system, which is highly regulated from a structural point of view and within parameters that are quite different from English regulations. According to the authors and to other research conducted in Dutch context, in the case of the Netherlands 'the absence of effects of group size and child-to-teacher ratio [...] is likely due to the limited variation in these structural characteristics' (p. 73). To conclude, more research is needed on the association between child to staff ratios and non-cognitive outcomes, and specifically in English context, as the limitations of the studies discussed above may be obscuring any real impacts.

Impact on process quality

Many studies have analysed the relationship between ratios and teacher-child interaction, teacher behaviour, sensitivity and stimulation. Gerber, Whitebook and Weinstein (2007) found that accreditation status—which prescribes certain child to staff ratios—was associated with teacher sensitivity. However, accreditation was also associated with several other teacher and setting characteristics, which could have contributed to this association. Howes et al. (cited in Smith et al., 2000) observed the effect of introducing stricter standards of training and ratios on staff-child interactions and global measures of quality (as measured by ECERS). The higher structural standards led to more positive interactions between children and adults and improvements on the ECERS scores of centres. Like several other texts we found through our search, this study did not isolate the effect of ratios, but supported the role of ratios as an important component of quality.

Research examining ratios separately from other aspects of structural quality has found that more favourable ratios are positively associated with process quality outcomes, though the extent of this effect is unclear. Russell (cited in Smith et al., 2000, p. 55) looked at the impact of changing ratios in 27 South Australian kindergartens and found significant relationships between ratios and both staff and child behaviour, indicating that more favourable ratios led to more friendly staff and better-behaved children. For example, they found that 'When ratios were more unfavourable staff spent more time in social interaction with other children and tended to be further away from staff members. Children spent more time annoying and teasing others and staff were less likely to interact and to be close to children with poorer ratios. More positive staff behaviours were less likely to occur when the group was larger and the ratio poorer' (cited in Smith et al., 2000, p. 55).

Additionally, Marshall et al. (2001) found that when child to staff ratios were lower, children had more age-appropriate stimulation. The study concluded that 'Preschool classrooms with fewer children per staff member and better educated teachers are also the classrooms that meet a range of quality standards, as evidenced in the ECERS-R and the Process Quality Index. The ECERS-R and the Process Quality Index include not only stimulation, warmth and sensitivity, and engagement, but also include policies that affect staff and parents, measures of space and furnishings, as well as other characteristics of quality early care and education' (Marshall et al., 2001, p. 30). Meanwhile, Smith, McMillan, Kennedy and Ratcliff (cited in Smith et al., 2000) looked at the impact of introducing a third teacher in a kindergarten classroom. The only effect on children's behaviour was 'to reduce the amount of negative "agonistic" behaviour between peers' (p. 55). However, several issues with how the change in ratios was introduced might have limited the impact. In a follow up study, the same authors looked into teachers' perspectives on the changes in ratios. Teachers reported that the effects of changing ratios depended on the dynamics between teachers. If the teachers' roles were not clear or power was distributed unfairly, adding another teacher was unproductive (Smith, McMillan, Kennedy & Ratcliff, 1992 cited in Smith et al., 2000).

Two studies examined the impact of child to staff ratios on classroom and childcare quality, with mixed results. Pianta et al. (2005) found no relation between child to staff ratio and quality. However, the researchers acknowledged that 'state-level factors not attributable to the teacher, program, and classroom factors examined accounted for the majority of explained variance in observed quality' (p. 144). Additionally, this study is only partially relevant to England, where the early years system is not separated between comparable federal- (Head Start) and state- (pre-kindergarten programmes) funded programmes. In another study conducted in Portuguese context and focused on provision for children below age two, child to staff ratio was found to be a statistically significant predictor of overall child care quality (Barros & Aguiar, 2010).

Lastly, a few studies reported long-term social outcomes, labour market participation and delinquency. Mitchell et al. (2008) cited a study by Yoshikawa, which reviewed 40 programmes for which there were comparison groups looking at the long-term effects of early childhood programmes on social outcomes and delinquency. There were some mixed results, but high quality ECE programmes with family support were generally associated with positive outcomes for parent and teacher ratings of behaviour, official delinquency, and criminal reports. Hayakawa and Reynolds (2014) found that ratios of at most nine children per staff member were most associated with long-term effects and high economic returns. Finally, Van Huizen and Plantenga (2015), who reported on a variety of long-term outcomes, determined that while child to staff ratios and high educational

qualifications were not always significantly associated with positive outcomes, the two in combination were. This finding supports the framework of the iron triangle—a combination of child to staff ratio, group size, and staff qualifications - being instrumental to structural quality.

Conclusions

The evidence on child to staff ratios is quite conclusive: favourable child to staff ratios are important to high quality provision. Having fewer children per staff leads to better children's outcomes as it provides the opportunity for more individualised attention and it leads to better teacher and child behaviour.

There is not a golden rule to determine the exact ratio. Nevertheless, both research and practice lead to strong suggestions that are consistent with those provided by accreditation agencies, such as the National Association for the Education of Young Children (NAEYC). NAEYC's requirements are: ratios of 3:1 or 4:1 for infants, a range from 3:1 to 6:1 for 12- to 36-month-old children, from 6:1 to 10:1 for three- and four-year-olds, from 8:1 to 10:1 for four- to five-year-olds and from 10:1 to 12:1 for kindergarteners (NAEYC, 2013). Accordingly, Cathy Wylie comments that 'A critical cut-off point, above which children are harmed by their experience or gain nothing from it is not clear but looks to be around 6-7:1 for infants and 11 or 12:1 for older children' (cited in Smith et al., 2000, p. 55). Similarly, Minervino's (2014) review of the most successful, effective and scalable pre-school programmes (i.e. for three- and four-year-old children) of the last few years identifies ratios of 10:1 or better a necessary feature, with higher ratios associated with poorer children's outcomes (Minervino, 2014).

The literature also points to two key elements that need to be kept in mind in the context of this review and of policy recommendations. First, both the appropriate ratio and the importance of it differ by child age, as we have seen in the previous paragraph. Second, in extrapolating learning lessons from international evidence, we need to remember that studies looking into the early years abroad may be dealing with ages that would equal primary school in England. This is clearly the case for Kindergarten children in the USA, Australia and New Zealand (to name a few countries), who are the same age as English Y1 children.

Finally, in many cases tighter ratios might be a necessary condition to improve the quality of provision but not a sufficient one. The literature we have analysed shows the clear connection between ratios and staff qualifications and group/classroom size. Tightening ratios while relaxing regulations on the other two elements of 'the iron triangle' might lead to disappointing outcomes. Some of the literature presented in this chapter illustrated the importance of not relaxing regulations on staff qualifications and training in order to make the most out of tighter ratios. In the next chapter, we will observe more in detail the connection between ratios and size.

Classroom and/or group size

The last structural element we investigated is size. The literature is sometimes confusing as ‘size’ can refer to: setting/school, classroom or group. To keep our work focussed, we have left out of our analysis the studies on settings size. School size is important - and the literature on this topic provides some interesting insights – but we wanted to keep our analysis within the realm of ‘the iron triangle’.²

Strictly speaking, group size is the third element of the ‘iron triangle’. It is widely recognised that small group size provides a better learning environment through improved opportunities for quiet and focused activities and for more positive and individualised teacher-student interactions. On the other hand, large groups lead to ‘lower global quality, more adult restriction and control, less positive caregiving, less social interaction, more crying, less talking and playing, more separation distress, lower language and complexity of play scores, lowered social competence, decreased creativity and co-operation, and fewer episodes of joint attention’ (Smith et al., 2000, pp. 58-59).

The effectiveness of working in small groups also seems to be generally appreciated at different levels of the school system – though probably at varying degrees depending on contexts and/or subject. It is also common practice in the early years to split the classroom into small groups throughout the day.

However, group size and classroom size are sometimes used interchangeably, and this can have important consequences. From a methodological point of view, in some of the studies we reviewed the authors did not set out by explicitly distinguishing one from the other. Therefore, we included studies that looked at either classroom or group size, particularly when it was clear that they could offer good quality evidence (through randomised controlled trials or semi-experiments).

In addition, looking at group size alone can lead to an inaccurate interpretation of ‘what works’, as group size and classroom size are often interrelated with each other and with ratios. Small group activities are sometimes used as remedy to classrooms that are too big in size. Moreover, the benefits might get confounded with other elements of provision, such as staff’s level of education. On the other hand, creating stricter child to staff ratios can lead to children being organised into larger groups across activities. From a regulatory point of view, international evidence supports the idea that regulations specifying child to staff ratios should also address the issue of classroom size (Munton et al., 2002, pp. 9-10).

We decided to include classroom size because, in England, there is no statutory limit to class size in the early years, while the maximum established for Reception Year is much higher than what is the norm for the corresponding age group in other countries. In fact, Iacovou (2002) comments that ‘under successive Conservative governments between 1979 and 1997, class sizes in state schools increased steadily, growing to become among the largest in the OECD’ (p. 3). The trend continued in subsequent years. Primary schools class size in the UK was already 25.8 on average in 2006 and

² Humann, C., & Griffin, S. (2014). *Preliminary report on the impact of school size*. Denver, CO: Augenblick, Palaich & Associates.

Ho, D., Lee, M., & Teng, Y. (2016). Size matters: The link between staff size and perceived organizational support in early childhood education. *International Journal of Educational Management*, 30(6), 1104-1122.

increased to 26 in 2015.³ Only three countries – Chile, Israel and Japan – had bigger average class sizes. Most of the other countries for which data are available have an average class size of around, or below, 20 students. In addition, the most common trend over the last decade was for a decreasing in class size, contrary to the trend for UK (OECD, 2018).

Finally, it is interesting to note that some authors are adamant that ‘Some topics, such as class-size reduction, have been the subject of rigorous and well-controlled experiments that have undergone the intense scrutiny of peer review and stood up to the test’ (Barnett, 2002). This, however, is more the case for studies done at primary school level. And while most of this literature seems conclusive about the impact of class size reduction on the learning environment, students’ outcomes and other success measures (Barnett, Schulman & Shore, 2004), the topic is still either contentious (for primary and secondary school level) or neglected (for the early years) in England. It is from these considerations that evidence gaps and promising areas for further study stem.

Message from research

Impact on children’s outcomes

Part of the literature on group/classroom size points to the fact that the approach to the topic can depend significantly on cultural and contextual factors. Boocock and Larne (cited in Smith, 2000, p. 59), for example, remarked that it is a variable that is stressed more in the United States than elsewhere and that views of quality in other countries - such as Japan, France and Italy - place less importance on group size. Also, context matters, above all if we do not isolate the impact of group/classroom size from the other elements of structural quality. In some cases, the positive effect of smaller classrooms might be lost by looser regulations of teachers’ qualifications, and vice versa. In this sense, cross-country comparisons are not straightforward. A study by Cryer et al. (cited in Munton, et al., 2002, p. 92) looked at the relationship between structural and process quality in four countries - USA, Portugal, Spain and Germany - and found no significant differences in process quality even though structural characteristics differed significantly. However, it seemed clear that structural elements can be manipulated into different combinations to reach similar levels of process quality. For example, in Germany, staff and centre directors have generally higher levels of education compared with the United States but they also have less favourable child to staff ratios.

Other international evidence might not be fully relevant to the English context because of significant differences in the way the child care system is regulated. For example, Slot et al. (2015) found that class size does not matter in Dutch child care provision, but the authors warn that this result was likely driven by the limited variation in the structural characteristics. In addition, The Dutch Childcare Act of 2005 prescribes a child-teacher ratio of 7:1 for two- and three-year-old children and a maximum group size of 12 for two- to three-year-old children and 16 for three- to four-year-old children. Also, teachers are required to have completed a minimum of three years vocational training in a relevant subject. Therefore, the conclusions reached by this study cannot necessarily be translated to the English context as Dutch regulations of structural elements are much stricter.

Nevertheless, when comparisons are appropriate and/or results are generalisable, the consensus is that a smaller group and classroom size has a positive impact on children’s outcomes, all else being equal. Barnett’s (2002) literature review concluded that all the sources were consistent in showing

³ Data from the Department for Education, which are specific to England, suggest similar figures.

the positive impact of smaller class sizes and better ratios as they provide opportunities for better teaching, more individual attention, and larger cognitive gains, especially for disadvantaged students.

Most of the international evidence we found focussed on class size for children in their first few years of primary school. For example, Finn (2002) investigated the impact of class-size reduction in grades K-3 (the equivalent of Y1-Y4 in England), drawing similar conclusions. This and other studies we reviewed make extensive reference to the STAR project, one of the most famous studies on the subject and one that, similarly, looks at the transition between early years and primary school. One particularly significant conclusion reached by STAR is that the timing and continuation of class-size reduction matter. In fact, Finn (2002) stated that 'The most recent analyses of STAR data show that the greatest initial impact on student achievement is obtained when students enter reduced-size classes in kindergarten or Grade 1. Pupils who attended small classes for at least three years had significant sustained benefits through Grade 8; the carry-over effects of fewer than three years were mixed. [In addition,] The STAR results do tell us about one alternative reduced-ratio arrangement: a full-size class with a full-time teacher aide does not work. Alternative class configurations, such as team-taught classes or classes with support teachers for reading and math instruction, need their own research to evaluate whether or not they offer viable options to increase student achievement' (p. 51). These conclusions make up a strong evidence base in favour of small classroom sizes, being a controlled scientific experiment.

Iacovou (2002) used data from the National Child Development Study (NCDS) and expanded on previous studies by isolating the impact of class size in infant school on children's outcomes and using an exogenous instrument for class size that had never been used before: the interaction between school size and school type. The use of this instrumental variable revealed a significant and sizeable positive association between small classes and higher reading scores, an effect that was true for all students but that also showed persistency through age 11 for some groups, such as girls and students from larger families. On the other hand, the authors found no impact on mathematics scores.

Finally, Yan and Lin (2005) explored the effects of length of school day and class size on kindergartners' reading, math and general knowledge achievement at the end of the kindergarten year using data from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), with a sample of 15,575 children. A slight positive relationship was found between small class size and children's achievement in reading and maths, particularly for disadvantaged children. On the other hand, no relationship linked class size and general knowledge achievement.

As mentioned above, several of these studies investigated the topic of class size for children beyond the early years. However, Barnett, et al. (2004) suggested that 'it seems reasonable to extrapolate from these findings to 3- and 4- year-olds [as] theory and evidence indicate that preschool children should benefit from small class size even more than do kindergarten children' (p. 6). Barnett and Ackerman (2006) reported of a considerable amount of research, including randomised trials at the preschool and kindergarten level, which found that smaller classes and better ratios are associated with better teaching and improved outcomes for children as they give teachers the opportunity to engage in more stimulating, responsive, and supportive interactions, and provide more individualised attention and dialogues. Teachers also spend less time managing behaviour and more time in educational activities. Nevertheless, rigorous evaluation of the impact of group and/or

classroom size on children's outcomes specifically for the early years and Reception Year is much less available, above all in the English context.

Impact on process quality

While the above studies focused specifically on the impact of classroom size on children's outcomes, several important analyses looked into the effects on teacher interactions, and classroom organisation and environment. Munton et al. (2002) reviewed a number of studies that found a positive impact of group size on a variety of outcomes. For example, Clawson (1997) considered the effect of group size and teacher education on adult:child interactions using a sample of 12 classrooms serving 194 pre-school children. Group size varied from five to 46 children. A measure of teacher education was based on an average of all teachers in a room and included years of experience in day care, specialised training in child development and care, and level of educational attainment. Results clearly showed that: in smaller classes with more favourable teacher-child ratios, there were more frequent adult:child interactions; social interaction was more likely to occur in classes with fewer children; and highly qualified staff were more likely to have more positive interactions with children (cited in Munton et al., 2002, p. 87).

Renwick and McCauley (cited in Munton et al., 2002, p. 92) examined teachers' perceptions of increases in the size of groups in 54 kindergartens in New Zealand. Teachers in larger groups (45 children compared to 30) concluded that the bigger size had negative consequences for both teachers and children. They believed that children were overwhelmed in larger groups, and had to compete more for equipment, space and teacher time. Teachers said they found it difficult to do individual or small-group work with children, were less able to provide varied learning experiences, and had to move constantly between groups of children. This had a negative impact on the quality and continuity of their interactions with children. Teachers working in classes with 45 children listed only negative consequences for children, staff and the quality of programs, from having to work with such large groups. Among the problems identified as a consequence of working in large classrooms were: excessive noise; more difficulties for quiet children; concern about accidents and frustration for children in trying to gain teacher attention; teachers' role becoming mainly supervisory; increased workload; difficulty finding time connecting with all the families; and low morale.

These conclusions are supported by data from the National Day Care Study (cited in Munton et al., 2002, p. 93), which also showed that 'in smaller groups, adults spent more time interacting with children and less time simply watching them. Consequently, children cared for in smaller groups were more verbal, more engaged in activities, less aggressive and performed better on tests of language and learning'. And again, a review of the research on quality provision conducted by Gillian Doherty in 1991 for the Child Care Branch of the Ontario Ministry of Community and Social Services - and updated in 1996 - concluded that 'despite a couple of studies that contradicted the general trend in research findings, it was safe to conclude that smaller groups facilitated caregiver behaviour which in turn encouraged positive child development. [...] Similarly, children in smaller groups tended to cry less frequently, do better on measures of social competence and exhibit more highly developed styles of play' (cited in Munton et al., 2002, p. 95).

Finally, we found only two studies that specifically looked into provision for children under two. Barros and Aguiar (2010), investigating the quality of Portuguese programmes for toddlers, found small statistically significant positive associations between overall quality and group size. Smith's

study (cited in Smith et al., 2000) looked at 200 under two-year-old children and staff in New Zealand child care centres and showed that ‘smaller group size was the best predictor of positive child initiations and total child initiations.’⁴ Class size was also strongly related to the number of joint attention episodes. Centres with group sizes of 14 or less had three times the number of joint attention episodes between children and staff than occurred in groups of 26 or more’ (pp. 58-59). Joint attention is an essential component of quality. It allows the teacher to work in the child’s zone of proximal development, provide opportunities for extension of language, mediate social interactions between peers, encourage the exploration and problem-solving and support the extension of children’s physical skills. Despite the limited literature regarding infants and toddlers, both research and practice indicates that smaller group and class size is extremely important for this age group.

Conclusions

International evidence strongly indicates that smaller class sizes for the entire school day are associated with improved children’s outcomes, greater educational effectiveness and other benefits at classroom level. Even within studies that focus only on preschool children, the effects of class size have been found to be larger for younger children (Barnett et al., 2004; Smith et al., 2000).

Just as in the case of child to staff ratios, there is not a golden rule for size. And yet, both practice and research point to some general guidelines. Barnett (2002) and Barnett et al. (2004) go as far as suggesting that the best practice is a class of 15 (or fewer) with a teacher and an aid. More recent evidence points to a maximum class of 20 children. The OECD data mentioned above show that the clear majority of the countries have an average class size that is around or below 20 students for primary grades. As Barnett et al. (2004) comment, it seems reasonable to extrapolate from these findings to three- and four-year-olds as both theory and practice indicate that preschool children should benefit from small class size even more than kindergarten children do. Several other studies suggest a classroom size of no more than 20 children (Hayakawa & Reynolds, 2014; Minervino, 2014).

Classroom size for children zero to four is not regulated in England but practice seems to be in line with what the international evidence establishes as best practice. What stands in starker contrast is that for children in Reception Year, 30 pupils per class is the norm despite international evidence that clearly points to a maximum average size of 20 children per class for this age group.

Finally, Barnett et al. (2004) have some practical recommendations on how to transition towards smaller classes, such as phasing in smaller classrooms gradually and providing adequate financial support. More importantly, and again showing the connection with other elements of ‘the iron triangle’, professional support and development for teachers of small classes is also needed to get the most out of small class size.

⁴ ‘Child initiations’ are play and learning activities that are chosen and initiated by the child or for which the child take ownership and ‘subvert’ them to a different purpose than the one intended.

Discussion and recommendations

Evidence gaps

As most of the research analysed with this literature review took place outside of England and/or in very specific settings conditions, care must be taken to account for other structural factors and context before generalising findings to England early years provision. For example, while the American childcare system presents many similarities to the English one - both being a mixed-market system with a relatively low qualified and low paid workforce - it also presents significant differences in terms of regulations, stakeholders, uptake and funding. In addition, some of the studies in this, and other literature reviews already feel out-dated. Below we will provide a list and short discussion of the main evidence gaps that emerged from our literature review.

Workforce training and professional development

The following are more specific areas that emerged as lacking in clear, robust and/or recent evidence:

- Quality level of the training providers: just as in other countries, in England there are many different routes to enter the early years sector and a considerable number of training providers. Yet 'little is known (except on paper) of how these training programs balance practice and theory, the content of their curriculum, or the duration, intensity and nature of the learning environments provided. The outcomes of different kinds of teacher education program in terms of teacher competency, where they are employed, how stable their employment history is needs investigation' (Smith et al., 2000, p. 78);
- Differential impact of education and training for teachers working with different age groups: particularly understudied is the workforce of children aged zero to two;
- Differential impact of CPD for teachers with different levels of qualifications and with different educational background: we need more opportunities to conduct tests across different types of provider (PVI versus maintained) so that we have enough variation in teachers' qualifications;
- Differential impact of CPD for teachers at different points in their career: we need more opportunities to test which CPD opportunities are more effective for new teachers (less than 5 years of experience) versus more experienced ones;
- Differential impact of education and training for teachers and managers: research points to the fact that teachers' and leaders' education level and focus, while being equally important, might exert their effects through different routes. For example, Munton et al. (2002) points to education and training of centre managers having a greater influence on global quality, while teachers' education and training having an impact on classroom quality. We need to know more about these topics in the English context and how it might differ between the maintained sector and PVIs;
- Definition and impact of specialised training in early childhood education: we need to know more about what exactly it encompasses and the mechanisms through which it advances teachers behaviour. A connected issue raised by Brown, Burr, Johnson, Krieger and Mihaly is that 'for many in the current early care and education workforce, specialized early childhood training and higher levels of formal education have gone hand in hand, making it challenging

to identify the differing contributions that formal education and less formal training make to teacher behaviour. There continues to be not only confusion in the literature, but also among practitioners and policy makers, about the particular role of training in early childhood education or child development, with large investments into a wide array of training programs targeted at a diverse group of teachers and providers' (as cited in Whitebook, 2003, p. 16). These are all important areas that need to be differentiated;

- The effectiveness of supervised teaching or mentoring as training strategies;
- Costs and benefits of alternative staffing structures deriving from different combinations of education and experience levels: for example, one teacher with a degree and assistants without, one very experienced teacher with less experienced assistants;
- Value added of a degree or an advanced degree: we need to better understand what additional value having a degree or an advanced degree provides to a teacher's competences, skills and knowledge as this could help us understand if there is a threshold of education and training that afford a certain degree of quality of provision (Whitebook, 2003);
- The aspects of the adult work environment that scaffold teachers' knowledge, enabling them to engage in effective strategies with children.

Child to staff ratios

The main evidence gaps fall into the following strands:

- Impact of ratios: very few studies have looked at the impact of more stringent ratios in isolation. In the case of children aged zero to four, statutory ratios in England are in line with both evidence and good practice. On the other hand, reception classes are the most divergent part of our education system and the impact of tighter ratios for Reception Year classes should be considered as an important area of further study;
- Impact of ratios on other practices at setting or classroom level (e.g. establishment of informal class size, staffing organisations): few studies have been conducted in England on this topic. This is important for two main reasons: first, in England - contrary to the majority of other countries and particularly the USA - there is no statutory class size for children zero-to four-year-old; second, Reception Year classes generally have a much higher child to staff ratios than what international evidence prescribes for the same age range.

Group/classroom size

While practitioners and parents' naturally feel that smaller class sizes were better, England-based literature has not been conclusive and presents some important gaps, such as:

- Impact of smaller group size in English context: the majority of the studies we analysed are not based in England; therefore, we need more studies specific to the English context and to both early years and Reception Year;
- Impact of smaller classroom sizes in isolation from other structural elements: More randomised controlled trials or quasi-experiments are needed that look at only classroom size effects.

Policy recommendations/Future research needed

Based on the evidence gaps identified above, several areas present the potential areas for further investigation and trials. Ideally, tests should be run across different types of providers to have enough variation, for example, in teachers' qualifications, as well as for different age ranges:

- Testing the impact of CPD on teachers with different levels of qualifications;
- Testing the effectiveness of different staffing structures in terms of levels of qualifications and years of experience;
- Testing different types of training routes for teachers working with children of different age ranges;
- Testing the effectiveness of supervised teaching or mentoring as training strategies;
- Testing the impact of smaller ratios and class size in nursery classes without changing teachers' qualification requirements;
- Testing the impact of smaller ratios in Reception Year classes;
- Testing staffing structures in Reception Year that allow for big classes to be organised into small groups so as to achieve smaller ratios at least during certain activities (e.g. more academic in focus);
- Testing different classroom sizes in newly opened nurseries (could also trial different thresholds).

In particular, we think that the key priorities should be:

- Testing whether CPD can be a substitute for pre-service training and, if so, to what extent;
- Testing a variety of nursery staff structures;
- Testing smaller classroom sizes in Reception Year (could also trial different thresholds).

Conclusions

The last twenty years have seen a flourishing of early years research focused on the topic of quality, recognising that high quality provision is necessary to improve children's outcomes, particularly for disadvantaged children. This evidence base is both dense and scattered. At times, it is also far from conclusive. It is often difficult to compare studies and to assess which of the key drivers of quality provide the greatest contribution to children's outcomes. This is due to multiple reasons, such as: the heterogeneity of interventions; their different purposes (targeted versus universal); the time period in which they were delivered; the measures used to calculate effect size; their different definitions and measures of children's outcomes; and, sometimes, fairly old data. Only more recently have studies moved to rigorous methodologies, such as randomised controlled trials and quasi-experiments. With this literature review we have attempted to systematise this evidence to gain a deeper understanding of what we know and where the evidence gaps are.

Both process and structural quality are needed to create the best environment for young children (Hayakawa & Reynolds, 2014; Marshall, et al., 2001; NICHD Early Child Care Research Network, 2002). We focused on structural elements, and specifically on 'the iron triangle', because they are the more easily observed, measured and regulated elements. And yet theory and practice are not always in agreement. The studies we analysed show how key these elements are, but also how hard it is to disentangle the impact of each individual factor. In order to fill in some of the gaps found through this review, we need more robust studies in the form of randomised controlled trials and/or quasi-experiments. At the same time, we need to remember to contextualise the results and remember that in practice these elements are interconnected and should not be seen in isolation.

Finally, it is also important to recognise the role played by other elements at the macro level, such as financing and costs, policies and procedures, entitlement to early years services, participation rates, and types of early years setting that exist in a country. Making significant changes to structural elements is costly, and the costs depend on the other macro elements listed above (Early et al., 2007; Slot et al., 2015). Policies can become controversial because of the price tag attached to them. One such example is the case of policies that mandate certain levels of education attainment for staff (Early et al., 2007; Whitebook, 2003). Changing ratios and/or classroom size can lead to difficulties in recruiting qualified teachers in newly created classrooms, which could lead to a level of disorganisation and other dynamics at setting level that could offset the benefits that small classes provide if not timed correctly (Finn, 2002). As recruitment and retention issues are already widespread in the early years sector, these are factors that cannot be disregarded.

In conclusion, and given the cost impact of structural changes, this literature review has identified some evidence gaps that need to be filled in order to understand how to direct public investments more effectively and sustainably. Yet, there are some strong pointers coming from both evidence and practice. While there is no unique recipe to create a high quality early years system, we also know that a systematic approach and political willingness to increase investments in the sector are needed. With this work we have attempted to make these pointers more visible.

References

- Barnett, W. S. (2002). Early childhood education. In A. Molnar (Ed.), *School reform proposals: The research evidence* (pp. 8-35). CT: Information Age.
- Barnett, W. S., Schulman, K., & Shore, R. (2004). *Class size: What's the best fit?*. New Brunswick, NJ: National Institute for Early Education Research.
- Barros, S., & Aguiar, C. (2010). Assessing the quality of Portuguese child care programs for toddlers. *Early Childhood Research Quarterly*, 25(4), 527-535.
- Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., . . . Gill, S. (2008). Promoting academic and social-emotional school readiness: The Head Start REDI program. *Child Development*, 79(6), 1802-1817.
- Brühwiler, C., & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction*, 21(1), 95-108.
- Buysse, V., Castro, D. C., & Peisner-Feinberg, E. (2010). Effects of a professional development program on classroom practices and outcomes for Latino dual language learners. *Early Childhood Research Quarterly*, 25(2), 194-206.
- 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.
- Chambers, B., Cheung, A., Slavin, R. E., Smith, D., & Laurenzano, M. (2010). *Effective early childhood education programmes: A systematic review*. Peru: Ministerio de Educación.
- Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, 1, 697-812.
- Domitrovich, C. E., Gest, S. D., Gill, S., Bierman, K. L., Welsh, J. A., & Jones, D. (2009). Fostering high-quality teaching with an enriched curriculum and professional development support: The Head Start REDI program. *American Educational Research Journal*, 46(2), 567-597.
- Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., . . . Henry, G. T. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development*, 78(2), 558-580.
- Felfe, C., & Lalive, R. (2014). *Does early child care help or hinder children's development?* IZA Discussion Paper, No. 8484. Bonn: Institute for the Study of Labor.
- Fenech, M. (2011). An analysis of the conceptualisation of 'quality' in early childhood education and care empirical research: Promoting 'blind spots' as foci for future research. *Contemporary Issues in Early Childhood*, 12(2), 102-117.
- Finn, J. D. (2002). Class-size reductions in grades k-3. In A. Molnar (Ed.), *School reform proposals: The research evidence* (pp. 37-33). CT: Information Age.

- Friendly, M., Ferns, C., & Prabhu, N. (2009). *Ratios for four and five year olds: What does the research say? What else is important?* Univ Tor Childcare Resour Res Unit. Retrieved from http://childcarecanada.org/sites/default/files/BN_ratios.pdf
- Greenfield, T. (2002). *Research methods for postgraduates*. London: Arnold.
- Hayakawa, C. M., & Reynolds, A. J. (2014). Key elements and strategies of effective early childhood education programs: Lessons from the field. In Ben-Arieh, A., Casas, F., Frønes, I., Korbin, J.E. (Eds.) *Handbook of Child Well-Being*, XXXV, 2993-3023.
- Howes, C., Phillips, D., & Whitebook, M. (1992). Thresholds of quality: Implications for the social development of children in center-based child care. *Child Development*, 63, 449-60.
- Hutchinson, J., Dunford, J., & Treadaway, M. (2016). *Divergent pathways: The disadvantage gap, accountability and the pupil premium*. London: Education Policy Institute
- Hutchinson, J., Robinson, D., Carr, D., Hunt, E., Crenna-Jennings, W., & Akhal, A. (2018). *Education in England: Annual report 2018*. London: Education Policy Institute.
- Iacovou, M. (2002). Class size in the early years: Is smaller really better?. *Education Economics*, 10(3), 261-290.
- Layzer, J. I., & Goodson, B. D. (2006). The “quality” of early care and education settings: Definitional and measurement issues. *Evaluation Review*, 30(5), 556-576.
- Marshall, N. L., Creps, C. L., Burstein, N. R., Glantz, F. B., Robeson, W. W., & Barnett, S. (2001). *The cost and quality of full day, year-round early care and education in Massachusetts: Preschool classrooms*. Wellesley, MA: Wellesley Centers for Women and Abt Associates.
- Milburn, T. F., Hipfner-Boucher, K., Weitzman, E., Greenberg, J., Pelletier, J., & Girolametto, L. (2015). Effects of coaching on educators' and preschoolers' use of references to print and phonological awareness during a small-group craft/writing activity. *Language, Speech, and Hearing Services in Schools*, 46(2), 94-111.
- Mims, S. U., Scott-Little, C., Lower, J. K., Cassidy, D. J., & Hestenes, L. L. (2008). Education level and stability as it relates to early childhood classroom quality: A survey of early childhood program directors and teachers. *Journal of Research in Childhood Education*, 23(2), 227-237.
- Minervino, J. (2014). *Lessons from research and the classroom*. Seattle, WA: Bill & Melinda Gates Foundation.
- Mischo, C., Wahl, S., Strohmer, J., & Wolf, C. (2013). Does early childhood teacher education affect students' cognitive orientations?. *Journal of Education and Training Studies*, 2(1), 193-206.
- Mitchell, L., Wylie, C., & Carr, M. (2008). *Outcomes of early childhood education: Literature review*. Wellington: New Zealand Council for Educational Research.
- Mitter, R. & Putcha, V. (2018). *Strengthening and supporting the early childhood workforce. Training and professional development*. Washington, D.C.: Results for Development

Munton, T., Mooney, A., Moss, P., Petrie, P., Clark, A., Woolner, J., . . . & Barreau, S. (2002). *Research on ratios, group size and staff qualifications and training in early years and childcare settings*. DfES Research Report: RR320. London: Department of Education and Skills.

National Association for the Education of Young Children. (2013). *NAEYC accreditation*. Washington, DC: NAEYC. Retrieved from <https://idahostars.org/portals/61/Docs/Providers/STQ/TeacherChildRatioChart.pdf>

Neuman, S. B., & Kaefer, T. (2013). Enhancing the intensity of vocabulary instruction for preschoolers at risk: The effects of group size on word knowledge and conceptual development. *The Elementary School Journal*, 113(4), 589-608.

Neuman, S. B., & Wright, T. S. (2010). Promoting language and literacy development for early childhood educators: A mixed-methods study of coursework and coaching. *The Elementary School Journal*, 111(1), 63-86.

NICHD Early Child Care Research Network. (2002). Child-care structure→ process→ outcome: Direct and indirect effects of child-care quality on young children's development. *Psychological Science*, 13(3), 199-206.

OECD. (2015). *Starting strong IV: Monitoring quality in early childhood education and care*. OECD Publishing: Paris.

OECD. (2018). *Engaging young children: Lessons from research about quality in early childhood education and care, starting strong*. OECD Publishing: Paris

OECD. (2018). *OECD.Stat*. Average class size. Retrieved from https://stats.oecd.org/Index.aspx?DataSetCode=EDU_CLASS.

Perlman, M., Fletcher, B., Falenchuk, O., Brunsek, A., McMullen, E., & Shah, P. S. (2017). Child-staff ratios in early childhood education and care settings and child outcomes: A systematic review and meta-analysis. *PloS one*, 12(1), e0170256.

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.

Rhodes, S., & Hennessy, E. (2000). The effects of specialized training on caregivers and children in early-years settings: An evaluation of the foundation course in playgroup practice. *Early Childhood Research Quarterly*, 15(4), 559-576.

Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B., Grabbe, Y., & Barreau, S. (2007). Summary report. Influences on children's attainment and progress in Key Stage 2: Cognitive outcomes in Year 5. Research Report RR828. London: Department for Education and Skills.

Scobie, G., & Scott, E. (2017). *Rapid evidence review: Childcare quality and children's outcomes*. NHS Health Scotland: Edinburgh.

Siraj-Blatchford, I., Taggart, B., Sylva, K., Sammons, P., & Melhuish, E. (2008). Towards the transformation of practice in early childhood education: The effective provision of pre-school education (EPPE) project. *Cambridge Journal of Education*, 38(1), 23-36.

Slot, P. L., Leseman, P. P., 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.

Smith, A. B., Grima, G., Gaffney, M., Powell, K., Masse, L., & Barnett, S. (2000). *Strategic research initiative literature review: Early childhood education. Literature review for the Ministry of Education*. Wellington: Ministry of Education.

Sylva, K., Melhuish, E., Sammons, P. Siraj-Blatchford, I., & Taggart, B. (2011). Pre-school quality and educational outcomes at age 11: Low quality has little benefit. *Journal of Early Childhood Research, 9*(2), pp. 109–124.

Van Huizen, T. M., & Plantenga, J. (2015). Universal child care and children's outcomes: A meta-analysis of evidence from natural experiments. *Discussion Paper Series/Tjalling C. Koopmans Research Institute, 15*(13).

Wasik, B. A., & Hindman, A. H. (2014). Understanding the active ingredients in an effective preschool vocabulary intervention: An exploratory study of teacher and child talk during book reading. *Early Education and Development, 25*(7), 1035-1056.

Waters, J., & Payler, J. (2015). The professional development of early years educators—achieving systematic, sustainable and transformative change. *Professional Development in Education, 41*(2), 161-168.

Werner, C. D., Linting, M., Vermeer, H. J., & Van IJendoorn, M. H. (2016). Do intervention programs in child care promote the quality of caregiver-child interactions? A meta-analysis of randomized controlled trials. *Prevention Science, 17*(2), 259-273.

Whitebook, M. (2003a). *Bachelor's degrees are best: Higher qualifications for pre-kindergarten teachers lead to better learning environments for children*. Washington, D.C.: The Trust for Early Education.

Whitebook, M. (2003b). *Early education quality: Higher teacher qualifications for better living environments. A review of the literature*. Berkeley, CA: Center for the Study of Child Care Employment, Institute of Industrial Relations.

Workman, E., Guernsey, L., & Mead, S. (2018). *Pre-k teachers and bachelor's degrees. Envisioning equitable access to high-quality preparation programs*. Washington D.C.: New America Foundation.

Appendices

A1. Search terms

Search terms were decided on the basis of the most common language used in the literature on structural quality. The second round of searches was run in order to identify studies that used more rigorous methodologies, such as randomised controlled trials and quasi-experiments.

Figure A1.1: First round of search terms

Group 1: stage of education	Group 2: structural quality	Group 3: outcomes and impact
early years OR ECEC OR early childhood education OR ECE OR early care and education OR early education	Child:staff ratio OR Child:teacher ratio OR Staff:child ratio OR Teacher:child ratio OR ----- Group size OR Classroom size OR Class size ----- Staff education OR Teachers education OR Workforce education OR Staff qualifications OR Teachers qualifications OR Workforce qualifications	Outcomes

Figure A1.2: Second round of search terms

Group 1: stage of education	Group 2: structural quality	Group 3: outcomes and impact	Group 4: methodology
early years OR ECE OR early childhood education OR early care and education OR early education	Child:staff ratio OR Child:teacher ratio OR Staff:child ratio OR Teacher:child ratio OR ----- Group size OR Classroom size OR Class size ----- Staff education OR Teachers education OR Workforce education OR Staff qualifications OR Teachers qualifications OR Workforce qualifications	Outcomes	Randomised controlled trials OR Quasi-experiment

A2. Analytic review template

Item number: _____

Type of resource:

- ☐ Book/book chapter
- ☐ Peer-reviewed article
- ☐ Non peer-reviewed article
- ☐ Thesis or dissertation
- ☐ Report
- ☐ Brief
- ☐ Other:

APA Citation:

Institutional Affiliation(s)/location(s) of author(s)/Department(s)/division(s) and college(s) of author(s):

Summary of the chapter/article/book: (add abstract if there is one)

Keyword(s) listed (Please note when keywords were listed on article and/or in search terms):

What type of research is it? (check one or more)

- ☐ Literature Review
- ☐ Empirical study
- ☐ Theoretical/position paper
- ☐ Practical perspective (literature written for teachers, child care providers or general audience)
- ☐ Other (specify):

What methodologies are used? (check one or more)

- ☐ Qualitative
- ☐ Quantitative
- ☐ Experimental Analysis
- ☐ Case study
- ☐ Ethnography
- ☐ Discourse Analysis

☐ Not Applicable/Other

Where is the research taking place? (list country(ies) or region(s))

Definition of structure(s)/structural quality (copy/paste exact quotes)

What structural elements are considered?

- ☐ Ratios
- ☐ Group size
- ☐ Wages/Pay
- ☐ Workload
- ☐ Staff qualifications/competences/knowledge/skills
- ☐ Training/Continuing/Continuous Professional Development
- ☐ Building/Infrastructure/Materials/Service facilities
- ☐ Transitions/Systems coordination/Reception year
- ☐ Other (specify):

What children's outcomes are considered? Which "outcome(s)" is(are) considered?

- ☐ Cognitive
- ☐ Non-Cognitive
- ☐ Both
- ☐ None

How is "outcome" defined? (examples of quotes)

Sign of the impact:

- ☐ Positive
- ☐ Negative
- ☐ Mixed

What other outcomes are considered?

- ☐ None
- ☐ Adult outcomes
- ☐ Family outcomes
- ☐ Society outcomes

- ☐ School outcomes
- ☐ School system outcomes
- ☐ Other (specify):

Sign of the impact:

- ☐ Positive
- ☐ Negative
- ☐ Mixed

What factors controlling structures are considered/mentioned (e.g., systems perspectives, political/policy contexts, historical contexts)?

- ☐ None
- ☐ Government (policy orientation/priorities/approach, e.g. child-care v education divide; market v public provision; types of ECEC settings that exist in a country)
- ☐ Financing and costs
- ☐ Curriculum
- ☐ ITT Systems
- ☐ Policies and procedures (e.g. benefits system, parental family leave policies, entitlement to ECEC – targeted v universal delivery)
- ☐ Other (describe; e.g. participation rates)

How, when, and to what extent were these larger, macrocontextual factors addressed?

- ☐ Not at all
- ☐ In general terms
- ☐ In depth

Additional Comments:

Sign of the impact:

- ☐ Positive
- ☐ Negative
- ☐ Mixed

What does the author cite as future directions?

Quotes For Literature Review in our Report:

References we May Want to Find and Read:

Additional Comments/Reflections on Article:

Contextualisation to England

Is this literature relevant for the English context?

☐ Yes

☐ No

☐ Only in part

How? (examples/quotes)

A3. Coded table and list of studies analysed

	Type of research					Methodology				Structural elements considered								Children's outcomes considered				Sign of impact			Other outcomes considered							Sign of impact				
Item Number	Literature Review	Empirical Study	Theoretical/position paper	Practical perspective	Other: usually meta-analysis	Qualitative	Quantitative	Experimental Analysis	Case Study	Other	Ratios	Group Size	Wages/Pay	Workload	Pre-service Training	In-service Training	Building/Infrastructure/Materials	Transition/Systems coordination	Other	Cognitive	Non-Cognitive	Both	None	Positive	Negative	Mixed	None	Adult Outcomes	Family Outcomes	Society Outcomes	School Outcomes	School Systems Outcomes	Other	Positive	Negative	Mixed
1		1					1				1					1						1		1			1									
2		1					1				1				1	1			1				1									1			1	
3	1					1											1		1			1				1			1	1				1		
5			1			1													1	1					1		1									
6	1					1					1	1			1	1						1				1	1									
7	1					1					1	1			1	1						1		1				1				1				
9					1		1					1										1				1	1									
25		1				1	1				1	1			1	1			1			1				1	1									
32		1					1				1	1	1		1				1				1				1									
33					1		1				1				1							1		1				1						1		
42	1					1									1	1						1														
44		1					1									1						1		1			1									
47		1						1				1								1						1	1									
56		1					1					1								1						1	1									
57	1					1						1										1				1						1	1			

58		1					1			1		1				1				1				1				1			1				
65	1					1				1	1	1	1	1	1			1			1			1		1	1	1	1	1	1			1	
66		1					1			1	1	1		1				1			1			1											
68	1					1								1							1									1		1	1		
69		1						1							1					1				1											
71		1					1							1	1			1				1							1		1	1			
74		1					1							1					1					1							1			1	
118				1			1								1						1									1		1	1		
119		1						1							1								1							1		1			1
120		1				1	1		1						1					1				1			1					1			
121		1						1							1					1					1		1								1
122		1					1								1					1					1					1			1		
123	1					1					1		1		1	1						1		1			1					1			
124		1					1							1									1								1	1			
125	1					1									1					1				1			1								
126		1					1	1							1							1		1			1								
129	1					1					1	1	1		1	1	1	1				1		1											
131	1					1	1				1											1			1	1									
132	1					1					1											1		1				1					1		

1 NICHD Early Child Care Research Network. (2002)

2 Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005)

3 Mitchell, L., Wylie, C., & Carr, M. (2008)

5 Layzer, J. I. & Goodson, B. D. (2006)

6 Munton, T., Mooney, A., Moss, P., Petrie, P., Clark, A., Woolner, J., . . . & Barreau, S. (2002)

- 7 Barnett, W. S. and Finn, J. D. (2002). In A. Molnar (Ed.), *School reform proposals: The research evidence*.
- 9 Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010)
- 25 Slot, P. L., Leseman, P. P., Verhagen, J., & Mulder, H. (2015)
- 32 Barros, S., & Aguiar, C. (2010)
- 33 Van Huizen, T. M., & Plantenga, J. (2015)
- 42 Whitebook, M. (2003b)
- 44 Rhodes, S., & Hennessy, E. (2000)
- 47 Neuman, S. B., & Kaefer, T. (2013)
- 56 Iacovou, M. (2002)
- 57 Barnett, W. S., Schulman, K., & Shore, R. (2004)
- 58 Brühwiler, C., & Blatchford, P. (2011)
- 65 Smith, A. B., Grima, G., Gaffney, M., Powell, K., Masse, L., & Barnett, S. (2000)
- 66 Marshall, N. L., Creps, C. L., Burstein, N. R., Glantz, F. B., Robeson, W. W., & Barnett, S. (2001)
- 68 Whitebook, M. (2003a)
- 69 Buysse, V., Castro, D. C., & Peisner-Feinberg, E. (2010)
- 71 Mims, S. U., Scott-Little, C., Lower, J. K., Cassidy, D. J., & Hestenes, L. L. (2008)
- 74 Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., . . . Henry, G. T. (2007)
- 118 Werner, C. D., Linting, M., Vermeer, H. J., & Van IJzendoorn, M. H. (2016)
- 119 Domitrovich, C. E., Gest, S. D., Gill, S., Bierman, K. L., Welsh, J. A., & Jones, D. (2009)

- 120 Neuman, S. B., & Wright, T. S. (2010)
- 121 Milburn, T. F., Hipfner-Boucher, K., Weitzman, E., Greenberg, J., Pelletier, J., & Girolametto, L. (2015)
- 122 Wasik, B. A., & Hindman, A. H. (2014)
- 123 Scobie, G., & Scott, E. (2017)
- 124 Mischo, C., Wahl, S., Strohmer, J., & Wolf, C. (2013)
- 125 Chambers, B., Cheung, A., Slavin, R. E., Smith, D., & Laurenzano, M. (2010)
- 126 Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., . . . Gill, S. (2008)
- 129 Hayakawa, C. M., & Reynolds, A. J. (2014)
- 131 Perlman, M., Fletcher, B., Falenchuk, O., Brunsek, A., McMullen, E., & Shah, P. S. (2017)
- 132 Friendly, M., Ferns, C., & Prabhu, N. (2009)

A4. Brief description of reviews and meta-analyses

Reference	Review focus	Time period covered	Number of studies included	Brief summary of the study and findings
Barnett, W. S. (2002). Early childhood education. <i>School reform proposals: The research evidence</i> , 1-26.	Effects of early care and education on children disadvantaged by social and economic circumstances		Narrative synthesis: number of studies not stated	The author conducted a narrative synthesis of the potential benefits and adverse effects of early care and education, placing an emphasis on the effects for disadvantaged children. The author found that early care and education can increase the cognitive abilities of disadvantaged children. Although increases in IQ were temporary, there were long-term increases in the abilities measured by standardised tests in reading and maths. Furthermore, early care and education can have long-term positive effects on children's social and emotional development. The author advises that, in order to reach their full potential, pre-kindergarten programs must be intensive and high quality, focusing on both cognitive and social development.
Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. <i>Teachers College Record</i> , 112(3), 579-620.	Effect of early childhood education programs on cognitive and affective gains in pre-schoolers	1960-2000	Meta-analysis: 123	The authors conducted a meta-analysis of experimental and quasi-experimental studies of early childhood interventions. The search identified 123 comparative studies, comprising 412 treatment-control (T/C) or treatment/alternative treatment (T/A) contrasts, which met inclusion criteria. The authors found significant effects, particularly on cognitive outcomes, for children who attended an early childhood education programme. In T/A contrasts, the authors found that teacher-directed instruction and small-group instruction were positively correlated with cognitive and affective gains, whereas the provision of additional services was negatively associated with these gains. However, the largest effect sizes were associated with high research design quality and the authors warned that weak research designs potentially underestimated the impact of some programmes.

Chambers, B., Cheung, A., Slavin, R. E., Smith, D., & Laurenzano, M. (2010). Effective early childhood education programmes: A systematic review.	Effectiveness of early childhood education programmes	1968-2008	Narrative synthesis: 38	The authors conducted a narrative synthesis of the effects of early childhood education programmes on language, literacy, phonological awareness, mathematical and/or cognitive outcomes. The search identified 38 studies, comprising 27 different programmes, which met the inclusion criteria. The authors found that of the 27 programmes reviewed, six had strong evidence of effectiveness and five had moderate evidence of effectiveness. Averaging across all studies, there were small positive short-term effects for all outcomes. A narrative synthesis was conducted of the long-term effects of four programmes. The authors found that early childhood education programmes that focus broadly on cognitive development had better long-term effects on social adjustment outcomes than programmes focused solely on academic skills.
Fenech, M. (2011). An analysis of the conceptualisation of 'quality' in early childhood education and care empirical research: Promoting 'blind spots' as foci for future research. <i>Contemporary Issues in Early Childhood</i> , 12(2), 102-117.	Conceptualisation of 'quality' in early childhood education empirical research	1980-2008	Critical analysis: 338 peer-reviewed journal articles	The author critically analysed the way empirical research in early childhood education and care has conceptualised 'quality'. The author identified 338 peer-reviewed journal articles that discuss quality and analysed them to uncover six themes that span across the literature: quality is an objective reality; quality enhances children's optimal development; quality is the domain of science/psychology; quality can be known from researchers' perspectives; quality can be understood using an ecological framework that is limited to child, family and early years provision variables; and quality in early years is more pertinent to three- and four-year olds than babies and infants.
Hayakawa, C. M., & Reynolds, A. J. (2014). Key elements and strategies of effective early childhood education programs: Lessons from the field. In <i>Handbook of Child Well-</i>	Key elements of effective early childhood education programmes		Narrative synthesis: number of studies not stated	The authors conducted a narrative synthesis of the key elements for delivering a high quality early childhood education program that will increase the return on investment. The authors identified 10 principles of effectiveness: target children at risk, early timing of participation, sufficient duration of participation, comprehensive services, well-trained and well-compensated teachers, diverse and language-focused instruction, high intensity, small class size and child-teacher ratio,

Being (pp. 2993-3023). Springer Netherlands.				continuity of service, and a strong accountability system. In addition to these key principles, the authors found that one pathway through which early childhood education programmes impact children's well-being is parental involvement. There is strong evidence that parental involvement has a positive effect on children's cognitive and socio-emotional well-being. Although the key principles of effectiveness have been identified, the author notes that the biggest challenge will be determining the combination of principles most effective for different populations and contexts.
Perlman, M., Fletcher, B., Falenchuk, O., Brunsek, A., McMullen, E., & Shah, P. S. (2017). Child-Staff Ratios in Early Childhood Education and Care Settings and Child Outcomes: A Systematic Review and Meta-Analysis. PloS one, 12(1), e0170256.	Association between child-staff ratios in early childhood education and care and children's outcomes	1980-2013	Meta-analysis: 3; Narrative synthesis: 29	The authors aimed to examine the association between child-staff ratios in ECEC classrooms and child outcomes. The search identified 29 articles which met the inclusion criteria, comprising 31 samples in which child-staff ratios ranged from five to 14.5 preschool-aged children per adult. A meta-analysis was conducted of a sub-set of three articles that examined the relationship between child-staff ratios and children's receptive language. The results of the meta-analysis were not significant. A narrative synthesis was conducted on all 29 studies. The authors found that there is little, if any, association between child-staff ratios and child outcomes. However, the lack of significant associations may be due to selection biases, specifically the family-level factors that influence child care selection.
Scobie, G., & Scott, E. (2017). Rapid evidence review: Childcare quality and children's outcomes.	Indicators of quality in early learning and childcare programmes	1960s onwards	Narrative synthesis: number of studies not stated	The authors conducted a narrative synthesis of the indicators of quality in early learning and childcare (ELC) programmes which are associated with children's outcomes. Indicators of quality were classified into two types: structural and process. Structural indicators included group size, child-staff ratios, and teachers' qualification, and process indicators included children's day to day experiences and interactions. Of the structural indicators, the authors found that highly qualified teachers and staff, good working conditions that reduced staff turnover, and the delivery of an age-appropriate curriculum increased the quality of ELC programmes. The process indicators that contributed to an increase in

				quality were closely linked to structural indicators. For example, highly qualified staff were more able to deliver the process indicators which provide the care, nurturing and support necessary for children's developmental needs. However, the authors noted that it was difficult to identify which indicators of quality contributed most to children's outcomes, primarily due to the heterogeneity of interventions. The authors also reviewed the effect sizes of ELC programmes and found that targeted interventions showed the largest positive effect on children's outcomes, whereas the effect of universal programmes was weaker, although still positive.
Smith, A. B., Grima, G., Gaffney, M., Powell, K., Masse, L., & Barnett, S. (2000). Strategic research initiative literature review: Early childhood education. Literature review for the Ministry of Education. Wellington: Ministry of Education.	Issues of quality, diversity and participation in early childhood services, and the outcomes associated with early childhood services	1993 onwards	Narrative synthesis: 797	The authors conducted a narrative synthesis of the effects of structural and process quality on child outcomes. The search identified 797 articles which met the inclusion criteria. The authors found that staff-child ratios, staff training and group size are three of the most important structural indicators of quality associated with high quality care and positive child development outcomes. Staff wages and working conditions were found to be just as powerful as staff-child ratio, teacher training and group size at predicting high quality care. An issue closely tied to wages and working conditions is teacher turnover, which is a concern internationally because high turnover is associated with lower quality care and poorer outcomes for children. Although structural indicators of quality were a necessary condition for high quality care, they were not sufficient. The authors found that process quality, particularly the quality of teacher/child relationships, was the most important indicator of quality.
Van Huizen, T. M., & Plantenga, J. (2015). Universal child care and children's outcomes: A meta-analysis of evidence from natural experiments.	Effects of early childhood education and care on children's development and	2005-2015	Meta-analysis: 30	The authors conducted a meta-analysis of quasi-experimental studies that examined the impact of universal early childhood education and care (ECEC) arrangements. The search identified 30 studies, comprising 253 estimates of children's outcomes, which met the inclusion criteria. The authors found that evidence on universal ECEC is mixed, most likely due to variations in quality. About a third of estimates indicated that

Discussion Paper Series/Tjalling C. Koopmans Research Institute, 15(13)	later life outcomes			universal ECEC has a positive impact on children's outcomes, 20 per cent indicated a negative impact and the remaining estimates were insignificant. Interestingly, the positive outcomes were concentrated within the group of children from parents with lower socio-economic status. Neither age of enrolment nor the intensity of the program had a significant impact on results. The authors found that quality was the crucial determinant of the effects of ECEC and recommended that policies focused on increasing coverage should not compromise on quality.
Werner, C. D., Linting, M., Vermeer, H. J., & Van IJzendoorn, M. H. (2016). Do intervention programs in child care promote the quality of caregiver-child interactions? A meta-analysis of randomized controlled trials. <i>Prevention Science</i> , 17(2), 259-273.	Effect of targeted intervention programmes in professional childcare on the quality of caregiver-child interactions	2003-2012	Meta-analysis: 18	The authors conducted a meta-analysis of randomised controlled trials of targeted intervention programmes in professional child care. The search identified 18 articles, comprising 19 study samples and 16 different intervention studies, that met the inclusion criteria. Nine of the 16 intervention studies focused on caregiver sensitive responsiveness, and the remaining seven mainly targeted verbal communication skills and peer interaction. The authors found that targeted interventions were moderately effective in improving caregiver-child interactions on the classroom level and caregiver level. To a lesser extent, targeted interventions were effective in improving child behaviour.
Whitebook, M. (2003). Early Education Quality: Higher Teacher Qualifications for Better Living Environments. A Review of the Literature.	Relationship between teacher preparation and child outcomes in early childhood education	1989-2001	Narrative synthesis: 13	The author conducted a narrative synthesis of the effect of teacher qualifications on outcomes for three- to five-year-olds. The search identified 13 eligible articles, eight of which specifically addressed the question of the effect of a bachelor's degree and specialised early childhood training on preschool program quality. The author found that teacher preparation at the four-year college degree level with some specialised content in early childhood education or child development was the best way to achieve high-quality preschool programs that help children reach their full potential. However, there is evidence that alternative pathways to teacher preparation exist and may also be

effective, particularly for ensuring linguistic and cultural diversity in preschool programmes.
