Carefree in child care? : child wellbeing, caregiving quality, and intervention programs in center-based child care

Werner, C.D.

Carefree in child care?

Child wellbeing, caregiving quality and intervention programs in center-based child care

Claudia Denise Werner

Carefree in child care?

Child wellbeing, caregiving quality, and intervention programs in center-based child care

Carefree in child care?

Child wellbeing, caregiving quality, and intervention programs
in center-based child care

Proefschrift

ter verkrijging van de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr.dr. C.J.J.M. Stolker
volgens besluit van het College voor Promoties
te verdedigen op woensdag 18 juni 2014
klokke 10:00 uur door

Claudia Denise Werner

geboren te Leiderdorp in 1982

Promotiecommissie
Promotor:
prof. dr. M. H. van IJzendoorn

Co-promotores:
Dr. H. J. Vermeer
Dr. M. Linting

Overige leden:
Prof. dr. F. Juffer
Prof. dr. L. W. C. Tavecchio (Universiteit van Amsterdam)
Dr. J. C. de Schipper (Vrije Universiteit Amsterdam)

Printing: drukkerij Mostert, Leiden
Cover design: Ontwerpmetpit.nl

The studies in this dissertation were supported by a grant from the Netherlands Organization for Scientific Research (NWO Spinoza prize), awarded to M.H. Van IJzendoorn.

Contents

CHAPTER 1

General introduction

General introduction

Child care services support millions of families and enable many parents with young children to take part in the labor market (Organization for Economic Co-operation and Development; OECD, 2013). The use of center child care in Western countries has increased over the last three decades, and is nowadays the most frequently used type of non-parental care for children aged zero to four (OECD, 2013). The aim of the current dissertation is to shed more light on indicators of child care quality in center child care and to answer the question whether narrow-focused caregiver interventions are effective in improving child care quality.

Since research consistently has shown that the parent-child relationship is not necessarily negatively affected by the use of non-parental care in general (Ahnert, Pinquart, & Lamb, 2006; Goossens & Van IJzendoorn, 1990; Howes, Rodning, Galluzzo, & Myers, 1988; Love et al., 2003; Spieker, Nelson, Petras, Jolley, & Barnard, 2003), most parents consider it no longer controversial to make regular use of center child care. What is more, when the quality of care is high, child care attendance can even be beneficial for the child's cognitive and language development (Dearing, McCartney, & Taylor, 2009). It has been shown that when taking into account moderators and contextual factors such as child age, type of care, hours in care and social-economic status of the families, high quality care has advantages for child development (Belsky, 2006; McCartney et al., 2010; Vandell et al., 2010). However, when the quality is low, children's development can be negatively affected in the short term as well as in the long term, as was evidenced in the groundbreaking longitudinal study by the National Institute for Child Health and Human Development Early Child Care Research Network (NICHD ECCRN) in the U.S.A. (Vandell, Belsky, Burchinal, Steinberg, & Vandergrift, 2010). Before answering the question how low child care quality can be improved, we first need to define the concept of child care quality more clearly.

Defining quality of center child care

Scientists, policy makers and parents agree that high quality care can be achieved through four fundamental goals, which are providing children with a sense of security, enhancing their personal and social competence, and stimulating their socialization process (Riksen-Walraven, 2004). The extent to which a child care center succeeds in reaching these goals determines the quality of care (Riksen-Walraven, 2004). Quality of child care can further be defined in terms of distal factors and proximal factors, which contribute to achieving these four main goals. Distal factors are the more 'structural' aspects of the care, for instance the use of space in the room and furniture, caregiver education level, and group size. However, the most important indicators of quality in the model (see Figure 1) are formed by proximal factors: caregiver-child interactions,

peer interactions, and the interaction of the child with the physical environment (Riksen-Walraven, 2004). These proximal factors may be positively or negatively affected by distal factors: for instance more caregiver training and lower staff turnover rates are related to higher quality caregiver-child interactions (De Schipper, Tavecchio, Van IJzendoorn, & Linting, 2003; Gerber, Whitebook, & Weinstein, 2007).

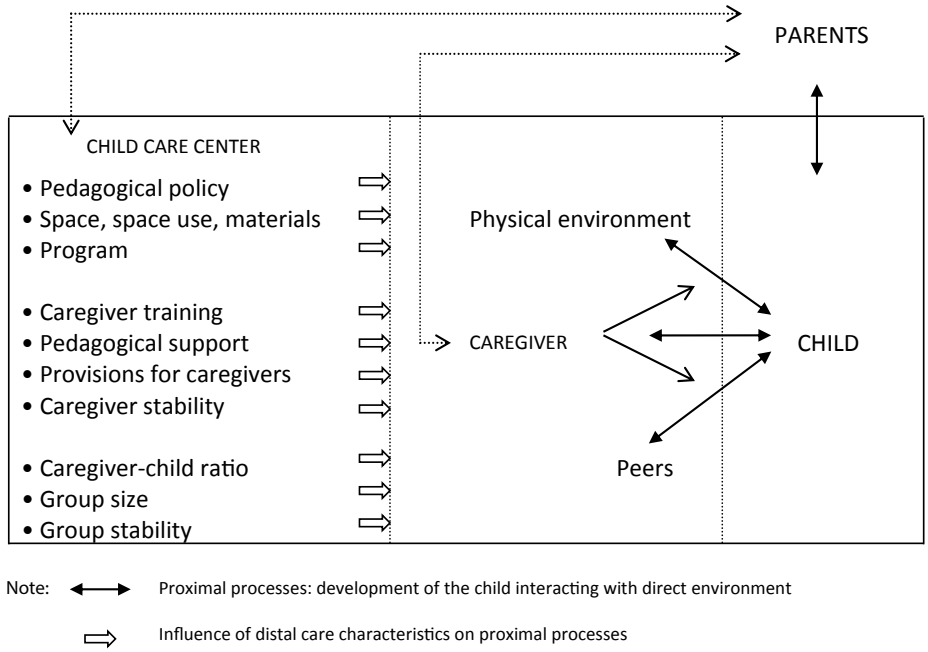


Figure 1. Model of factors that have direct or indirect effects on child wellbeing and development in center child care (Riksen-Walraven, 2004)

Noise as an additional indicator of child care quality

Although the model in Figure 1 provides a comprehensive picture of indicators of child care quality, we suggest combining this model with the environmental chaos theory (Evans & Wachs, 2010). This theory states that people's physical and psychological health can be negatively affected by high levels of environmental chaos. The most important indicators of environmental chaos are noise, crowding (the number of people in one place), and chaos itself, as indicated by a lack of routines, planning of activities and order (Matheny, Wachs, Ludwig, & Philips, 1995). These indicators are highly relevant to child care settings and may be important predictors for child wellbeing (Evans & Wachs, 2010). While not yet much research has been done with regards to environmental chaos in the child care setting, in family environments high levels of noise, chaos and crowding have been associated with more frequent negative caregiver-child interactions (Evans & Wachs, 2010). Moreover, indicators of crowding, such as higher child-to-caregiver ratios

and larger group sizes, have been shown to be related to fewer positive caregiver-child interactions (De Schipper, Riksen-Walraven, & Geurts, 2006).

A core element of environmental chaos is noise. Particularly, previous research has demonstrated high levels of noise in child care centers to be detrimental to child cognitive performance (Evans, 2006; Shield & Dockrell, 2003) and stress regulation (Evans, 2006). In home-based child care, higher noise levels have been associated with lower levels of child wellbeing (Linting, Groeneveld, Vermeer, & Van IJzendoorn, 2013). In light of the above, noise fits in well with the proximal processes described in the Riksen-Walraven model (2004). That is, on the one hand, noise results from caregiver-child interactions or peer interactions in the physical environment, but on the other hand noise may also affect these processes. Moreover, noise is likely to be influenced by distal factors such as group size, materials and space. Although the specific details regarding how to best integrate particular indicators of environmental chaos theory into the model in Figure 1 needs to be discussed further, it is clear that their associations with child care quality and child wellbeing requires more attention. In the current dissertation we therefore also investigate the association between child care quality and indicators of environmental chaos.

Our starting point is that quality of child care includes indicators at the center level, the group level, the caregiver level, and the child level, that all interact with one another in their prediction of child wellbeing. However, the indicator of child care quality that consistently has been shown to be most predictive of optimal child development is the caregiver-child relationship (Ahnert et al., 2006; Goossens & Van IJzendoorn, 1990; Howes & Smith, 1995).

Why is the caregiver-child relationship so important?

Children who attend child care depend on other caregivers than their parents for emotional support and physical care. Similar to the child in the family environment, the child in child care needs to be nurtured, supported and calmed throughout the day by a stable, sensitive caregiver. The relationship between the professional caregiver and the child is fundamental for the child's emotion regulation, feelings of security and social development (Ahnert et al., 2006). A positive relationship can be reached through consistent sensitive caregiving, that is, professional caregivers who observe and interpret children's signals correctly and respond to them in prompt and adequate ways (Ainsworth, Bell, & Stayton, 1974). Just like parental sensitivity in the home environment, caregiver sensitivity is a main predictor of child social-emotional development for children in child care (Ahnert et al., 2006; Goossens & Van IJzendoorn, 1990). Frequent positive caregiver-child interactions with a sensitive caregiver help children to build a secure relationship with the caregiver. In this vein, caregivers may function as attachment figures (De Schipper, Tavecchio & Van IJzendoorn, 2008), that

is, trusted persons who are available and responsive to the children when they are distressed or anxious (Bowlby, 1969). Positive relationships with sensitive caregivers are especially important for children growing up in home environments of poor quality, because these relationships may protect against adverse developmental outcomes of these home environments (Badanes, Dmitrieva, & Watamura, 2012; Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2011; Vesely, Brown, & Mahatmya, 2013; Watamura, Morrissey, Philips, McCartney, & Bub, 2011). Whilst parental sensitivity and adequate quality of the home environment for child social-emotional development are fundamental (Rijlaarsdam, 2014) over and beyond child care quality (Belsky et al., 2007; Hungerford & Cox, 2006), the quality of care and sensitivity of professional caregivers are also essential. Children receiving poor quality care both at home and in their out-of-home child care are specifically at increased risk for adverse outcomes, because of the so-called double risk situation (Watamura et al., 2011). This finding implicates that one mechanism to protect these children from negative life outcomes is to improve the quality of child care. The question arises how the caregiver-child relationship and general quality in center child care may be improved.

Improving child care quality

Internationally, two types of programs have addressed the issue of improving child care quality: narrow-focus and broad-focus intervention programs. The main difference between the two types of programs is that in the broad-focus programs, also referred to as Early Childhood Education (ECE) programs, child care is viewed as an intervention in itself, aiming to provide children, especially children from disadvantaged families, with better opportunities (Hungerford & Cox, 2006). In these programs, child care curricula are often combined with home visits, parent training and child health care, thus involving multiple social environments, often in a longitudinal design. ECE programs are not so much initiatives to improve child care quality itself, but these programs merely try to enhance child cognitive and emotional development through high quality child care. Well-known ECE programs in the U.S.A. are the High/Scope Perry Preschool program (Belfield, Nores, Barnett, & Schweinhart, 2006), the Carolina Abecedarian (Campbell et al., 2008; Campbell et al., 2012), the Chicago Longitudinal School Readiness program (Jones, Bub & Raver, 2013) and Head Start (Shager et al., 2013).

In contrast, narrow-focus programs which are characterized by their relatively short-term design are often more specifically designed to improve one or two specific indicators of child care quality. In the ongoing debate on how to reach high quality child care with low costs a variety of narrow-focus programs have been developed. Many narrow-focus programs aim at improving indicators of child care quality at the child level, such as child cognitive development (for a review see Chambers, Cheung, Slavin, Smith, & Laurenzano, 2010). Other intervention programs focus on specific indicators of quality at the center

level, such as regulations regarding infection prevention (Binns & Lee, 2010), sleep safety (Moon, Calabrese, & Aird, 2008), nutrition, and physical activity (see e.g. Larsson, Ward, Neelon, & Story, 2011; Ward, Vaugh, McWilliams, & Hales, 2010). Other programs target the proximal indicators of child care quality: the caregiver-child relationship and child social-emotional wellbeing. Two well-known programs in this domain are the Incredible Years Teacher Programs (Webster Stratton, 2004), and the Tools of the Mind program (Bodrova & Leong, 1996; Diamond & Lee, 2011).

In this dissertation we provide an overview of randomized trials with narrow focus programs in child care settings. Our aim is to provide the current state-of-the-art regarding interventions in child care and to examine whether these programs succeed in improving child care quality as indicated by caregiver-child interaction and child social-emotional wellbeing. Finally, after reviewing the literature, we present two empirical studies conducted in the Netherlands. To provide some context to these studies specific characteristics of Dutch child care services and child care quality are presented next.

Child care in the Netherlands

In the Netherlands, three main types of formal child care services for children under four years of age can be distinguished. The first type, center child care (in Dutch: *kinderdagverblijf* or *kindercentrum),* is an out-of-home arrangement where two or three professional caregivers are responsible for a group of maximum 16 children. Home-based child care (in Dutch: *gastouderopvang*), the second type, is an arrangement more comparable to the home environment: One professional takes care of a maximum of six children in her or his own home. Children may attend center care and home-based care on a part-time or full-time basis, although most children attend part time (OECD, 2013). In contrast, preschool playgroups (in Dutch: *peuterspeelzaal*), the third type of child care service, are part-time-only arrangements for children between 2.5 to 4 years old. Preschool playgroups were originally designed to prepare children from disadvantaged backgrounds for primary school; therefore the focus is mainly on cognitive educational goals. Group sizes are comparable to those in center child care. Recently, official initiatives have been put forward to integrate all formal care and education services for children 0 to 12 years at the organizational level as well as in a spatial sense, including center child care and preschool playgroups (Sardes, 2012; www.nji.nl). The current dissertation focuses primarily on center child care.

Although Dutch center child care took an internationally leading position in terms of quality in the early 1990s, a steep decline in general quality was observed from 1995 to 2008, leaving the Netherlands with very few centers of high quality and a majority of centers of only minimal to adequate quality (Fukkink, Gevers Deynoot-Schaub, Helmerhorst, Bollen, & Riksen-Walraven, 2013; Vermeer et al., 2008; Vermeer, 2012). Over the years, more centers of moderate and low quality have emerged. Although an

increase in average quality has been observed in the most recent child care assessment by the Netherlands Consortium for Child Care Research in 2012 (Fukkink et al., 2013), a substantial part of the centers continued to be of inadequate to moderate quality, with only a few of high quality. Moreover, the stability of quality within the centers was low (Fukkink et al., 2013), strengthening the argument that even centers of adequate to high quality should be cautiously monitored given the fact that their quality may fluctuate over the years.

Especially after formal child care was opened to the market forces with the implementation of the Dutch Child Care Act of 2005, which resulted in a large increase in the number of for-profit centers (Noailly & Visser, 2009), the quality of center care has come under scrutiny in the Netherlands. Worries about the lack of capacity to monitor all centers by local authorities and the high demand for child places have resulted in speculations about the causes of the relatively low quality level of child care in the Netherlands. More recently concerns have been raised about for-profit centers that have started to further compromise the quality of care they provide, primarily due to the economic downturn (www.fnv.nl). Until 2010 demands for child care increased steadily (Centraal Plan Bureau, 2011). However, since 2011 the financial crisis has resulted in high drop-out rates of children in child care. As a consequence of their sudden decrease in income, many child care centers were faced with bankruptcy and many of the remaining organizations are struggling to keep their staff and provide adequate care (www.fnv.nl). The abovementioned developments are reason for concern, particularly in light of the fact that high quality child care ensuring good physical and emotional care for young children is needed in order to provide optimal child social-emotional wellbeing and cognitive development.

When we look at initiatives for quality improvement in the Netherlands, the discussion about program effectiveness is mostly limited to the preschool playgroup programs (in Dutch: *Voor- en Vroegschoolse Educatie)* that focus on language development and cognitive stimulation for minority children (Driessen, 2003). In center child care, programs that supposedly enhance child cognitive development and caregiver-child interaction (e.g. Kaleidoscoop, Ben ik in Beeld, Piramide, Uk & Puk; www.nji.nl) are used on a large scale, yet these programs still lack solid evidence considering their effectiveness (Meij, Mutsaers, & Penning, 2009; Rutten, 2009). Without well-designed randomized controlled trials, selection bias may obscure results, causes and effects remain unclear and a best practice approach (as opposed to an evidence based approach) leads to implementation of costly programs that may not enhance child care quality and child wellbeing at all.

An attachment-based intervention in center child care

There is a clear need for more solid effectiveness studies to intervention programs focusing on the caregiver-child relationship in center child care, internationally as well

as in the Netherlands. Therefore we designed a randomized controlled trial to test an attachment-based program that has been proven effective in various family settings and in home-based child care. The Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD, Juffer, Bakermans-Kranenburg, & Van IJzendoorn, 2008a) was originally designed to be used in families with children in the preschool age. The program aims to improve the parent-child relationship and was tested in several randomized trials in different populations. Maternal sensitivity improved as a result of the intervention for mothers with insecure attachment representations (Klein Velderman, Bakermans-Kranenburg, Juffer, & Van IJzendoorn, 2006), insensitive mothers (Kalinauskiene et al., 2009), mothers of adopted children (Juffer, Bakermans-Kranenburg, & Van IJzendoorn, 2005, 2008b), mothers with children high in externalizing behavior problems (Mesman et al., 2008; Van Zeijl et al., 2006), mothers with eating disorders (Stein et al., 2006; Woolley, Hertzmann, & Stein, 2008), and mothers of low SES at risk for maltreating their children (Negrão, Pereira, Mesman, & Soares, 2013). Groeneveld and colleagues adapted the program for use in home-based child care (VIPP-CC) and showed that the program enhanced general caregiving quality and the caregiver's attitudes towards sensitive caregiving and limit setting (Groeneveld, Vermeer, Van IJzendoorn, & Linting, 2011). These findings support the hypothesis that this intervention method could also be suitable for use in center child care. However, as home-based care and center-based care are quite different from one another in terms of group sizes, routines and physical environment, we cannot generalize the findings by Groeneveld et al. (2011) to center child care. Therefore a randomized controlled trial to evaluate the effectiveness of the VIPP-CC in center child care is needed.

Outline of the dissertation

This dissertation begins with an overview of randomized controlled trials of narrow-focus intervention programs in child care through a meta-analysis on the effectiveness of caregiver training programs regarding child care quality outcomes at the classroom level, caregiver level, and child level in Chapter 2. Next, Chapter 3 presents the randomized controlled trial of the VIPP-CC in center child care and discusses results on caregiver sensitive responsiveness, caregiving attitudes and general quality. In the correlational study in Chapter 4 we investigate how indicators of environmental chaos in the child care centers, in particular noise, can be combined with other indicators of child care quality and how they are related to child emotional wellbeing. Finally, the findings of the meta-analysis and empirical studies are reflected upon and integrated in the theoretical framework in the general discussion in Chapter 5, where implications for practice and future research are provided.

CHAPTER 2

Do intervention programs in child care promote the quality of caregiver-child interactions? A meta-analysis of randomized controlled trials

Claudia D. Werner, Mariëlle Linting, Harriet J. Vermeer, Marinus H. Van IJzendoorn

Manuscript submitted for publication

Abstract

In this meta-analysis we report on the effectiveness of narrow-focus preventive intervention programs in child care settings targeted at child care professionals to improve child care quality, caregiver skills, and child social-emotional development. We exclusively included randomized controlled trials. From a series of four meta-analyses, we conclude that programs are moderately effective in improving overall caregiver-child interactions (k = 19, Hedges' g = 0.35) and in improving child care quality on the classroom level (k =11; Hedges' g = 0.39), the caregiver level (k = 10; Hedges' g = 0.44), and the child level (k = 6; Hedges' g = 0.26). Based on this meta-analysis, policy makers may reconsider current preventive intervention programs in child care: implementing low-cost evidence-based programs with a narrow focus may lead to better social-emotional development for millions of children under age five. Still, there is an urgent need for more and larger randomized controlled trials with a solid design and high quality measures, in order to shed more light on which child care components for which children are most critical in supporting children's socio-emotional development.

Keywords: meta-analysis, child care quality, intervention, randomized controlled trials, child social-emotional development

Introduction

How can the quality of caregiver-child relationships in professional child care be improved? This meta-analysis is the first to focus on randomized controlled trials of narrow-focus preventive intervention programs in professional child care, targeting caregiver-child interactions. The benefits of enhancing child care quality and preventing child social-emotional problems by implementing effective programs are substantial, given the fact that millions of children under the age of five could be reached. In the U.S.A. and most European countries, around 33% of children under the age of three and around 70% of children aged three to five are enrolled in formal child care (Organization for Economic Co-operation and Development, 2013). Therefore, preventive programs for child care that aim to improve caregiver-child interactions require attention. This meta-analysis is highly relevant to the field of preventive intervention science, because knowledge about effectiveness of programs with a narrow focus and therefore relatively low costs could provide governments and other funders with vital background information for their investments.

Empirical studies have shown that more positive caregiver-child interactions in professional child care are related to fewer behavior problems and higher cognitive-academic achievement in primary school (Belsky et al., 2007; Peisner-Feinberg et al., 2001). In many western countries there is still much debate on how to arrive at high quality care (Chambers, Cheung, Slavin, Smith, & Laurenzano, 2010) and how to find a balance between the costs and benefits of intervention programs, especially since more and more children are attending child care (Barnett, 2011; Chambers et al., 2010; Nelson, Westhues, & MacLeod, 2003; Weikart, 1998). Quality of child care includes indicators at the classroom level, the caregiver level, and the child level. At the classroom level, quality of care is indicated by the amount of space in the room, general classroom atmosphere, and general supervision of children (Harms, Clifford, & Cryer, 1998; Riksen-Walraven, 2004). At the caregiver level, caregiver-child interactions and the stimulation of peer interactions by the caregiver are important indicators of child care quality (Harms et al.,1998; Riksen-Walraven, 2004), whereas at the child level, child care quality is indicated by child cognitive development, social-emotional wellbeing, and peer interaction (Riksen-Walraven, 2004).

An important reason for governments to invest in interventions aimed at improving the quality of child care lies in the indirect cost reduction to society by providing children with better educational and professional opportunities (Burger, 2010; Heckman, 2006). A paradox in the cost-benefit debate is that authorities and funders call for evidence-based intervention programs, whereas they continue spending large amounts of money on the implementation of non-evidence-based programs and on ill-designed research projects that cannot answer the question of program effectiveness (Barnett, 2011;

Janus & Brinkman, 2010). So far, two fundamentally different routes have been taken to improve child care quality: (a) broad-focus programs, and (b) narrow-focus programs.

Bronfenbrenner's ecological model (1979) formed the basis of many broad-focus child care programs, also referred to as Early Childhood Education (ECE) programs. The core assumption of the model is that children are affected by experiences in multiple social environments simultaneously, which all contribute to child development. Broad-focus programs target various social environments at once, involving child care providers, parents, and teachers, to reach optimal outcomes (Ramey & Landesman Ramey, 1998). The programs are generally long-term (several years) and aim at disadvantaged children from low SES families (Hungerford & Cox, 2006). Well-known ECE programs in the U.S.A. are the High/Scope Perry Preschool program (Belfield, Nores, Barnett, & Schweinhart, 2006), the Carolina Abecedarian (Campbell et al., 2012), the Chicago Longitudinal School Readiness program (Jones, Bub, & Raver, 2013) and Head Start (Shager et al., 2013). Most of these programs have focused on the children's cognitive development (Blok, Fukkink, Gebhardt, & Leseman, 2005; Burger, 2010; Nelson et al., 2003) and mixed results regarding child outcomes have been reported (Barnett, 2011; Clarke & Campbell, 1998; Nelson et al., 2003). It is difficult to disentangle effective aspects when many program components and 'players' (e.g. parents, teachers, trainers) are involved. It should be noted that within some broad-focus programs great efforts have been made to systematically examine the programs' effects, using control groups and randomized assignment of participants (Campbell et al., 2012; Belfield et al., 2006). Impressive long-term advantages of early childhood education programs up to the age of 40 have been shown regarding participants' psychological well-being, employment and non-criminal behavior (Campbell et al., 2012; Belfield et al., 2006).

The goal of this meta-analysis is to evaluate the second type of preventive intervention programs in child care: narrow-focus programs. Narrow-focus programs target only one 'ecological environment' at a time and may be as effective on the child cognitive and social-emotional domains as broad-focus programs (Barnett, 2011; Burger, 2010; Nelson et al., 2003). The former are generally less costly because they are short-term, i.e. they are completed within a time span of several weeks or months. Moreover, these programs are generally easier to implement and allow researchers to study the effectiveness of particular aspects of the program, because of the narrow focus and the single environmental setting. Studying the effectiveness of narrow-focus programs may still be a complex task given the variation in how programs are implemented and how program fidelity is ensured. However, due to a lack of well-designed experimental evaluations (Barnett, 2011; Burger, 2010), it is unclear whether such narrow-focused interventions actually succeed in improving child-care quality on the caregiver level and classroom level, for instance caregiver-child interactions and general classroom atmosphere. Whereas child-focused trainings aimed at improving children's school

readiness are widespread, the caregiver-child relationship and child social-emotional development as targets of intervention in child care have been less thoroughly investigated (Blok et al., 2005; Chambers et al., 2010). Yet, for very young children caregiver-child interactions are highly important, because caregivers can provide them with feelings of security and may stimulate their development (Burchinal, Cryer, Clifford, & Howes, 2002; Vermeer & Bakermans-Kranenburg, 2008). Caregivers thus play a crucial role in children's social-emotional development. Addressing this role, Riksen-Walraven (2004) defined six important skills for professional caregivers: Sensitive responsiveness, respecting children's autonomy, structuring and limit setting, enhancing verbal communication, peer interaction, and developmental stimulation. These skills can be assessed from caregiver practices, attitudes, beliefs, or knowledge about caregiving. The question arises how these skills can be improved. With respect to intervention programs, one might ask whether 'training-on-the-job', that is *additional* caregiver training, may improve caregiver-child interaction skills even further, in order to enhance child social-emotional outcomes (Burchinal et al., 2002; Howes et al., 1992). However, researchers have struggled to reach solid conclusions about effective elements of caregiver training, because of a wide variety in focus, scale and design of the programs (Fukkink & Lont, 2007). With regard to their design, interventions on caregiver interaction skills and child social-emotional competence have been reported far less often in randomized trials than school readiness programs (Blok et al., 2005; Burger, 2010; Chambers et al., 2010).

Over the past decade, some meta-analyses on early childhood interventions in child care have been published (Blok et al., 2005; Fukkink & Lont, 2007; Nelson et al., 2003). Still, it is difficult to distill clear conclusions on the effectiveness of narrow-focus child care programs for child care quality and child outcomes. For instance, Blok et al. (2005) and Nelson et al. (2003) included both narrow-focus and broad-focus programs. Moreover, they focused not specifically on the caregiver-child relationship, but on child emotional well-being, parent-child relationships (Nelson et al., 2003) or education (Blok et al., 2005). Second, outcomes were not reported on the classroom level, but on the child level (Blok et al., 2005; Nelson et al., 2003) or the caregiver level (Fukkink & Lont, 2007). Therefore, no conclusions can be drawn on the effectiveness of programs on the classroom level, an indication for general quality of care. Finally, the three meta-analyses were restricted by including quasi-experimental studies, potentially confounding internal validity issues with conclusions about the effectiveness of the programs.

Research objectives

We focus on the effectiveness of narrow-focus intervention programs targeted at child care professionals in improving child care quality, caregiver skills, and child social-emotional development. Our study is different from previous meta-analytic reports in the field in at least two ways: (1) it only includes randomized controlled trials of

narrow-focus programs, and (2) beyond reporting overall effectiveness of the programs, results are reported separately for the three levels that represent child care quality: The classroom level, the caregiver level, and the child level. We conduct a series of meta-analysis to answer the main research question and additionally investigate which aspects may account for the effectiveness of interventions. In addition, we investigate which aspects of the care settings and the intervention programs may account for the effectiveness of preventive interventions in professional child care. For that purpose, we will distinguish different types of child care settings. Center child care and preschool are both arrangements with relatively large groups and multiple caregivers, whereas home-based care is more similar to the family setting, with only one caregiver present in a home-like environment. We will further distinguish intervention programs that were conducted in the context of subsidized Head Start settings, because Head Start centers share the same standard program and serve children from low income families. Regarding intervention characteristics, we focus on possible differential effects dependent on the focus of the intervention, the duration and intensity of the programs, the type of sessions that are provided to caregivers (group sessions, individual sessions, or both), and the use of video (see e.g. Bakermans-Kranenburg, Van IJzendoorn, & Juffer, 2003; Blok et al., 2005; Fukkink & Lont, 2007; Nelson et al., 2003). In addition, some studies may involve child curricula implemented by the caregiver, besides caregiver training. Child curricula generally consist of weekly activities around a certain theme for all children in the classroom that are described in a detailed activity manual for caregivers. The caregiver leads the activities, for instance, by inviting the children to discuss theme-related topics during circle time, role play or storybook reading. The child curriculum may be the basis of some intervention programs, whereas the caregiver training part may be small. In other programs, there may be no specific child curriculum, so that activities may be less structured, less frequent, or not directed at the group of children as a whole, but the caregiver training part may be more extensive. We would like to distinguish programs with and without such an intense child curriculum. We will examine whether the use of these child curricula influences the outcomes of the interventions. Finally, effect sizes may be different depending on the types of treatment for the control group.

Method

Literature search

To identify relevant studies the following electronic databases were systematically searched for articles published until 2013: Web of Science, SCIRUS, PsychInfo, WorldCat, ERIC, Google Scholar and Dissertation.org. The following keywords were entered: *child care, daycare, preschool, center-based care, home-based care*, *family-based care*, in combination with one or more of the following keywords: *intervention, staff training,*

teacher training, caregiver training, and child development. Subsequently, the reference lists of collected studies were searched for relevant studies. A flow chart of the literature search is shown in Figure 1. To be included in the meta-analysis, articles had to meet the following inclusion criteria: (1) the study design was a randomized controlled trial, (2) the language of publication was English, (3) the study was published (or available online) as an article in a research journal, or a doctoral thesis, (4) the topic of study was an intervention or training, targeting professional caregivers or teachers for typically developing children aged 0-5 in professional child care. The intervention may encompass the implementation of a standard curriculum for the children, if it was in combination with or through caregiver/teacher training. In addition, (5) the article should report on at least one of the following outcomes considering child care quality: Classroom quality of the child care setting; caregiver quality as indicated by practices, attitudes, beliefs, or knowledge about caregiving; or quality at the child level, as indicated by child social-emotional development, or child communication skills. Finally, (6) the results section should allow for calculation of effect sizes for outcome measures.

Considering the first criterion, a control group and randomized assignment to groups were required. Studies reporting placebo interventions for the control group were allowed, but only if the contents of these placebo interventions were unrelated to the programs in the intervention groups. Considering criterion 4, studies were excluded when they focused on caregivers for preschool children beyond the targeted age range, e.g. children in kindergarten; or if the intervention program targeted children directly, i.e. not through training of the caregiver. Community-based programs including parental involvement were also excluded. Regarding the outcome measures, studies were excluded if the intervention was exclusively aimed at aspects other than listed in criterion

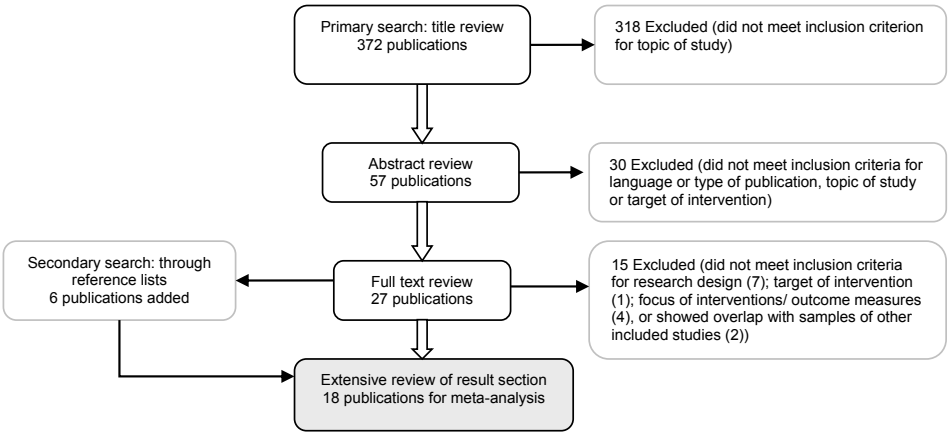


Figure 1. Flow chart of study selection process

5; for instance, classroom physical environment for health and hygiene practices, or indicators of school readiness. Following these criteria, 18 articles were included (see Figure 1). The studies were reviewed and coded by two independent coders. Inter coder reliability was satisfactory with .70 or higher (range .71 to 1.00) for both categorical variables (kappa) and numerical variables (intra class correlation) (Fukkink & Lont, 2007; Nelson et al., 2003). When information was reported unclearly, consensus was reached between the two coders through discussion.

Coding system

The studies were coded on study and intervention characteristics. We coded year of publication, publication outlet, country of the study, and research design, by distinguishing pretest- posttest design and posttest-only design. An overview of the most important characteristics of the studies is presented in Table 1. The most relevant aspects of the intervention programs are presented in Table 2. For type of care three categories were used: preschool, center-based care and home-based care. To distinguish clearly between different levels of randomization and outcomes, we refer to centers as well as preschools as 'centers'; to caregivers (in home-based care or center-based care) as well as teachers in preschools as 'caregivers'; and to groups (center-based care as well as classrooms in preschools) as 'classrooms'. The level of randomization was coded on the highest level: district level, center level, classroom level, caregiver level, or child level. Sample sizes (differentiated by intervention and control group) were coded on all levels for pretest and posttest. Regarding characteristics of the intervention programs, we coded the name of the program, number and duration (in hours) of intervention sessions and whether these were group sessions or individual sessions. We calculated the total number of hours of the intervention (individual sessions and group sessions apart and their sum), and the duration of the total intervention period (in months). If the authors only reported that the intervention lasted for 'a school year', we coded the intervention period as eight months. The time between pretest and posttest was coded (in months). In case authors reported that 'pretests were carried out at the beginning of the school year and posttest at the end of the school year', we coded the time between pretest and posttest as nine months, calculating one month extra on top of the time coded for the duration of the intervention. In addition, we coded the use of an additional child curriculum (yes or no), the use of video (video-feedback, video-modeling or no video), and the "treatment" for the control group (placebo intervention, waitlist, or care-as-usual).

Sample description

The included articles (n = 18) were published between 2003 and 2012. None of the relevant articles published before 2003 reported on a randomized controlled trial. There

was overlap between studies in two cases. Girolametto, Weitzman, and Greenberg (2004) and Girard and Girolametto (2011) reported on the same sample, but on different outcome levels. In a different study, the same holds for Raver et al. (2008) and Zhai, Raver, and Li-Grining (2011). Neuman and Cunningham (2009) reported in one publication on two intervention programs in two settings, providing four study samples. Finally, 19 study samples were distinguished, describing sixteen different intervention studies. We clustered the studies according to the focus of the intervention, using the description of caregiver skills provided by Riksen-Walraven (2004). Nine studies targeted mainly caregiver sensitive responsiveness, and most of these also included aspects of respecting children's autonomy, structuring and limit setting. Seven studies were mainly focused on enhancing verbal communication and peer interaction. Developmental stimulation defined as stimulation of physical or cognitive development was not the main focus of any of the interventions described here, because for the current meta-analysis we selected studies that aimed at improving caregiver-child interactions and children's socio-emotional development. In Appendix 1, we briefly describe the theoretical background and goals of the intervention programs.

Meta-analytic procedures

We conducted meta-analyses using the Comprehensive Meta-Analysis software (Borenstein, Hedges, Higgins, & Rothstein, 2009). Four meta-analyses were conducted: an overall effect size was calculated, and effect sizes were calculated for the three separate levels (classroom level, caregiver level and child level). Within four different sets of studies we calculated Hedges' g, a variant of Cohen's *d* that is more appropriate for small samples.

The first set of studies considered the overall effect size of the randomized controlled trials on social interaction of caregivers and children. To ascertain similarity of outcome measures we excluded variables not reflecting social behavior or interaction, and very specific caregiver aspects, such as stress and job satisfaction. One study (Zhai et al., 2011) was not included because the only reported outcome considered caregiver reported job stress.

The second, third and fourth sets of studies included studies on the classroom level, caregiver level, and child level, respectively. However, it was not possible to use 'outcome level' as a moderator in the overall meta-analysis, because the majority of studies reported on more than one outcome level. Therefore, separate datasets were created on the three different levels and three meta-analyses were conducted. Within the three separate datasets outcomes were meta-analytically combined using CMA, which leads to a conservative estimate of the overall effect size (Borenstein et al., 2009). Random-effects models were applied, accounting for sampling error between as well as within studies. A random-effects model allows for the possibility that there are random

Table 1
Sample characteristics (moderators in bold)

Study	Country	**SES of the children**	**Type of care**	N of centers		N of classrooms		N of caregivers		N of children		Level of randomization	Outcome level(s)
				I	C	I	C	I	C	I	C		
Baker-Henningham et al. (2009)	Jamaica	Low SES	preschool	3	2	15	12	15	12	22 per classroom		center	class, caregiver
Barnett et al. (2008)	U.S.A.	Low SES	preschool	1 in total		7	11	7	11	81	117	child	class, child
Cain et al. (2007)	U.S.A.	NR	center	27 in total		NR	NR	16	19	NR	NR	caregiver	caregiver
Domitrovich et al. (2007)	U.S.A.	Low SES: Head Start	preschool	NR	NR	10	10	NR	NR	100	100	center	child
Domitrovich et al. (2009)	U.S.A.	Low SES: Head Start	preschool	NR	NR	22	22	42	42	NR	NR	center	class, caregiver
Driscoll et al. (2010)	U.S.A.	Low SES: Head Start	preschool	29 in total		NR	NR	NR	NR	38	40	center	child
Fukkink & Tavecchio (2010)	The Netherlands	NR	center	NR	NR	NR	NR	52	43	NR	NR	center	caregiver
Girard et al. (2011)[1]/ Girolametto et al. (2004)	Canada	NR	center	4	3	NR	NR	8	9	32	36	center	caregiver / child[1]
Girolametto et al. (2003)	Canada	NR	center	2	2	8	8	8	8	32	32	center	caregiver
Groeneveld et al. (2011)	The Netherlands	NR	home	-	-	24	24	24	24	7 per caregiver		caregiver	class, caregiver
Izard et al. (2004)	U.S.A.	Low SES: Head Start	preschool	NR	NR	7	9	NR	NR	58	58	caregiver	child
Neuman & Cunningham (2009)[2]	U.S.A.	Low SES	center (PD)	53	35	53	35	53	35	NR	NR	caregiver	class
Neuman & Cunningham (2009)[2]	U.S.A.	Low SES	center (PD+ CO)	53	35	53	35	53	35	NR	NR	caregiver	class
Neuman & Cunningham (2009)[2]	U.S.A.	Low SES	home (PD)	32	22	32	22	23	22	NR	NR	caregiver	class
Neuman & Cunningham (2009)[2]	U.S.A.	Low SES	home (PD + CO)	34	22	34	22	34	22	NR	NR	caregiver	class
Raver et al. (2008)[1]/ Zhai et al. 2011[1]	U.S.A.	Low SES: Head Start	preschool	9	9	18	18	48	42	455 in total		center	class / caregiver[1]
Rusby et al. (2004)	U.S.A.	NR	home	-	-	18	20	18	20	NR	NR	caregiver	class
Rusby et al. (2008)	U.S.A	NR	home	-	-	18	20	33	30	8.7 per caregiver		caregiver	class
Snyder et al. (2011)	U.S.A.	Low SES: Head Start	preschool	3	2	9	5	18	10	76	33	center	caregiver, child

N= number at posttest; I = Intervention group; C = control group; NR= Not reported; PD = professional development; CO = coaching; 1 Two publications report on one study but on different outcome levels; 2 Publication covers two settings and two intervention programs

Table 2
Characteristics of intervention programs and study designs (moderators in bold)

Study	Name of the program	**Focus**	**Program duration (months)**	**Pretest to posttest (months)**	**Group training (total hours)**	**Individual coaching (total hours)**	**Use of child curriculum**	**Use of video**	**Treatment control group**
Baker-Henningham et al. (2009)	IY Teacher Training	SR	7.00	9.00	56.00	14.00	yes	modeling	placebo
Barnett et al. (2008)	Tools of the mind	SR	7.00	9.00	41.00	20.00	yes	no	placebo
Cain et al. (2007)*	Focus-Follow-Talk	SR	3.00	-	4.00	2.25	no	no	placebo
Domitrovich et al. (2007)	PATHS	VC/PI	9.00	9.00	24.00	NR	yes	no	waitlist
Domitrovich et al. (2009)**	PATHS-REDI	VC/PI	12.00	13.00	34.00	160.00	yes	modeling	care-as-usual
Driscoll et al. (2010)	Banking Time	SR	1.50	NR	NR	NR	no	no	waitlist
Fukkink & Tavecchio (2010)	Video Interaction Guidance	SR	NR	NR	0.00	4.00	no	feedback	care-as-usual
Girard et al. (2011)[1] / Girolametto et al. (2004)[1]	Learning Language & Loving it	VC/PI	1.50	2.00	7.00	1.50	no	modeling + feedback	placebo
Girolametto et al. (2003)	Learning Language & Loving it	VC/PI	3.00	4.00	20.00	3.00	no	modeling + feedback	waitlist
Groeneveld et al. (2011)	VIPP-SD	SR	5.00	6.00	0.00	9.00	no	feedback	placebo
Izard et al. (2004)	Emotions Course	VC/PI	5.00	7.00	2.00	4.00	yes	no	care-as-usual
Neuman & Cunningham (2009)[2] - center based	Professional Development	VC/PI	3.50	8.00	45.00	0.00	no	modeling	waitlist
Neuman & Cunningham (2009)[2] -home based	Professional Development	VC/PI	3.50	8.00	45.00	0.00	no	modeling	waitlist
Neuman & Cunningham (2009)[2] - center based	Professional Development + Coaching	VC/PI	8.00	8.00	45.00	48.00	no	modeling	waitlist
Neuman & Cunningham (2009)[2] - home based	Professional Development + Coaching	VC/PI	8.00	8.00	45.00	48.00	no	modeling	waitlist
Raver et al. (2008)[1] / Zhai et al. (2011)[1]	IY Teacher Training	SR	7.00	8.00	38.00	45.00	no	modeling	placebo
Rusby et al. (2004)	Carescapes – 1 module	SR	0.03	1.00	3.00	0.00	no	modeling	waitlist
Rusby et al. (2008)	Carescapes – 3 modules	SR	1.50	2.50	9.00	0.00	no	modeling	waitlist
Snyder et al. (2011)	IY Teacher Training	SR	3.00	8.00	15.00	2.25	no	no	care-as-usual

*Posttest-only design;**Only posttest results reported; NR = Not reported; SR= Sensitive Responsiveness; VC= Verbal communication; PI=Peer interaction; [1] Two publications report on one study but on different outcome levels; [2] Publication covers two settings and two intervention programs

differences between studies that are associated with variations in procedures, measures, settings, that go beyond subject-level sampling error and thus point to different study populations. Homogeneity of the sets of effect sizes were tested with Q-statistics. It should be noted that in these meta-analyses the number of studies was relatively small (ranging from k = 6 to k =19), so that interpretation of the Q-value as an indication of the homogeneity of outcomes should be done with caution, another reason to use the random model. We used the trim-and-fill method to calculate the effect of potential data censoring or publication bias on the outcome of the meta-analyses. Using this method, a funnel plot is constructed of each study's effect size against the sample size or the standard error (usually plotted as 1/*SE* or precision). If no publication bias were present, this plot would show the shape of a funnel, because studies with smaller sample sizes and larger standard errors are expected to have increasingly large variation in estimates of their effect sizes as random variation becomes increasingly influential, whereas studies with larger sample sizes have smaller variation in effect sizes (Duval & Tweedie, 2000; Sutton, Duval, & Tweedie, 2000). However, since smaller non-significant studies are less likely to be published, studies in the bottom left hand corner of the plot are often absent. With the 'trim and fill' procedure, the k right most studies considered to be symmetrically unmatched are trimmed and their missing counterparts are imputed or 'filled' as mirror images of the trimmed outcomes. This leads to an adjusted estimate of the combined effect size taking into account potential publication bias. Regarding reported statistics the following decisions were made. We based the analyses as much as possible on raw data (pre and post means, standard deviations, and sample size). However, none of the studies, except the study by Groeneveld, Vermeer, Van IJzendoorn, and Linting (2011), reported on correlations between pre- and posttest outcomes. Therefore, we used a default estimate of 0.5 for the pre-post correlations, which is in accordance with the empirical values reported by Groeneveld et al. (2011) (0.43. and 0.56). An estimate of .50 was also applied in the meta-analysis by Fukkink and Lont (2007). When means and standard deviations were reported in combination with an F-value for the interaction, but no correlation between pre- and posttest, we entered the data with F-values to avoid uncertainty about the pre-post correlation. When reported sample sizes differed for pretest and posttest, posttest sample sizes were used. When only regression coefficients were reported, we selected the unstandardized b-values, corrected for pretest score and as few as possible covariates. When only significance levels were reported instead of exact p-values, we used the significance levels as a conservative estimate of the p-values. In one study, two different treatment effects were investigated in two different contexts: home-based care and center-based care (Neuman & Cunningham, 2009). We analyzed this sample as four separate studies, equally dividing the control group across experimental groups within each context to prevent individuals from being included more than once in the meta-analysis. Regarding the robustness of the effect sizes, we

applied Rosenthal's criterion, implying that if the fail safe number is larger than 5*k* +10 (with *k* the number of studies in the meta-analysis), it can be concluded that the effect size might be rather robust.

Moderator analysis

In the overall meta-analysis, we investigated the role of potential moderators. With respect to the type of care we compared center-based (including preschools, *k* = 14) versus home-based settings (*k* = 5) and programs within Head Start settings (*k* = 6) versus those without Head Start (*k* = 13). Considering program characteristics, we compared interventions that used video (aggregating in one variable the use of individual video-feedback and/or video modeling, *k* = 12) with those without video (*k* = 7), and programs including a child curriculum (*k* = 5) versus those without (*k* = 14). We compared program duration, distinguishing between 'shorter than 6 months' (*k* = 11) and 'longer than 6 months' (*k* = 7). For one study duration was not reported. The intensity, that is the total amount of hours dedicated to training (group training and individual training combined), was categorized as 'less than 10 hours' (*k* = 7) and '10 hours or more' (*k* = 11). The number of individual training hours was not reported in two studies, so that total training hours could not be calculated. Furthermore, we distinguished programs with individual training sessions (irrespective of the number of hours) (*k* = 15) from those without (*k* = 4). We used Riksen-Walraven's model (2004) to compare programs by their focus of intervention: caregiver sensitive responsiveness (*k* = 9) versus verbal communication and peer interaction (*k* = 10). Finally, we regarded treatment for the control group as a moderator, distinguishing two categories: 'placebo' (*k* = 6) versus 'no placebo' (*k* = 13, including categories 'care-as-usual' and 'waitlist'), e.g. Blok et al. (2005).

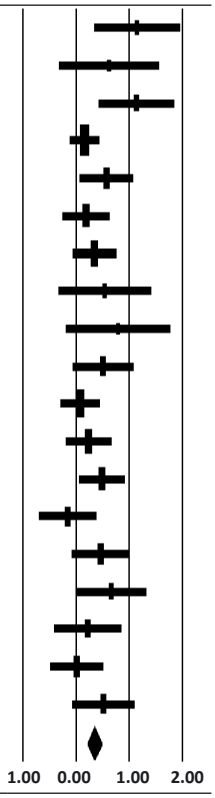
Results

Overall effect

The combined effect of the 19 randomized controlled studies with combined outcome measures on all levels was Hedges' *g* = 0.35 (SE = 0.07), CI= 0.21 - 0.48, *p* <.001) and there was no indication for heterogeneity (Q = 22.50, *p* = 0.21). The fail-safe number was 171, indicating that 171 studies with null results would be needed to reduce the overall significant effect to non-significance. After applying the trim and fill method, the adjusted effect size was Hedges' *g* = 0.25 (CI = 0.10 - 0.40, Q = 40.96), including six trimmed studies. The necessity to trim studies pointed into the direction of publication bias against small studies with small effect sizes (Borenstein, Hedges, Higgins, & Rothstein, 2009). The effect sizes of the studies are presented in Table 3.

Table 3
Overall outcomes on social-interaction measures from RCTs in child care: Effect sizes (Hedges' g, 95 % confidence interval) and significance

First author	*Hedges' g*	*SE*	*lower*	*upper*	*p*	*Relative weight*
1. Baker-Henningham	1.15	0.41	0.35	1.95	0.005	2.51
2. Barnett	0.62	0.48	-0.31	1.56	0.193	1.90
3. Cain	1.14	0.36	0.43	1.84	0.002	3.17
4. Domitrovich (2007)	0.16	0.14	-0.11	0.44	0.256	12.58
5. Domitrovich (2009)	0.57	0.26	0.07	1.07	0.026	5.63
6. Driscoll	0.19	0.23	-0.26	0.63	0.408	6.86
7. Fukkink	0.35	0.21	-0.06	0.76	0.093	7.73
8. Girard /Girolametto (2004)	0.54	0.44	-0.33	1.41	0.226	2.17
9. Girolametto (2003)	0.79	0.50	-0.19	1.77	0.115	1.73
10. Groeneveld	0.51	0.29	-0.06	1.08	0.079	4.61
11. Izard	0.08	0.19	-0.29	0.44	0.679	9.02
12. Neuman - center based [PD]	0.24	0.22	-0.19	0.66	0.273	7.27
13. Neuman - center based [PDC]	0.49	0.22	0.06	0.92	0.026	7.16
14. Neuman - home based [PD]	-0.16	0.27	-0.70	0.38	0.567	5.03
15. Neuman - home based [PDC]	0.46	0.27	-0.08	1.00	0.092	5.04
16. Raver	0.66	0.34	0.01	1.32	0.048	3.58
17. Rusby (2004)	0.22	0.32	-0.41	0.85	0.491	3.88
18. Rusby (2008)	0.02	0.25	-0.48	0.51	0.954	5.78
19. Snyder	0.51	0.30	-0.07	1.10	0.087	4.35
Total	*0.35*	*0.07*	*0.21*	*0.48*	*<.001*	



Note. [P] = Professional development group; [PDC]= Professional development plus coaching group

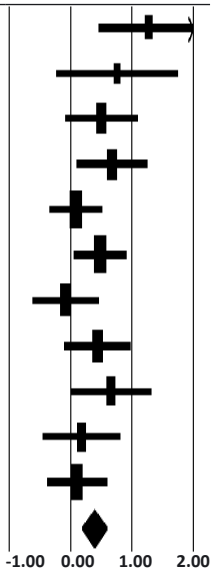
Effects on the classroom level, caregiver level and child level

The combined effect of 11 randomized controlled studies with outcome measures on the classroom level was Hedges' *g* = 0.39 (*SE* = 0.10), CI = 0.19 - 0.59, *p* <.001) and there was no statistical indication for heterogeneity (Q = 13.56, *p* = 0.19). The fail-safe number on the classroom level was 52. After applying the trim and fill method, the adjusted effect size was Hedges' *g* = 0.33 (CI = 0.11- 0.54, Q = 20.09), including two trimmed studies. The effect sizes on the classroom level are presented in Table 4. For the caregiver level, the combined effect of ten randomized controlled studies was Hedges' *g* = 0.44 (*SE* = 0.12), CI = 0.21 - 0.68, *p* <.001) and there was no statistical indication for heterogeneity (Q = 13.43, *p* = 0.14). The fail-safe number on the caregiver level was 48. After applying the trim and fill method, the adjusted effect size was Hedges' *g* = 0.30 (CI = 0.04- 0.55, Q = 24.95), including three trimmed studies. The effect sizes on the caregiver level are presented in Table 5. For the child level, the combined effect of six randomized controlled studies was Hedges' *g* = 0.26 (*SE* = 0.08), CI = 0.11 - 0.40, *p* = .001) and there was no statistical indication for heterogeneity (Q = .70, *p* = 0.98). The fail-safe number on the child level was 11. After applying the trim and fill method, the adjusted effect size was Hedges' *g* = 0.24 (CI = 0.10- 0.38, Q = 1.25), including one trimmed study. Effect sizes on the child level are shown in Table 6.

Table 4

Classroom level outcomes from RCTs in child care: Effect sizes (Hedges' g, 95 % confidence interval) and significance

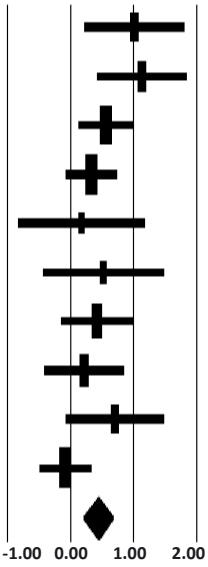
First author	*Hedges' g*	*SE*	*lower*	*upper*	*p*	*Relative weight*
1. Baker-Henningham	1.28	0.42	0.46	2.01	0.002	5.13
2. Barnett	0.76	0.51	- 0.23	1.75	0.130	3.66
3. Domitrovich (2009)	0.51	0.30	- 0.08	1.10	0.093	8.67
4. Groeneveld	0.68	0.29	0.11	1.25	0.020	9.06
5. Neuman - center based [PD]	0.09	0.22	0.34	0.51	0.680	13.63
6. Neuman - center based [PDC]	0.48	0.22	0.06	0.91	0.027	13.45
7. Neuman - home based [PD]	- 0.08	0.27	- 0.61	0.46	0.779	9.96
8. Neuman - home based [PDC]	0.44	0.27	- 0.09	0.98	0.106	9.98
9. Raver	0.66	0.34	0.01	1.32	0.048	7.33
10. Rusby (2004)	0.18	0.32	- 0.45	0.81	0.569	7.85
11. Rusby (2008)	0.11	0.25	- 0.38	0.60	0.654	11.28
Total	*0.39*	*0.10*	*0.19*	*0.59*	*.000*	



Note. [PD]= Professional development group; [PDC]= Professional development plus coaching group

Table 5
Caregiver level outcomes from RCTs in child care: effect sizes (Hedges' g, 95% confidence interval) and significance

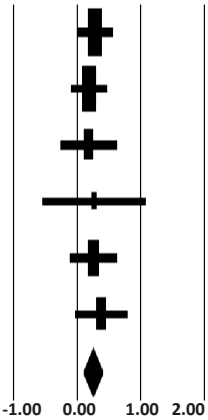
First author	*Hedges' g*	*SE*	*lower*	*upper*	*p*	*Relative weight*
1. Baker-Henningham	1.02	0.43	0.23	1.81	0.012	6.88
2. Cain	1.14	0.36	0.43	1.84	0.002	8.17
3. Domitrovich (2009)	0.57	0.22	0.13	1.00	0.011	15.25
4. Fukkink	0.33	0.21	- 0.07	0.74	0.108	16.27
5. Girolametto (2004) [a]	0.18	0.51	- 0.83	1.19	0.726	4.62
6. Girolametto (2003)	0.53	0.49	- 0.43	1.49	0.282	4.99
7. Groeneveld	0.42	0.29	- 0.14	0.99	0.141	11.15
8. Rusby (2004)	0.22	0.32	- 0.41	0.85	0.496	9.63
9. Snyder	0.71	0.40	- 0.07	1.49	0.075	7.01
10. Zhai	-0.07	0.21	- 0.49	0.34	0.724	16.03
Total	*0.44*	*0.12*	*0.21*	*0.68*	*0.000*	



Note. [a] Outcomes on child level of this study reported by Girard et al. (2011)

Table 6
Child level outcomes from RCTs in child care: Effect sizes (Hedges' g, 95% confidence interval) and significance

First author	*Hedges' g*	*SE*	*lower*	*upper*	*p*	*Relative weight*
1. Barnett	0.28	0.15	0.00	0.57	0.05	27.39
2. Domitrovich (2007)	0.19	0.14	- 0.08	0.47	0.17	28.54
3. Driscoll	0.19	0.23	- 0.26	0.63	0.41	11.28
4. Girard (2011) [a]	0.27	0.42	- 0.54	1.09	0.51	3.34
5. Izard	0.26	0.19	- 0.11	0.63	0.17	16.39
6. Snyder	0.38	0.21	- 0.03	0.79	0.07	13.06
Total	*0.26*	*0.08*	*0.11*	*0.40*	*0.001*	



Note. [a] Outcomes on caregiver and classroom levels of this study reported by Girolametto et al. (2004)

Moderator analysis

We investigated the role of potential moderators. No moderator effects were found for type of care (Q = 1.52, *p* = 0.22), Head Start versus no Head Start (Q = .67, *p* = .41), use of video (Q = .06, *p* = .80), use of a child curriculum (Q = .07, *p* = .79), program duration (Q = 1.60, *p* = .21), and focus of intervention (Q = 1.16, *p* = .28). No moderator effects were found for overall program intensity: Programs with less than ten hours in total were not significantly different in their effectiveness than programs with ten or more hours of training (Q = 0.750, *p* = 0.39). The presence of an individual training component was a significant moderator (Q = 4.198, *p* = .040). Programs with individual training sessions for caregivers led to higher effect sizes (Hedges' *g* = 0.41 (*SE* = 0.07), CI = 0.27 - 0.55, *p* <.001) than programs without individual training (Hedges' *g* = .09 (*SE* = 0.14), CI = -0.18 - 0.36, *p* = .52). Treatment for the control group was also a significant moderator (Q = 9.431, *p* = 0.002), showing that programs with a placebo training for the control group (Hedges' *g* = 0.75 (*SE* = 0.15), CI = 0.46 – 1.05, *p* <.001) were more effective than programs without a placebo treatment (Hedges' *g* = 0.25 (*SE* = 0.06), CI = 0.12 - 0.37, *p* <.001).

Power analysis

We performed a priori power analyses with G*power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007), calculating the sample size needed to detect the aggregated effect size (i.e., the assumed population effect size) with a power of 0.80 and a two-sided significance level of 0.05. For the overall meta-analysis on outcomes related to social behavior and interaction (aggregated Hedge's *g* = 0.35), a minimum sample size of *N* = 260 would be needed. Note that for the trim-and-fill adjusted effect size of *g* = 0.25, we would even need more than 500 subjects. Post hoc power analysis showed that the power of the included studies to detect the aggregated effect size ranged from 0.10 for the study with the smallest sample size (Girolametto et al., 2003; *N* = 16, combined reported effect size *g* = 0.79) to 0.69 for the study with the largest sample size (Domitrovich et al., 2007; *N* = 200, combined reported effect size *g* = 0.16). We also performed power analyses considering the meta-analytic outcomes on the classroom, caregiver, and child level separately. Results showed that we would need 210 classrooms, 166 caregivers, and 468 children to detect the aggregated effect sizes of 0.39, 0.44, and 0.26, respectively. On all three levels, post hoc power to detect the aggregated effect size was far below the required 0.80, the highest value being 0.56 (Fukkink & Tavecchio, 2010; *N* = 95; caregiver outcomes with combined reported effect size *g* = 0.33).

Discussion and conclusions

In this meta-analysis we showed that child care preventive intervention programs with a narrow focus on caregiver-child interactions are moderately effective in improving child care quality on three levels: classroom quality, caregiver skills, and, to a lesser extent,

child behavior. According to Cohen's criteria the reported effects are small to medium. For our overall meta-analysis, which should be considered the most important one, we found a rather robust effect size. The significant effect sizes on the three separate levels are based on fewer studies, resulting in larger confidence intervals. Specifically, when we also take into account possible unpublished studies (with trim-and-fill) the effect sizes become rather small and relatively unstable. Still, meta-analysis has the advantage compared to a narrative review that it is replicable and takes trends in primary studies into account in computing combined effect sizes. We consider it informative that caregiver training seems to be indeed most effective for the caregiver, and less for the classroom level quality and child level quality.

Although it is remarkable that only in the last ten years randomized controlled trials on narrow-focus intervention programs have been published, it is promising to see a shift towards more solid research designs in the field of child care and early childhood education. Still, there is much room for improvement. For instance, the studies included in our meta-analyses were systematically under-powered as a result of the small number of subjects. Moreover, in many studies we were confronted with lack of information needed to perform a meta-analysis, for instance, sample size, randomization procedures, raw data (means and standard deviations), and pretest-posttest correlations. In addition, we were confronted with missing information on intervention characteristics such as duration and spacing of training sessions (e.g. Driscoll et al., 2010; Domitrovich et al., 2007; Fukkink & Tavecchio, 2010). Furthermore, although the studies were rather homogeneous in design, they were at the same time heterogeneous with regard to other aspects such as sample size, SES of the sample, focus of the intervention programs, duration and frequency of training sessions, and outcome measures. The relatively small set of studies ($k = 19$) restricted our exploration of identifying effective components within and between studies. A remarkable moderator was the presence of placebo training for the control group. Programs without a placebo intervention were less effective, which is in contrast with Blok et al. (2005). It is possible that the studies with a placebo intervention for the control group do not report more effective programs, but merely represent methodologically higher quality studies with better outcome assessments. Our conclusion that overall program duration and intensity did not moderate program effectiveness should be considered with caution, because the small number of studies forced us to dichotomize these moderators in our analyses. Our findings are in line with those of Blok and colleagues (2005), but in contrast with those of Nelson et al. (2003) who concluded that more lengthy and more intense programs in preschool are more effective. It should be noted that studies in the meta-analysis by Nelson et al. (2003) directly targeted children, not caregivers.

Unfortunately, we were unable to test the '*less is more'* hypothesis which states that short-term intervention programs with relatively few sessions are more effective

than long-term programs with many sessions (Bakermans-Kranenburg et al., 2003). In our meta-analysis it was not possible to distinguish exact numbers of training sessions, because of missing data and variation across studies in type (group and/or individual sessions), contents and duration of sessions. Instead, we distinguished studies with and without an individual training component. We cautiously conclude that there seems to be added value of individual coaching on top of group training sessions. However, the small number of studies did not allow us to further analyze whether intensity of the individual training component also matters. It would not be surprising that individual attention for the caregivers leads to improvement of their skills. Still, it is important to have better understanding of the minimal dose that is needed for individual training, so that costs can be reduced.

Some limitations of the current meta-analysis should be mentioned. First, the number of pertinent studies is rather small which restricts moderator analyses and prevented us from conducting multivariate meta-regression. Second, we found evidence for publication bias that might have led to inflated estimates of effect sizes. With the trim-and-fill method we have tried to create a more valid estimate. Third, the studies included in the meta-analyses were underpowered which might reflect on the overall meta-analytic outcome. We suggest taking the current findings as a tentative description of the current state-of-the-art that shows how promising the intervention approach might be.

An important conclusion for the field and policy makers is that focused training programs to improve caregiver skills by targeting caregiver-child interactions are moderately effective. In effect, not only caregivers improve their behaviors, also general quality of the classroom increases. Ultimately, even though children were not directly targeted in the intervention programs described here, they do seem to benefit from these types of trainings. The effect on the child level regarding child social-emotional behavior was small, yet significant. For policy makers in the field of preventive intervention, our findings implicate that relatively low-cost programs in child care can be effective in preventing child problem behavior for millions of children. Funders and authorities may want to reconsider their current prevention programs or caregiver trainings to improve child care quality. We point out that there still is a need for more, and especially larger, randomized controlled trials. It is time for a multi-site randomized trial to improve child care quality, as a next step from the ground-breaking study by the National Institute of Child Health and Human Development Early Child Care Research Network (NICHD-ECCRN; Belsky et al., 2007; Burchinal et al., 2013). Well-designed intervention studies may teach us what critical components for which children and their families are most critical in terms of socio-emotional development. Only when studies are conducted in a solid design with enough power and high quality measures can we start to advise policy makers which evidence based programs to implement nation-wide to increase child care

quality. Effective preventive intervention programs with a narrow focus could then start to play a key role in improving the wellbeing of many young children in professional child care.

Appendix 1:
Original name, authors, background and aims of the programs evaluated in studies included in current meta-analysis.

Name of original program	Program evaluated by	Goal and underlying theory
		Focus on sensitive responsiveness
Banking Time (Pianta & Hamre, 2001)	Driscoll et al. (2010)	Based on the assumption that caregiver-child relationships serve as resources for children to enhance development and emotion-regulation skills. A key aspect is non-directive interaction between caregiver and child. In a one-on-one setting caregivers practice observing the child's actions and describing them aloud, and labeling the child's feelings and emotions. Relational themes are brought up by the caregiver, to build a supportive relationship.
Carescapes – Setting up to Support Social Development (Rusby et al., 2004; Rusby et al., 2008); *Proactive approaches for managing children's behavior, Understanding and dealing with problem behavior* (Rusby et al., 2008)	Rusby et al. (2004) Rusby et al (2008)	The underlying theory is that caregiver-child relationships and child social competence can be promoted by improving the physical environment. Topics of discussion are how a safe and orderly environment, the arrangement of the physical space, selection of materials and toys, and setting schedules and routines may elicit children's social skills. In two additional modules structuring and limit setting were added to caregiver target behavior.
Focus-Follow-Talk (Cain et al., 2007)	Cain et al. (2007)	To improve joint attention in caregiver-child interactions, which is seen as an important precursor for social-emotional skills and language development. Caregivers are coached to focus on the object of the child's interest, follow the child's lead by engaging in non-interfering ways and talk about the object in ways that promote language development.
Incredible Years Teacher Training Program (IYTTP) (Webster-Stratton, 2004)	Baker-Henningham et al. (2009) Raver et al. (2008) Snyder et al. (2011) Zhai et al. (2011)	Aim is to prevent and reduce child problem behaviors. Caregivers are trained to use labeled praise to promote positive behavior, and to use clear rules and routines to improve the caregiver-child relationship and child social-emotional skills. Stress-reduction workshops for caregivers and mental health coaching for children are available in adapted versions. The child curriculum consists of 14 weekly lessons of 30 minutes: Circle time discussions, songs, role play and group activities, with the goal to capture children's anger management, their recognition of emotions, and the use of school rules.
Video Interaction Guidance	Fukkink & Tavecchio (2010)	Program rooted in a somewhat general communication theory implying that positive caregiver-child relationships enhance child social-emotional development. During video guidance sessions, there is a focus on caregiver's interactive behaviors such as turning towards the child, making eye contact, recognizing the child initiatives, following the child, and acknowledge actions of the child and oneself.

Video-feedback Intervention to Promote Positive Parenting and Sensitive Discipline (VIPP-SD) (Juffer, Bakermans-Kranenburg, & Van IJzendoorn, 2008)	Groeneveld et al. (2011)	Originally developed for families and adapted to a version suited for child care. The program is rooted in Bowlby's (1969) attachment theory and Patterson's (1982) coercion theory. The caregivers are trained to recognize children's signals of exploration and contact seeking behaviors, to take the child's perspective, and to use empathy, induction and praise in order to set limits in a sensitive manner.
		Focus on verbal communication and peer interaction
Early Childhood Experiences in Language Arts (Machado, 2003)	Neuman & Cunningham (2009)	The program stresses the importance of child care providers' knowledge of child language and literacy development, as well as general interaction skills. Caregivers are advised how to ask children open-ended questions, elicit discussions, and support interaction.
Emotions Course (Izard, 2001)	Izard et al. (2004)	The program is rooted in Emotion Theory. The author reasons that improvements in emotion knowledge and regulation will lead to pro-social behavior, because children learn how to communicate and empathize from their emotions. The program stresses self-regulation and aims to improve child social-emotional development by enhancing communication skills. The child curriculum consists of 22 lessons, including activities with puppets, storybooks, posters and games to foster emotion recognition, expression and utilization.
Learning Language and Loving It – The Hanen Program for Early Childhood Educators (Weitzman, 1994)	Girard et al. (2011) Girolametto et al. (2004) Girolametto et al. (2003)	Caregivers are trained to follow the child's lead, encourage turn-taking, and use responsive labels, thereby expanding and extending the children's use of language.
Promote Alternative Thinking Strategies (PATHS) (Domitrovich, Greenberg, Kusché, & Cortes, 2005)	Domitrovich et al. (2007) Domitrovich et al. (2009)	The rationale is that by improving awareness, expression of emotions, and self-regulating skills in young children, problem behaviors can be prevented. The program was extended by adding *Research-based Developmentally Informed* (REDI) components, including interactive reading activities, sound games and alphabet activities to improve caregiver strategic use of language. Stimulation of verbal communication by the caregiver is emphasized. Additional goals are to improve child vocabulary, narrative, and syntax. The child PATHS curriculum consists of 30 weekly lessons, including circle time activities and games, art projects and storybook reading.
Tools of the mind (Bodrova & Leong, 1996)	Barnett et al. (2008)	Vygotsky-based program that aims to improve child language skills, social-emotional development and self-regulation. Key elements are the caregivers' scaffolding behavior and facilitation of learning opportunities. A child curriculum of 40 weekly activities is the core of the program, including circle time activities, dramatic play, and elements to enhance self-regulation, memory, cognition, and attention.

CHAPTER 3

Video-feedback intervention in center-based child care: A randomized controlled trial

Claudia D. Werner, Harriet J. Vermeer, Mariëlle Linting, Marinus H. Van IJzendoorn

Manuscript submitted for publication

Abstract

In the current study, we aimed to improve center-based child care quality with a short-term, attachment-based program: The Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline in Child Care (VIPP-CC). Caregivers (*N* = 64) from child care centers in urban areas in the Netherlands participated in our pretest-posttest randomized controlled trial. The VIPP-CC was effective for increased observed caregiver sensitive responsiveness in the group setting and led to a more positive attitude towards caregiving and limit setting. Post hoc analyses revealed that the intervention effect was apparent for caregiver sensitive responsiveness in structured play situations. The VIPP approach can now be expanded from the family setting to out-of –home group settings with larger groups of children. This is a promising conclusion for millions of children enrolled in center child care from a very young age.

Keywords: Center-based child care, intervention, caregiver training, randomized controlled trial

Introduction

Center-based child care is an important support system for parents of young children. Quality of center child care has been subject of heated debates, and most of the research efforts have concentrated on trying to describe the consequences of center care on the development of children (Love et al., 2003; Lowe Vandell, Belsky, Burchinal, Steinberg, & Vandergrift, 2010; Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2011; Votruba-Drzal, Levine Coley, Maldonado-Carreno, Li-Grining, & Chase-Lansdale, 2010). Much less research has been devoted to the improvement of quality of center care although the need for careful experimental work showing how to enhance child care quality has been emphasized by parents, professionals, and policy makers (Besharov & Morow, 2006; Janus & Brinkman, 2010) . In the present study we report one of the few randomized control trials aimed at enhancing quality of center child care for children aged 0-4 years. A video-feedback intervention to promote positive parenting with an emphasis on sensitive discipline is adapted to center care and its effects on caregiver sensitivity, caregiver attitudes and general child care quality are tested.

For optimal social-emotional and cognitive development children need stable attachment figures who are available and responsive to them when they are distressed or anxious (Bowlby, 1969). As many young children across western countries attend formal child care, this need extends to child care settings. For instance, in the Netherlands, 40% of the children aged zero to four years are in formal child care (OECD, 2013). A sensitive caregiver in the child care setting, who responds promptly and adequately to the child's signals and provides comfort and security (Ainsworth, Bell, & Stayton, 1974) may fulfill the crucial role of secondary attachment figure (Ahnert, Pinquart, & Lamb, 2006; Badanes, Dmitrieva, & Watamura, 2012; Barnas & Cummings, 1994; De Schipper, Tavecchio, & Van IJzendoorn, 2008; Goossens & Van IJzendoorn, 1990; Howes & Spieker, 2008; Vermeer & Bakermans-Kranenburg, 2008). The quality of attachment relationships between children and their professional caregivers can be predicted by caregiver sensitivity and frequency of interactions (De Schipper et al., 2008). Ahnert and colleagues (2006) suggested in their meta-analysis that 'group sensitivity', directed at the group of children, but not caregiver sensitivity directed at the individual child, was predictive of the child's attachment security towards the professional caregiver. Higher group sensitivity is related to higher general child care quality (Gerber, Whitebook, & Weinstein, 2007), which covers both structural aspects and process quality. Interactions between caregivers and children are an important aspect of process quality (Howes, Philips, & Whitebook, 1992a; Riksen-Walraven, 2004). Structural aspects refer to the more "fixed" elements of the child care setting, such as space and furnishing, play materials, the number of children and caregivers in the room and the daily care routines (Howes, Philips, & Whitebook, 1992a). Higher caregiver sensitivity and general quality

are related to more formal and specialized training (Gerber et al., 2007; Howes, Philips, & Whitebook, 1992b; Torquati, Raikes, & Huddleston-Casa, 2007). Thus, an intervention program involving caregiver training may be effective in improving caregiver sensitivity, and, to a lesser extent, general child care quality. Center-based care is the most common type of non-parental care for children in the age range of 0-4 (OECD, 2013), which points out the possible impact of improving the quality of center-based child care.

Only few of the many intervention programs in child care have been tested in randomized controlled trials (Besharov & Morrow, 2006). Moreover, the focus of interventions is often school readiness or child behavior, but not the caregiver-child (attachment) relationship. The Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD; Juffer, Bakermans-Kranenburg, & Van IJzendoorn, 2008) is an attachment –based intervention that is effective in enhancing parental sensitivity. Moreover, a recent randomized control trial showed that the VIPP for Child Care (VIPP-CC), a variant of the VIPP-SD for group child care, was also found to be effective for caregivers in home-based child care (Groeneveld et al., 2011), where mostly one professional caregiver takes care of a small group of children (generally up to seven) in her own home. As a result of the intervention, general caregiving quality improved and caregivers had more positive attitudes toward sensitive caregiving and limit setting (Groeneveld et al., 2011). Given the need for quality improvement (Burchinal, Cryer, Clifford, & Howes, 2002) and the emphasis on the caregiver-child attachment relationship in center-based care, we applied the VIPP-CC to center-based child care. The current randomized controlled trial aims to investigate whether the VIPP-CC is effective in enhancing caregiver sensitive responsiveness, improving general child care quality and increasing caregiver's attitudes towards sensitive caregiving and limit setting in center-based care.

Method

Recruitment and randomization

Participants in this study were professional caregivers in child care centers. For recruitment we targeted child care centers in the western part of the Netherlands, making special efforts to include centers in less privileged urban areas. Letters of invitation were sent to 180 organizations that exploited a wide range of centers. Managers were informed that the aim of the study was to compare the effectiveness of two types of training: Video-feedback training and coaching by telephone. In our design, the latter was used as a placebo intervention for the caregivers in the control group. Child care centers were eligible for participation if they met the following inclusion criteria: at least 50% of the children within a group were to be older than 2.5 years, because the instrument to assess general quality of the centers is aimed at children 2.5 to 5 years old. Thus, toddler groups

(for children 2-4 years) and mixed age groups (for children 0-4 years) were included. Furthermore, parents of at least three children attending the group on the same day of the week should provide written consent. One group per center was randomly selected for participation. Professional caregivers from eligible groups had to have a degree in early childhood education, be working for a minimum of two days (16 hours) per week in a fixed group, and be available and willing to participate during the study period of approximately six months. From each group, one caregiver who met these inclusion criteria was randomly selected. An overview of recruitment and participants is shown in Figure 1.

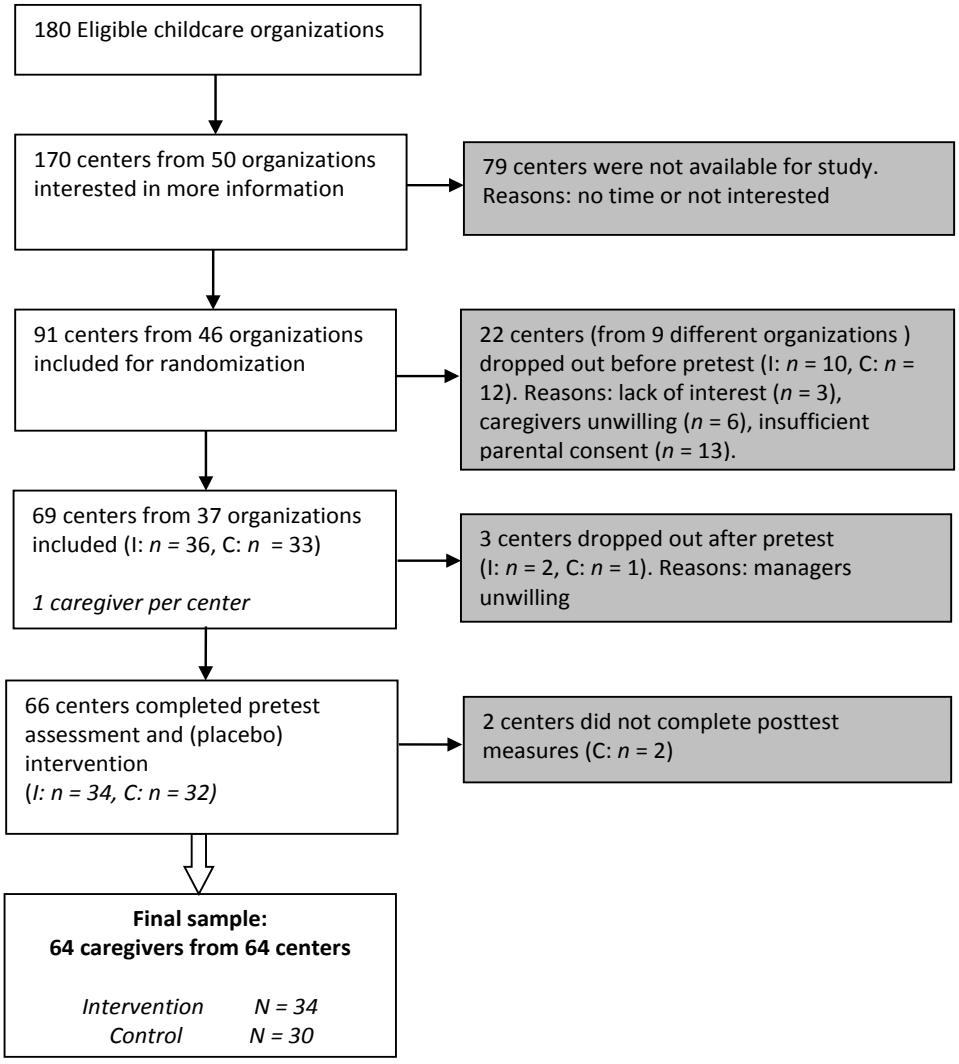


Figure 1. Flow chart of recruitment (I=intervention group, C=control group)

Randomization was performed hierarchically: Centers were first randomly assigned to either the intervention condition or the control condition. Next, within each center, one group was randomly selected to participate, and third, within each group, one caregiver was randomly selected. Finally, 64 caregivers completed both pre- and posttest (34 in the intervention group and 30 in the control group) and are included in this study.

Participants

The intervention group and control group were compared on background characteristics on center level, caregiver level and child level. In total, 27% of the centers (n = 17) were located in less privileged areas. Percentages did not differ significantly between the experimental group (21%, n = 7) and control group (33%, n = 10) (X^2 (1) = 1.33, p = .25). Background questionnaires were not returned by 11 managers, resulting in information from 29 centers in the intervention group and 24 centers in the control group. No significant differences were found between the intervention group and control group on years of existence, number of groups per center, the number of children in the center, and child places per center. Regarding age range of the children, 43 toddler groups and 18 mixed age groups participated. For three groups the age range was not reported. The distribution of the age groups was not significantly different for the intervention group (24 toddler groups, 8 mixed age groups) and the control group (19 toddler groups and 10 mixed age groups) (X^2 (2) = 2.30, p = .32).

For caregiver characteristics, no significant differences were found between the intervention group and control group on age, years of experience in child care, years of working on the particular group, number of working hours, level of education and nationality (see Table 1). The majority of caregivers had the Dutch nationality (93% in the

Table 1.
Characteristics of the caregivers

	Intervention (*N*= 29)		Control (*N*= 24)		
	M	*SD*	*M*	*SD*	*p- value*
Caregiver age	32.45	8.87	31.42	8.68	.67
Years working in childcare	8.71	6.17	6.66	7.67	.28
Years working on group	3.76	3.37	3.66	6.31	.94
Working hours per week	30.50	6.33	29.15	4.96	.40
Highest level of education					
high school (%)	3%		12%		
vocational training (%)	83%		76%		
bachelor or master degree (%)	14%		12%		.49*

*Note. * Pearson Chi square test:* X^2 (2) = 1.44, p = .49

intervention group and 84% in the control group; X^2 (5) = 6.38, p = .27) and were born in The Netherlands (in the intervention group 89% and in the control group 79%; X^2 (4) = 8.93, p = .06).

With respect to child characteristics, no significant group differences were found for age of the children or the number of children attending the group, neither at pretest nor at posttest. On average, ten children were present during the pretest (intervention group M = 10.12, SD = 2.95; control group M = 10.28, SD = 2.37, p = .82) and the posttest (intervention group M = 9.52, SD = 4.00; control group M = 9.90, SD = 2.67, p = .66). Age of the children ranged from three to 48 months. At pretest the mean age of the youngest child in the group was 21.42 months (SD = 9.46), whereas the mean age of the oldest child was 44.06 months (SD = 3.90). These values did not differ significantly between groups (for youngest and oldest child p =. 84 and p = .96, respectively).

Procedure

All caregivers were visited for a pretest assessment from 9:00 AM until 1:00 PM. Afterwards, questionnaires were sent to the caregivers and the center managers. During the pretest, a shortened version of the Early Childhood Environment Rating Scale-Revised (ECERS-R; Harms, Clifford & Cryer, 1998; Vermeer, 2012) was administered. In addition, caregiver behavior was videotaped with a digital video camera at three predetermined time points during regular activities. All videotaped episodes were rated afterwards on caregiver sensitive responsiveness. The first session of the (placebo) intervention was held within three weeks after the pretest. Posttests took place 2 to 4 weeks after the last session of the (placebo) intervention. Procedures for posttest assessments were similar to those of the pretests, including the administration of the shortened ECERS-R and the videotaped observation. In addition, three structured play tasks of the caregiver with four children were videotaped, i.e. storybook reading, putting together a jigsaw puzzle and tidying up the jigsaw pieces. These episodes were rated afterwards on caregiver sensitive responsiveness. After posttest, caregivers filled out a questionnaire on their attitude towards caregiving. Caregivers and managers completed evaluation forms. The duration of the project from pretest to posttest was approximately six months.

The intervention. The Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD; Juffer et al., 2008) was adapted for use in home-based child care in a previous study, resulting in the VIPP-SD for Child Care or VIPP-CC (Groeneveld et al., 2011). The intervention aims to improve sensitive behavior of the caregiver by providing personal video-feedback on sensitive responsiveness in daily situations as well as the use of sensitive discipline in challenging caregiver-child interactions. The program elaborates on four themes regarding sensitivity: (1) recognizing the child's exploration versus attachment behavior, (2) recognizing the child's signals, which is accomplished by

taking the child's perspective, i.e. "speaking for the child", (3) explaining the relevance of prompt and adequate response to the child's signals, and (4) sharing emotions. In addition, four themes of sensitive disciplining are addressed: (1) using inductive discipline and distraction methods, (2) using positive reinforcement, (3) giving sensitive time-outs, and (4) showing empathy towards the child in disciplining situations. The program consists of six biweekly visits that are carried out according to an elaborate protocol. The last two visits, so-called 'booster sessions' are used to repeat the themes of all previous sessions.

Each visit starts with videotaping the caregiver on the group in three or four semi-structured five-minute play tasks with a group of four children. Afterwards, the 15- to 20-minute video-taped observations of the previous visit are reviewed and discussed for approximately one hour according to the theme of that visit. Generally, the video feedback hour is planned directly after the videotaping, or during lunch break of the caregivers.

The play tasks are protocol-led for each visit and the materials are brought along by the intervener. Examples of the play tasks are story book reading, playing with Lego, playing with hand puppets, singing songs together and building a tower of blocks. Play tasks especially designed for the discipline themes are, for example, letting the children take turns during play and letting them wait before they can play with new toys. During the third visit a mealtime is discussed.

In between two visits, the video-taped sessions are reviewed by the intervener and prepared for the next intervention session, using micro-analysis. That is, the intervener writes comments in the form of a "script" on the observed behavior of the caregiver and the children for every 10 to 30 seconds per video clip. These comments are directed by guidelines in the intervention manual and the general theme of the next visit. Video clips are not edited or cut. Positive and successful interaction moments are emphasized, to reinforce positive behavior of the caregiver. General themes of child development and daily routines on the group are discussed as well.

A main issue in adapting the family intervention to the child care setting (for home-based and center-based care) was the focus on several children at the same time in a group setting. A sensitive caregiver in child care should be able to divide her attention adequately among the children and respond to them in a developmentally appropriate way. Perspective taking was modified so that speaking for the *child* became speaking for the *children*. Especially in mixed age groups this is an important topic, given the varying developmental levels. Another important difference is the nature of the relationship of the caregiver with the children, and accordingly, the relationship of the caregiver with the intervener. In The Netherlands, caregiver in child care are professionals who have completed a degree in early childhood education, so that during the intervention comments and themes for discussion should also be related to general

child developmental themes on a more professional level. The caregiver is viewed and referred to as the expert on this particular group of children. Moreover, no 'secondary' caregiver was invited during the booster sessions, which is common for the use of VIPP-SD in families. In the current study, we used the manual for the VIPP-CC in home-based care (Groeneveld et al., 2011) as a basis.

Finally, some minor adaptations – based on a pilot study in two centers – were made from the home-based day care version to a center-based version of the VIPP-CC in play material and tasks. First, the tidying-up task used in home-based care proved too easy as a discipline task for children in center-based care, and was therefore replaced by an adaptation of the "Do not touch"- task from the original VIPP-SD. A transparent, open box with attractive toys is placed in front of the children. The caregiver is instructed to have the children sit and wait for four minutes without touching the toys. In addition, instead of a TV screen and DVD player, we used laptops for viewing the video clips with the caregivers, because in the majority of child care centers TV equipment is not available. The outline and themes of the intervention were unchanged (for an overview see Appendix 1).

Program fidelity. Interveners were the first author and four MA students in child development who were trained as VIPP-SD interveners by experts at Leiden University. The training consisted of five days (40 hours) of group lectures, discussions and practice with micro-analyzing video clips. In addition, the interveners practiced the original intervention in a family setting with expert feedback. Next, they practiced the adapted VIPP-CC for use in a child care center. Weekly three hours supervision meetings were held with the first and third authors of this paper, who are both certified VIPP-SD interveners with experience in applying the intervention. Program fidelity during the research project was assured by two hour supervision meetings every three weeks. In addition, prepared "scripts" were reviewed and discussed by supervisors and co-interveners through email. Finally, interveners videotaped themselves during each third feedback session with the caregivers. These videos were viewed and discussed during supervision meetings.

Control group. The placebo intervention for the control group consisted of six biweekly protocol-led telephone calls of approximately 15 minutes, and was provided by three MA students and the first author. During the telephone calls, caregivers were interviewed about general developmental issues of four target children in their group, concerning playing, eating and sleeping habits at the center. After the fourth interview, caregivers were sent a brochure about play material for children 0-4 years old. The researchers carrying out the placebo intervention attended a two hours instruction meeting to get familiar with the protocol prior to the study. During the research project, bimonthly one-hour supervision meetings were held to discuss issues concerning contact with

caregivers or with the use of the manual. Caregivers in the control group were invited to attend a free, full-day workshop on the research project and key elements of the VIPP-CC approximately three months after completion of the study.

Measures

Caregiver sensitive responsiveness. Caregiver sensitive responsiveness was coded for nine videotaped episodes: Three ten-minute episodes in naturalistic situations ('unstructured') during pretest and posttest, and three five-minute play tasks ('structured') during posttest. For the structured play tasks the caregiver was asked to sit with a small group of four children and (1) read a storybook, (2) let the children put together a jigsaw puzzle, and (3) let the children tidy up the jigsaw puzzle. For pretest, an average score was computed across the three unstructured episodes, Cronbach's alpha = .67. A combined score for posttest sensitive responsiveness was computed, averaging the scores for unstructured and structured situations. Internal consistency was adequate, Cronbach's alpha = .69. The scale to code caregiver sensitive responsiveness was developed and validated by the Dutch Consortium for Child Care Research (NCKO, 2006; De Kruif et al., 2007). This is a group rating scale based on scales developed to measure sensitivity in the parent-child context (Ainsworth, Bell, & Stayton, 1974; Erickson, Sroufe, & Egeland, 1985).

Caregiver sensitive responsiveness refers to the degree to which the caregiver provides adequate and sufficient emotional support to all children in her care who need it, during stressful and non-stressful situations. It also refers to the level to which a caregiver is able to adequately divide attention among the children, showing interest in the children's activities and acknowledging their needs, emotions and competences. Sensitive responsiveness ratings are presented on a seven-point scale, ranging from (1) very low to (7) very high. A caregiver scoring high on the scale is very much involved with the children, and responds promptly and adequately to the signals of all children in her care, by taking the children's perspective. A caregiver scoring low on this scale may show either emotional distance or indifference towards the children, or she may be uninvolved with the children, because of administrative or cleaning tasks in the group and thus missing the signals of the children.

Two independent observers were trained to be reliable coders, using the official NCKO dataset. Intra class correlation for both coders was .75 (absolute agreement). Approximately 75% of the tapes were coded by the two observers. The remaining 25% were coded by the second author of this paper, who was involved in developing the scale. To obtain independency in ratings, researchers who administered the ECERS-R did not code video material from that particular setting. In addition, caregiver sensitive responsiveness was coded by different researchers for pretest and posttest situations and coders were blind to the condition of the group (intervention or control).

General child care quality. To asses general child care quality a shortened version of the ECERS-R (Vermeer, 2012; Harms et al., 1998) was administered in all participating groups at pretest and posttest. Training and administration of the full ECERS-R are time-consuming; therefore in the current study a shortened version with 18 items was used, which showed good psychometric properties (Vermeer, 2012). Internal consistency of the shortened ECERS-R was adequate, with Cronbach's alpha of .85 at pretest and .81 at posttest. Seven observers (six BA students and one MA student) were trained by the second author to be reliable on the shortened ECERS-R. The training included reviewing and discussing the items and field observations. Inter rater reliability was established to a criterion of 80% agreement for three consecutive observations for all observers. The mean percentage of agreement for these three observations was 90% (range 87% to 92%). To guarantee the independence of ratings, observers administered the ECERS-R in one group only once (either pretest or posttest) and coders were blind to the condition of the group (intervention or control).

Caregiver-child ratios and group size. During pretest and posttest, numbers of children and caregivers present were registered by the observers. The caregiver-child ratio was calculated as the number of children divided by the number of qualified caregivers in the room. Group size refers to the total number of people in the room (both children and caregivers).

Attitude towards caregiving. Immediately after the posttest, caregivers were handed a questionnaire regarding their attitude towards sensitive caregiving and limit setting (Bakermans-Kranenburg & Van IJzendoorn, 2003). Caregivers were asked to indicate their opinion on 15 items in a 5-point Likert format, ranging from (1) strongly disagree to (5) strongly agree. Examples of items are "Playing together with the children will prevent difficult behavior" and "The children need to learn that I will get angry when they do not listen to me" (reversed). Internal consistency was adequate, with Cronbach's alpha of .60.

Evaluation forms. To evaluate the research project and the intervention program, caregivers received an evaluation form. Caregivers were asked to express on 5-point Likert scales how much they learned from the training, and how informative and useful the training was to them (ranging from 1 *"not at all"* to 5 *"very much"*). They were also asked whether they would recommend the training to colleagues, to inexperienced and experienced caregivers and to the standard vocational curriculum for prospective caregivers (ranging from 1 *"certainly not"* to 5 *"certainly"*). Caregivers and center managers were invited to provide comments on the research project and the training.

Missing data

In our sample, 64 caregivers were included. However, due to missing data sample sizes are somewhat different among the analyses. Two caregivers did not have a score for posttest sensitive responsiveness, so they were not included in the analyses (so that *N* = 62). Of these caregivers, one caregiver in the intervention group could not be filmed during any of the unstructured episodes because on the day of observation she and the children went outside the whole morning and it was not possible to videotape her with the children in the public areas. One caregiver in the control group did not find the time to carry out the structured play tasks. Caregiver questionnaires on attitude and program evaluation were returned by 20 caregivers in the intervention group and 19 caregivers in the control group.

Statistical analysis

To investigate intervention effects on general child care quality and caregiver sensitive responsiveness, repeated measures ANOVAs were conducted to examine changes from pretest to posttest. Caregiver attitude towards sensitive caregiving and limit setting, and program evaluations were measured during the posttest only. Therefore, independent samples *t*-tests were used to investigate differences between the experimental and control group.

Results

Bivariate correlations

In Table 2 bivariate correlations between the main variables are presented. General quality during pretest was significantly associated with caregiver sensitive responsiveness during pretest, indicating that a higher score on the ECERS-R was associated with a higher score on the NCKO sensitivity scale. Furthermore, caregivers with more years of experience on their group showed higher general quality during pretest. Caregiver working hours were negatively associated to general quality and sensitive responsiveness during pretest. Caregivers who worked more hours per week had lower scores on general quality and on sensitive responsiveness during pretest. Surprisingly, a higher number of children per caregiver (child-caregiver ratio) was associated with higher scores of sensitive responsiveness during pretest.

Descriptive statistics

Descriptive statistics of the outcome variables are provided in Table 3.

Table 2
Pearson correlations between the main variables

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1. Group condition[a]	1.00														
2. Group size pretest	-0.01	1.00													
3. Group size posttest	-.06	**.66****	1.00												
4. Child/caregiver-ratio pretest	-.15	**.30***	.19	1.00											
5. Child/caregiver-ratio posttest	-.23	**.43****	**.52****	**.29***	1.00										
6. General quality pretest	-.13	-.18	-.11	-.05	-.14	1.00									
7. General quality posttest	-.16	.02	-.02	.20	.04	.21	1.00								
8. Cg. yrs. experience child care	.01	-.20	.02	.02	.06	.20	.07	1.00							
9. Cg. yrs. experience on group	.15	-.24	-.10	-.19	.01	**.28***	.15	**.69****	1.00						
10. Cg. work hours per week	.12	-.09	-.04	-.20	-.11	**-.37***	-.25	-.17	-.06	1.00					
11. Cg. Sens. resp. (U) pretest	**-.26***	.05	.20	**.30***	.05	**.45****	.20	.02	-.07	**-.29***	1.00				
12. Cg. Sens. resp. (U) posttest	-.12	-.14	-.08	.15	.13	.18	.06	.03	-.05	-.06	**.30***	1.00			
13. Cg. Sens. resp. (S) posttest	.07	.15	.05	.20	-.07	.14	.22	-.07	-.12	-.03	.20	.20	1.00		
14. Cg. Sens. resp. (C) posttest	.01	.03	-.04	.24	.05	.23	.18	-.01	-.08	-.07	**.32***	**.76****	**.79****	1.00	
15. Cg. Attitude	**.37***	-.01	-.03	.18	-.09	.21	.18	.07	-.17	-.21	.18	.14	.10	.17	1.00

*Note. * p <.05, ** p<.01* [a] Intervention group = 1, control group = 0 ; Cg = caregiver; U = unstructured; S =structured; C= Combined

Table 3
Descriptive statistics for the intervention group and control group during pretest and posttest

	Intervention group			Control group		
	N	*M*	*SD*	*N*	*M*	*SD*
General quality pretest	34	3.85	.81	30	4.04	.76
General quality posttest	34	4.10	.71	30	4.33	.73
Sensitive responsiveness U pretest	34	4.54	1.05	30	5.03	.76
Sensitive responsiveness U posttest	32	4.58	.99	28	4.80	.82
Sensitive responsiveness S posttest	33	4.90	.90	29	4.79	.85
Sensitive responsiveness C posttest	32	4.78	.73	30	4.76	.68

Note. U= unstructured episodes ; S= structured episodes; C= combination of structured and unstructured episodes, at least one episode of each type

Intervention effects

Caregiver sensitive responsiveness. At pretest there was a significant difference between the intervention group and the control group for sensitive responsiveness (p = .04), indicating that despite the randomization the control group scored significantly higher. To investigate intervention effects, we conducted three repeated measures ANOVAs for three different models. Table 4 presents the results for the models.

In the first analysis we combined the unstructured and structured episodes in the posttest, to test the overall effect of the VIPP-CC. The repeated measures ANOVA of this model (Model 1) showed that there was no significant main effect for time or group. However, there was a significant interaction effect (p = .03), indicating that the change over time in caregiver sensitive responsiveness was different for the two groups. The sensitive responsiveness of the caregivers in the intervention group increased from pretest to posttest, whereas the control group showed a decline over time.

Table 4
Results of repeated measures ANOVAs for three models

	Model 1 Posttest combined			Model 2 Posttest unstructured episodes			Model 3 Posttest structured episodes		
	F(1,60)	*p*	η^2	*F(1,58)*	*p*	η^2	*F(1,60)*	*p*	η^2
Time	.00	.96	.00	.88	.35	.02	.18	.67	.00
Group	2.06	.16	.03	3.89	.05	.06	1.06	.31	.02
Group*Time	4.72	.03	.07	1.26	.27	.02	4.37	.04	.07

To test the differential effects for sensitive responsiveness assessed in the structured and unstructured episodes, two additional analyses were conducted. The repeated measures ANOVA of Model 2, including unstructured episodes only, showed that there was no main effect for time (p = .35), nor group (p = .053), nor an interaction effect (p = .27). In the repeated measures ANOVA of the third model, including only the structured episodes of the posttest, no main effects for time (p = .67) or group (p = .31) were found, but a significant interaction effect emerged (p = .041). The intervention group showed an increase over time, whereas the control group showed a decline from pretest to posttest.

General child care quality. The intervention group and the control group did not differ significantly on ECERS-R scores at pretest (p = .33). The repeated measures ANOVA showed that there was a significant main effect for time (Pillai's F (1, 62) = 5.18, p =.03), indicating that general quality in both groups increased over time. We did not find a significant main effect for group (Pillai's F (1, 62) = 2.09, p = .15), nor an interaction effect (Pillai's F (1, 62) = 0.17, p = .90).

Caregiver attitude. After the intervention, caregivers who had participated in the intervention reported a more positive attitude towards caregiving and limit setting (M = 4.11, SD = .42, n = 20) than caregivers in the control group (M = 3.80, SD = .38, n = 19) (t (37) = 2.39, p = .02).

Evaluations. Caregivers in the intervention group reported significantly higher scores than caregivers in the control group on finding the training informative (t (34) = 3.60, p < .01), interesting (t (35) = 2.66, p = .01), and useful (t (30) = 2.20, p = .04) (see Table 5). Regarding the willingness to recommend the training to others, the intervention group scored significantly higher than the control group for recommending the training to starting caregivers (t (35) = 3.06, p < .01) and the standard vocational curriculum for prospective caregivers (t (33) = 2.58, p = .01).

Table 5
Caregiver evaluations of the training on 5- point Likert scales (1= "not at all / certainly not", 5 = "very much / certainly")

	Intervention group N = 20		*Control group N = 18*		
	M	*SD*	*M*	*SD*	*p*
The training was informative to me	3.89	.96	2.83	.79	<.01
The training was interesting to me	3.95	.62	3.28	.90	.01
The training was useful to me	3.88	.70	3.27	.88	.04
I would recommend training to starting caregivers	4.70	.73	3.64	1.32	<.01
I would recommend training to experienced caregivers	3.21	1.69	3.00	1.37	.69
I would recommend training to colleagues	3.63	1.46	3.31	1.30	.50
I would recommend training to the standard vocational curriculum for prospective caregivers	4.47	.77	3.50	1.41	.01

Discussion

The effects of the VIPP-CC in center-based child care on caregiver sensitive responsiveness, general child care quality and attitude towards sensitive caregiving and discipline were tested in a randomized controlled trial. With respect to caregiver sensitive responsiveness, we showed that the VIPP-CC was indeed effective for caregivers in center-based care: after the intervention, observed sensitivity increased, but only in the intervention group. Analyses regarding the intervention effect for caregiver sensitivity in unstructured and structured situations showed that the structured play situations accounted for the increase in sensitivity over time. The structured situations were more focused at in the VIPP training than unstructured settings. Therefore, an intervention effect may become most apparent in these settings, because learning experiences are probably activated more automatically in situations somewhat similar to the training. For caregivers, it may be easier to perceive and adequately interpret signals in a smaller group of children present during structured play than in a larger group in unstructured settings, because there are fewer children needing attention and signaling their needs, in line with findings from Ahnert et al. (2006). Still, the intervention effect may transfer to unstructured situations in a later stage, once the caregivers have consolidated their newly learned behavior in structured settings (Tziner, Haccoun, & Kadish, 1991). Our findings are promising, because even in our sample of qualified staff with quite some years of experience and relatively high scores for sensitivity at baseline, the intervention was effective.

Considering caregiver attitude towards sensitive caregiving and discipline, an intervention effect was found, which is in accordance with previous findings in home-based care (Groeneveld et al., 2011). After the VIPP-CC training, caregivers showed a more positive attitude towards caregiving and limit setting than the control group. We speculate that the intervention led to a change in caregiving attitude first, which in turn led to observable changes in behavior. Because of the change in underlying attitude, a generalization of the improved sensitive responsiveness to other situations than structured play tasks might be expected. These findings are in line with those reported by Susman-Stillman et al. (2013), who showed that for center-based caregivers, more positive caregiver attitudes and beliefs are related to higher quality caregiving practices.

With regard to general quality, moderate general quality at both time points in both groups was observed, representative for Dutch child care centers (Fukkink et al., 2013). In contrast with results of the VIPP-CC in home-based care (Groeneveld et al., 2011), no intervention effect was found. One reason may be that structural aspects that contribute to general quality (e.g. space and furniture, the presence of adequate play material, safety and hygiene practices) were not the focus of our intervention. More importantly, these aspects are not easily influenced by caregivers, but rather by center managers and through authority regulations. The situation is different for caregivers in home-based care, where individual caregivers are directly responsible for the physical environment in their own homes. This may explain the different results on general quality in our study compared with the Groeneveld et al. (2011) study. It was remarkable that for general quality, both intervention and control group showed a small, yet significant increase over time. A possible explanation is that during posttest, mean group sizes were slightly smaller, so that adequate supervision may have been easier to accomplish. Another explanation could be that participating in research observations may have led to caregiver's increased awareness of their classrooms. As a consequence, caregivers may have changed certain classroom aspects such as child decorations, the offering of specific play materials and hygiene activities from pretest to posttest. We should stress the fact that the intervention effect that we found for sensitive responsiveness was apparent only in structured play situations, whereas general quality is reflected by both unstructured and structured situations.

Limitations

The sample size was modest, so that subgroup analyses for more detailed investigation of effects were not possible. Although randomization was carried out carefully, in relatively small samples group differences may occur by chance. This may explain why at baseline, there was a difference between the control group and the intervention group regarding sensitive responsiveness. To control for these initial differences, repeated measures ANOVAs were conducted. Internal validity was further ensured by careful

procedures for recruitment, the pretest-posttest design, the use of reliable instruments, and blinding of the coders to avoid experimenter bias. With respect to external validity, our sample was representative for the population of child care centers in urban areas of the Netherlands, comparable to other study samples in the field. A downside of our multi-method approach (live observations, videotaping, questionnaires) may have been that it caused some non-response in posttest questionnaires about caregiving attitudes and caregiver evaluations.

Conclusion

We have shown that the short-term, attachment-based intervention VIPP-CC is effective in improving professional caregivers' behavior in center-based child care. Caregiver sensitivity to the group, which is an important predictor of a secure child-caregiver attachment, can thus be improved. With this study, the effectiveness of the VIPP approach has expanded from family settings (including home-based child care) to child care environments with larger groups of children. Future studies might focus on adapting the intervention program even further for group settings in which quality of care is highly disadvantageous for young children, for instance in institutionalized care and orphanages.

Appendix 1:
Overview of the VIPP-CC sessions

Visit	Video-taping episodes	Themes during intervention
0	• Building blocks together • Clearing up the toys • Turn taking	*Only filming, no intervention*
1	• Children playing by themselves • Playing together • Clearing up the toys • Storybook reading	Sensitivity: Exploration versus contact seeking Discipline: Inductive discipline and distraction
2	• Lunch or snack	Sensitivity: Speaking for the children Discipline: Positive reinforcement
3	• Caregiver responds only to invitation from children • Singing songs together • Building a tower together	Sensitivity: Sensitivity chain Discipline: Sensitive time-out
4	• Playing together with dolls • Do-not-touch-task • Reading a storybook	Sensitivity: Sharing emotions Discipline: Empathy for the children
5	• Building together • Clearing up the toys • Playing together	*All of the above*
6	-	*All of the above*

CHAPTER 4

Noise in center-based child care: Associations with quality of care and child emotional wellbeing

Claudia D. Werner, Mariëlle Linting, Harriet J. Vermeer, Marinus H. Van IJzendoorn

Manuscript submitted for publication

Author note:
We are grateful to Joost van Ginkel (Centre for Child and Family Studies, Leiden University) for his assistance in the multilevel multiple imputation procedure.

Abstract

Many children attending center-based child care are exposed to high noise levels on a daily basis. In the present study, the associations between noise levels, noise variability, caregiving quality and child wellbeing were investigated in child care centers (*N* = 64) involving children of 0 to 4 years (*N* = 245; *M* = 34.50 months). We examined whether thresholds for noise could be found for center child care; that is, whether minimum and maximum levels of noise and noise variability were required for optimal child wellbeing. Nonlinear regression analysis confirmed the threshold hypothesis: Optimal child emotional wellbeing was observed for noise levels over approximately 60dbA and below 65 dbA, and for noise variability over approximately 6.69 dbA and below 7.44 dbA. Linear multilevel regression analysis showed that more hours in care, higher child age and higher general child care quality were related to higher levels of wellbeing. A closer examination of the sources of noise showed that indoor activities were associated with lower noise levels and outdoor activities were related to higher noise levels. We conclude that noise, a major aspect of environmental chaos, has adverse outcomes on child wellbeing in center child care. The regulation of noise levels in child care centers is needed to provide optimal care and child wellbeing.
Keywords: Center child care, noise, child social-emotional wellbeing, quality of care, environmental chaos, nonlinear analysis

Introduction

Noise levels in child care centers can be overwhelming. In fact, these levels have been reported to range from 45 decibel, comparable to levels of adult normal conversation, up to more than 90 decibel, which is comparable to high way noise and, when sustained, can cause hearing damage (Manlove, Frank, & Vernon-Feagans, 2001). Center-based child care for children under five is an important rearing and care environment for millions of children in Western countries (Organization for Economic Co-operation and Development, 2013), implying that a large proportion of children may be daily exposed to high noise levels. In this study, we investigate how children's emotional wellbeing is associated with noise levels in center-based child care. A recent study by Linting, Groeneveld, Vermeer, and Van IJzendoorn (2013) showed that in home-based child care there is a threshold for noise beyond which lower child wellbeing can be observed. At regular child care centers, noise levels are generally higher than in home-based child care settings (Groeneveld, Vermeer, Linting, & Van IJzendoorn, 2010). In this study we examine noise levels in center child care and their associations with quality of care and children's emotional wellbeing.

Noise is a central component of the environmental chaos theory that originates from family research and states that environmental chaos is detrimental for child development (Evans & Wachs, 2010). Noise can be defined according to intensity (i.e. low or high average noise levels), variability (i.e. the differences in peaks and lows), and duration: Occasional versus chronic noise (Enmarker & Boman, 2004; Kjellberg, Landström, Tesarz, Söderberg, & Åkerlund, 1996). Linting et al. (2013) showed that not only noise intensity but also noise variability beyond certain levels is a predictor for lower levels of child wellbeing. In addition, the sources of noise can be categorized as non-social (e.g. noise from traffic, roads, aircraft, and transport) or social (e.g. chatter, classroom sounds, music, and activities). In this study, both social and non-social sources of noise will be considered. Environmental chaos in the home setting is reflected in households with high levels of crowding, noise, and the lack of family routines, order, and regular planning of activities (Matheny, Wachs, Ludwig, & Philips, 1995). High levels of chaos and crowding in the family setting have been associated with unfavorable outcomes, for instance with less parental talk, more negative parent-child interactions, more child social withdrawal and aggression, more child helplessness, less optimal child cognitive development, and more stress for adults and children (Evans, 2006; Evans & Wachs, 2010).

Noise can affect various ecological environments of a person (Evans & Wachs, 2010). Investigations relevant to child development focused on the effects of noise on children in the home environment (e.g. Babisch, Schulz, Seiwert, & Conrad 2012; Evans, 2006), teachers and adolescents in schools (e.g. Enmarker & Boman, 2004), and children in primary schools (for reviews see Evans, 2006; Shield & Dockrell, 2003). Fewer studies

addressed the effects of noise in child care settings, targeting toddlers in child care (Corapci, 2010; Groeneveld et al., 2010; Hambrick-Dixon, 1986; Linting et al., 2013; Manlove et al., 2001; McAllister, Granqvist, Sjölander, & Sundberg, 2009) or child care professionals (Lindstrom, Persson Waye, Södersten, McAllister, & Ternström, 2011; Sala et al., 2002).

The effects of noise encompass three major domains: health, cognition, and psychological wellbeing. In the health domain rather consistent adverse outcomes have been shown for noise levels and physiological stress indicated by higher blood pressure and higher levels of stress hormones in children and adults (for a review see Evans, 2006). Studies in the domain of cognition showed adverse outcomes of noise on children's memory and attention (Evans, 2006; Klatte, Bergström, & Lachmann, 2013; Shield & Dockrell, 2003), psychomotor performance, reading skills, and central information processing (Klatte et al., 2013; Evans, 2006). Studies considering the third domain, psychological wellbeing, showed negative associations between noise and parental communication and responsiveness (Evans, 2006), emotional wellbeing (Stansfeld et al., 2009), and positive associations with annoyance (Babisch et al., 2012; Enmarker & Boman, 2004; Haines et al., 2001; Maxwell, 2010), lack of motivation, helplessness, lack of patience, and aggression (Evans, 2006). It should be stressed that findings were mixed and that child participants in the studies varied in age. More importantly, most studies regarded non-social noise, which may have different effects on children than social noise. It can be argued that social noise may have lower average levels and lower maximum peaks than noise from aircrafts or subways passing. The former may be more controllable. Although unpredictable and uncontrollable noise are generally perceived as more annoying, Enmarker and Boman (2004) found that social noise from chatter in classrooms was perceived as more annoying than non-social noise.

Some researchers showed that negative associations between developmental outcomes and noise were stronger for older children, because older children may be more aware of the noise and therefore experience it as more intrusive or distracting than younger children (see Evans, 2006; Eysel-Gosepath, Daut, Pinger, Lehmacher, & Erren, 2012). Younger children, on the other hand, may be more vulnerable to negative environmental influences and may therefore more negatively affected by higher noise levels (Evans, 2006; Eysel-Gosepath et al., 2012). Apart from age, individual differences in sensitivity to noise may explain why some children may be more affected than others (Maxwell & Evans, 2000; Enmarker & Boman, 2004). To our knowledge only three studies have been conducted to noise in relation to child emotional wellbeing in the child care setting, two with a correlational design (Groeneveld et al., 2010; Linting et al., 2013) and one case study (Kishimoto, 2012).

Aim of the study. The aim of this study is to examine associations between noise and child emotional wellbeing in center child care. In our study average noise levels and noise variability are main predictors for the outcome variable child emotional wellbeing. Group size and child-caregiver ratio are taken into account as additional indicators of environmental chaos, in particular of crowding. The associations between noise and quality of care, caregiver working experience in child care, caregiver working hours per week, child gender, age, and hours in care are explored. We specifically investigate whether a threshold of noise can be found beyond which lower child wellbeing can be observed in center-based child care. We expect that the threshold theory of Linting et al. (2013) that was applied in home-based child care also applies to center child care, implying that beyond certain noise levels noise and noise variability are associated with lower levels of child emotional wellbeing; over and above associations with child, caregiver, and child care characteristics.

Method

Recruitment

This study is part of a larger investigation into the effectiveness of a video-feedback intervention aimed at professional caregivers in center-based child care. Here, we only present the pretest data. Participants in this study were children attending center-based child care. For recruitment we targeted child care centers in an urban area in the western part of The Netherlands. Letters of invitation were sent to180 child care organizations. Initially, 91 centers from 35 organizations agreed to participate. However, 22 centers dropped out before assessment because of a lack of interest from the managers (n = 3), unwillingness of professional caregivers to participate (n = 6), or lack of parental consent (n = 13). Two centers withdrew from the study directly after the assessment; the managers did not allow us to use the observational data in our study.

To avoid selection bias, one group per center was randomly selected for participation. Toddler groups (for children 2- 4 years) and mixed-age groups (for children 0 – 4 years) could be included. Furthermore, groups within centers were eligible for participation if parents of at least three children attending the group on the same day of the week provided written consent. Three centers that agreed to participate only had eligible groups of young infants (0 to 18 months), so they were excluded from the study. Professional caregivers from eligible groups had to be working for a minimum of two days (16 hours) per week in a fixed group and be available and willing to participate during the study period of approximately six months. From each group, one caregiver who met these inclusion criteria was randomly selected. Selection of target children per group was based on parental consent and their attendance on the day scheduled for assessment. If more than four children with consent were present at the assessment

day, selection was done randomly. There was a minimum of three and a maximum of four target children per group.

Participants

In total, 245 children and 64 caregivers from 64 centers were included for analysis. Approximately 49% of the children were boys (*n* = 121). The children had a mean age of 34.50 months (*SD* = 7.78) at the time of the assessment and they attended toddler groups (67%) and mixed age groups (33%). For more descriptive statistics, see Table 1. The majority of caregivers (*n* = 43) had a degree in vocational training. A minority (*n* = 8) had finished a higher educational degree on a bachelor's or master's level. Few caregivers (*n* = 3) had low educational levels, i.e. only primary school or high school degree. Educational level was not reported by 10 caregivers (16%). The majority of caregivers (65%) were born in the Netherlands. The other caregivers were born in Surinam (*n* = 4), the Dutch Antilles (*n* = 2), Turkey (*n* = 1), Cape Verde (*n* = 1); or this was not reported (*n* = 14).

Table 1
Descriptive statistics (prior to winsorizing and imputation of missing data)

		N	*M*	*SD*	*Min*	*Max*
Child level	Child age (months)	227	34.50	7.78	9.00	47.00
	Hours in care per week	180	27.91	9.55	8.00	55.00
	Child wellbeing	242	4.50	.48	3.00	6.33
Group level	Caregiver age (yrs)	53	32.09	8.63	22.00	52.00
	Caregiver work hours per week	54	29.42	5.74	16.00	40.00
	Caregivers years of experience	55	7.70	6.95	.60	35.00
	Quality of care (ECERS-R)	64	4.00	.78	1.89	5.39
	Caregiver sensitivity	64	4.84	.91	2.33	6.33
	Noise level average dB(A)	61	61.69	3.30	54.95	70.83
	Noise variability	61	7.39	.81	6.10	9.68
	Group size	63	12.14	3.11	4.00	18.00
	Child-caregiver ratio	63	5.41	1.76	2.00	13.00

Procedure

The child care centers were visited for assessment once from 9:00 AM until 1:00 PM. Afterwards, questionnaires were sent to the professional caregivers and the center managers. During the visit, a shortened version of the Early Childhood Environment Rating Scale-Revised (ECERS-R; Harms, Clifford & Cryer, 1998; Vermeer, 2012) was administered. The professional caregiver and four target children were recorded with a

digital video camera at three predetermined time points (at 9.30 AM, 10.30 AM and 11.30 AM) during regular activities. Each target child was recorded with a camera during three 5-minutes fragments, and the caregiver was recorded for three 10-minutes fragments. The target children and the caregiver were recorded successively, so that fragments did not overlap. During these recordings noise levels were registered for 30 minutes. All recorded fragments were rated afterwards on caregiver sensitivity and child emotional wellbeing. To obtain independence in ratings, researchers who administered the ECERS-R did not code video material from that particular setting. In addition, caregiver sensitivity and child emotional wellbeing were coded by different researchers.

Measures

Child emotional wellbeing. Child emotional wellbeing was measured with the Wellbeing Scale, developed and validated by the Dutch Consortium for Child care Research (NCKO; De Kruif et al., 2007). This scale contains several indicators of the child's wellbeing, such as pleasure, self-confidence, and relaxation. Scores were based on three video fragments of five minutes each of the target children at the child care center. Every two and a half minutes a score was registered, resulting in six intervals for each child. Wellbeing scores are presented on a seven-point scale, ranging from (1) very low wellbeing (signals of discomfort are clearly present, e.g. crying, screaming) to (7) very high wellbeing (signals of comfort are clearly present, e.g. enjoyment, smiling). Scores were aggregated across the time periods. Eight observers were trained to reliably assess child wellbeing. All observers met the criteria of reliability: mean intra-class correlation (two-way mixed, absolute agreement) was .79 (ranging from .71 to .80). Internal consistency (Cronbach's alpha) across the six intervals was .70.

Noise levels and variability. Data Logger Sound Level Meters (type CEN) were used to measure noise levels in decibels at the child care centers. We used dB (A) weighting which simulates the response of the human ear. Noise recording was conducted in parallel to the three predetermined videotaped sessions of 30 minutes; the sound level meter was placed in the room or playground where the caregiver and children were at the time of the observations. Noise levels were automatically recorded every second, and recordings were averaged across the three observation sessions to obtain a final score. Noise variability was computed as the average standard deviation of noise levels across the three observation sessions (Linting et al., 2013).

Sources of noise. Sources of noise were coded for all child care groups, using three 10-minutes videotaped observations of the caregiver. By using the caregiver observations rather than child observations, sources of noise at the group level may be more accurately identified because the scope of the observations included the whole group setting, rather

than just one target child. The observation scheme developed by Linting et al. (2013) for home-based child care formed the basis for our coding with some minor adaptations for center child care. For instance, it was highly unlikely that pets were sources of noise in center child care, because of regulations. Five main categories for sources of noise were distinguished: (1) outdoor noise, (2) background noise, (3) children's noise, (4) adults' noise, and (5) noise from the handling of toys. The coding form and instructions for coding are presented in Appendix A. Moreover, taking into account centers with multiple groups, we further defined 'background noise' as sounds that originated from inside the target group (e.g. 'background media' for music from a CD player, and telephone ringing) versus sounds from outside the target group ('background: human conversation' and 'background: other'), for instance when caregivers and parents of children from other groups were talking in the corridor, or children of other groups were playing in an area nearby the target group. We further defined sounds originating from children and adults as 'positive vocalizations' (talking, laughing, singing), 'negative vocalizations' (for children: crying, screaming, fighting; for caregivers: scolding, shouting), 'neutral sounds' (clapping, coughing, drinking), and sounds of 'moving objects' (plates, chairs, toys, or kitchen tools). For children's sounds, we added a separate category of 'children moving around', because the running, jumping and climbing could be a separate source of noise. Finally, there was a category for sounds originating from the handling of different types of toys. For each of the five main categories there was also an option for coders to add an item beyond the other definitions ('other'). In addition, we coded whether the observation was inside or outside the building for the majority of the time. For each variable, coders reported on a four-point scale whether the described source of noise was (0) absent, (1) occasionally present (less than 25% of the time), (2) often present (between 25 and 75% of the time), or (3) constantly present (more than 75% of the time) during the 10-minutes observation. In total, 20 variables for sources of noise were rated for three sessions per child care center. Four coders were trained to be reliable on the coding instrument by using a training set of 14 observations from five centers. On the basis of absolute two-way agreement with the first author, intercoder reliability (ICR) for the training set was good (M = .81, range .78 to .85). For the remaining DVDs inter coder reliability was established on the basis of double coding of 10% of the DVDs (6 sessions for each coder) by the first author; intercoder reliability for these 18 sessions was good (M = .78, range .76 to .82).

Caregiver sensitivity. Caregiver sensitivity was coded for three video recorded fragments during regular child care activities. The scale to code caregiver sensitivity was developed and validated by the Netherlands Consortium for Child care Research (NCKO, 2007; De Kruif et al., 2007; Helmerhorst, Riksen-Walraven, Vermeer, Fukkink, & Tavecchio, 2014). This is a group rating scale based on scales developed to measure sensitivity in

the parent-child context (Ainsworth, Bell, & Stayton, 1974; Erickson, Sroufe, & Egeland, 1985). Caregiver sensitivity refers to the degree to which the caregiver provides adequate and sufficient emotional support to all children in her care who need it, during stressful and non-stressful situations. It also refers to the level to which a caregiver is able to adequately divide attention among the children, showing interest in the children's activities and acknowledging their needs, emotions and competences. Sensitivity ratings are presented on a seven-point scale, ranging from (1) very low to (7) very high. A caregiver scoring high on the scale is very much involved with the children, and responds promptly and adequately to the signals of all children in her care, by taking the children's perspective. A caregiver scoring low on this scale may show either emotional distance or indifference towards the children, or she may be uninvolved with the children, because of administrative or cleaning tasks in the group and thus missing the signals of the children. Two independent observers were trained to be reliable coders, using the NCKO reliability set (Helmerhorst et al., 2014). Intra class correlation for both coders was .75 (absolute agreement). Approximately 75% of the video recordings were coded by the two observers. The remaining 25% were coded by the third author of this paper, who was involved in developing the scale.

General child care quality. To asses general child care quality a shortened version of the ECERS-R (Vermeer, 2012; Harms, Clifford & Cryer, 1998) was administered in all participating groups. The ECERS-R is a revised version of the original instrument and has been used extensively across countries. It contains seven subscales with 43 items and has demonstrated its validity and reliability (Perlman, Zellman, & Le, 2004). Training and administration of the full ECERS-R are quite time consuming; therefore in the current study a shortened version with 18 items was used. Psychometric properties of the shortened version of the ECERS-R are satisfactory (Vermeer, 2012). In the reported study, internal consistency of this instrument was adequate, with Cronbach's alpha of .79. Seven observers were trained by the third author of this paper to be reliable on the shortened ECERS-R. The training encompassed reviewing and discussing the items and field observations. Inter rater reliability was established to a criterion of 80% agreement for three consecutive observations for all observers. The mean percentage of agreement for these three observations was 90% (range 87% to 92%).

Observed group size and caregiver-child ratio. The number of children and caregivers present during the three observation sessions was registered by the observers. Group size refers to the total number of people in the room (both children and caregivers). The child-caregiver ratio was calculated as the number of children divided by the number of qualified caregivers in the room.

Demographic information. Center managers provided background information on the child care centers through a questionnaire. In a background questionnaire for the caregivers, information was gathered on their age, level of education, birth country, years of working experience in child care, and working hours per week. They were also asked to report the date of birth and number of hours of child care attendance per week for each of the target children in their care.

Data analysis

Data inspection. Data were collected on the group level (*N* = 64) and the child level (*N* = 245). The dataset was inspected for normality of distributions and outliers. On the group level, the scores for caregiver years of experience and child-caregiver ratio were not normally distributed. For caregiver years of experience, one outlier accounted mainly for skewness of the distribution. We applied a procedure similar to winsorizing: We replaced the influential outlier with a value closer to the distribution, keeping the rank order of the observed scores. After dealing with the outliers, all predictors and outcome variables were normally distributed.

Percentages of missing data ranged from 0% to 17% for variables on the group level. Caregiver questionnaires were not returned by ten caregivers (16%), leading to missing data on caregiver age, years of experience and working hours per week for these subjects. Regarding data on the child level, percentages of missing data ranged between 1% for observed wellbeing and 27% for hours in care. To obtain a complete dataset prior to the analyses, multiple imputation was performed (ten times) (Van Buuren, 2010; Goldstein & Woodhouse, 1996) including available variables in the data set on the child level and the group level, using predictive mean matching to impute missing data (Little, 1988; Rubin, 1986) and taking the nested structure of the data into account. Finally, the pooled imputed dataset (*N*= 245) was used for subsequent multilevel analysis. Before applying multilevel regression analysis, all predictor variables were centered by using the average score for each imputed dataset.

Multilevel analyses. The sample consisted of children (*N* = 245) who were nested within child care groups (*N*= 64). This dependency among children was taken into account by performing multilevel analysis, or hierarchical linear modeling (HLM), using Mixed models in SPSS 21.0. A two-level random intercept model was used to predict wellbeing on the individual child level (level 1) from child characteristics. Differences between the groups on child wellbeing were investigated with variables on the group level (level 2). Full maximum likelihood was used for estimating the model parameters. To investigate model fit we used – 2 log likelihood (-2LL) values. However, imputed datasets do not provide pooled values of -2LL. Therefore, we used the imputed dataset with the highest – 2LL in the first model (Model 0) to evaluate model fitness (see Table 3).

Nonlinear data analysis. We used categorical regression analysis (CATREG; Meulman, Heiser, & SPSS, 2012) in SPSS 21.0 to examine possible nonlinear relations between child wellbeing and noise characteristics, similar to Linting et al. (2013). Nonlinear regression is an alternative to linear regression, developed for performing ordinary least squares regression on possibly nonlinearly related variables (Van der Kooij, 2007). The method is quite similar to using linear regression with transformed data (e.g. log transformation). An advantage of CATREG is that an *optimal* transformation is performed, that is, a transformation that best reflects the relation between the transformed predictor and the response, given particular restrictions imposed by the researcher. We applied a spline nominal analysis level to our predictor variables average noise level and noise variability, which means that the transformation of these variables follows a smooth curve that may go up and down with the original order of the values. We used the SPSS default settings, estimating a quadratic curve in three data intervals. To ensure interpretability of the results, the response variable wellbeing was analyzed numerically (without transformation). The reported *p*-values from the CATREG models are based on 50 bootstraps.

Currently, nonlinear regression in SPSS does not allow for hierarchically structured (multilevel) data, nor does it allow pooling of results across imputed data sets. Therefore, analyses could only be performed on the group level (N = 64) and for one (randomly selected) imputation. We aggregated wellbeing scores for the selected imputed dataset by computing the mean across children within the same child care group. The nonlinear regression analysis was repeated for all other imputed datasets and similar results were yielded. We therefore present the results for one randomly selected imputed dataset only.

Nonlinear principal component analysis. We used nonlinear (categorical) principal component analysis (CATPCA; Linting & Van der Kooij, 2012) in SPSS 21.0 to examine multivariate relations between the ratings for sources of noise. In contrast to linear principal component analysis, nonlinear PCA allows researchers to handle nonlinear relationships between variables and to specify analysis levels separately for each variable, so that these can be in accordance with the measurement level of the variables (Linting, Meulman, Groenen, & Van der Kooij, 2007a, 2007b). Sources of noise were measured at an ordinal measurement level, as these variables were scored on four-point Likert scales. For each of the 20 sources of noise variables, we aggregated scores over the three 10-minutes observations. The aggregated scores were then rounded off to the original four-point scale, so that child care centers' average scores were categorized as 0, 1, 2, or 3 for each variable. Next, we examined whether numerical, (spline) ordinal or (spline) nominal scaling were the best fitting transformations for each of the variables. We finally chose to use ordinal transformation, taking into account the quantification

plots and the original four-point ordinal rating scale. In Appendix B the procedures for nonlinear PCA are described in more detail.

In order to relate the solution of the nonlinear PCA to the average noise levels and average noise variability of the centers, we created three categories for these variables: centers fell in the category of "low noise" (< 60.00 dbA, *n* = 22), "middle noise" (60.00-65.00 dbA, *n* = 26) or "high noise" (> 65.00 dbA, *n* =16), and for variability centers could be of "low variability" (< 7.00 dbA, *n* = 25), "middle variability" (7.00-8.00 dbA, *n* = 28) or "high variability" (> 8.00 dbA, *n* = 11). These categorized variables of noise were plotted as supplementary variables in the solution defined by the sources of noise.

Results

Descriptive statistics

Means, standard deviations and the range of scores for variables on the child level and the group level are provided in Table 1. Bivariate Pearson correlations between background characteristics and outcome variables are presented in Table 2.

Higher levels of wellbeing were associated with more hours in care and with higher age of the child. Moreover, higher child wellbeing was associated with higher levels of caregiver sensitivity, caregiver experience and quality of child care. Higher quality of care was associated with higher caregiver sensitivity levels. Larger group sizes and more children per caregiver were associated with lower general child care quality, but with higher caregiver sensitivity. Higher noise levels were related to higher caregiver sensitivity levels. In contrast, higher noise variability was associated with lower quality of care and lower caregiver sensitivity. Average noise levels and noise variability were negatively associated, indicating that groups with higher average noise levels showed less fluctuation in noise levels.

Multilevel analyses

The intra class correlation for wellbeing (calculated on a model with just a random intercept and no predictors) was 0.22, indicating that multilevel analyses are indeed preferable beyond regular linear regression. As a comparison, in educational research, intraclass correlations of around 0.10 are quite common (Hox, 2002; Twisk, 2006). We hierarchically added terms to the model, first on level 1 and then on level 2. Results of the Models 1 to 6 are presented in Table 3. The multilevel analyses showed that more hours in care, higher child age and higher general quality in the group were related to higher levels of wellbeing. The other predictors in the model did not significantly contribute to the prediction of wellbeing.

Table 2

Pearson correlations between characteristics of the children, caregivers and child care groups before and after imputation[1] (N =245)

	1	2	3	4	5	6	7	8	9	10	11	12
1. Child gender[a]	-	.10	.00	.12	-.03	-.04	.09	.13	-.07	-.03	.02	-.03
2. Child age	.10	-	.03	.28	.10	.11	-.07	.04	.11	.18	.11	-.04
3. Child hours in care	- .04	.04	-	.16	.03	.10	-.12	.08	-.13	.03	.07	.14
4. Child wellbeing	.12	.29	.20	-	.18	-.03	.18	.17	-.06	.06	.06	-.05
5. Caregiver experience (yrs)	- .04	.10	.05	.21	-	- .12	.12	.02	-.08	.15	-.01	- .05
6. Caregiver working hours	.03	.11	.11	-.03	- .14	-	-.25	-.21	-.02	-.16	.12	-.04
7. Quality of child care	.09	-.08	-.16	.14	.13	- .28	-	.44	- .15	- .07	- .01	- .14
8. Caregiver sensitivity	.13	.02	.11	.17	.00	- .22	.44	-	.08	.30	.16	- .19
9. Group size	- .07	.13	-.18	- .04	- .14	-.10	-.15	.08	-	.35	.28	-.25
10. Ratio[b]	- .03	.19	.03	.07	.14	-.16	-.07	.30	.36	-	.22	-.12
11. Noise level (dbA)	.01	.10	.12	.07	- .01	.10	.00	.14	.27	. 22	-	-.28
12. Noise variability	- .02	-.03	.16	.05	- .04	-.03	-.15	- .18	- .24	- .14	- .28	-

Note. [1] bivariate correlations before imputation are presented under the diagonal, correlations after imputation above the diagonal; [a] boy = 0, girl =1; [b] number of children per caregiver

Table 3
Results of two-level hierarchical linear regression analysis to predict child emotional wellbeing (N = 245)

Parameter	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
				Fixed effects			
	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)
Intercept	4.50 (.04)	4.46 (.05)	4.46 (.05)	4.50 (.04)	4.50 (.03)	4.50 (.04)	4.50 (.03)
Level 1 (child)							
Gender		.08 (.06)	.08 (.06)				
Age		.01* (.00)	.01* (.00)	.01* (.00)	.02* (.00)	.01* (.00)	.02 * (.00)
Hours in care			.01* (.00)	.01* (.00)	.01* (.00)	.01 (.00)	.01* (.00)
Level 2 (group)							
Caregiver yrs of experience				.01 (.01)			
Caregiver working hours				-.00 (.01)			
Quality of care (ECERS-R)					.12* (.05)		.11* (.05)
Caregiver sensitivity					.03 .(04)		.03 (.05)
Group size						-.02 (.01)	-.01 (.01)
Caregiver-child ratio						.01 (.03)	.01 (.03)
Noise levels						.00 (.01)	.00 (.01)
Noise variability						-.04 (.05)	-.01 (.05)
				Random parameters			
Intercept (variance)	.05* (.02)	.04* (.02)	.04* (.02)	.04* (.02)	.03* (.01)	.04* (.02)	.03 (.01)
-2 Log likelihood[a]	332.12	316.22	300.05	300.81	294.19	299.50	292.82

Note. [a] We selected the imputed dataset with the highest value on this variable in Model 0. The relative changes in -2LL were comparable for all imputed datasets. * $p < .05$

Moderators. The interaction of child age with hours in care was not significant (β = -.00, *SE* = .00, *t* = -.32, *p* = .75). As a next step, we tested cross-level interactions. None of these interactions turned out to be significant (see Appendix C, Table C1).

Nonlinear analyses

To test nonlinear relations between noise characteristics and wellbeing, we used CATREG on the aggregated data. The child care groups (*N* = 64) had a total mean score for aggregated wellbeing of 4.50 (*SD* =. 32), ranging from 3.79 to 5.25. The linear model in CATREG (no transformation, numerical level of analysis for all variables) only including average noise and noise variability explained 3% of the variance in wellbeing (R^2 = .03, *p* = .43). The model with a spline ordinal transformation level for the noise variables did not show notable improvement (R^2 = .05, *p* = .24). The model allowing for nonmonotonic nonlinear relations, however, showed much improvement over the linear and ordinal models (R^2 = .24, *df* = 8, *F*= 2.12, *p* = .049) and wellbeing could be significantly predicted by average noise levels (β = .36, *p* = <.001) and noise variability (β = .33, *p* = <.001). Transformation plots of the nonlinear regression analyses are displayed for average noise levels (Figure 1) and noise variability (Figure 2).

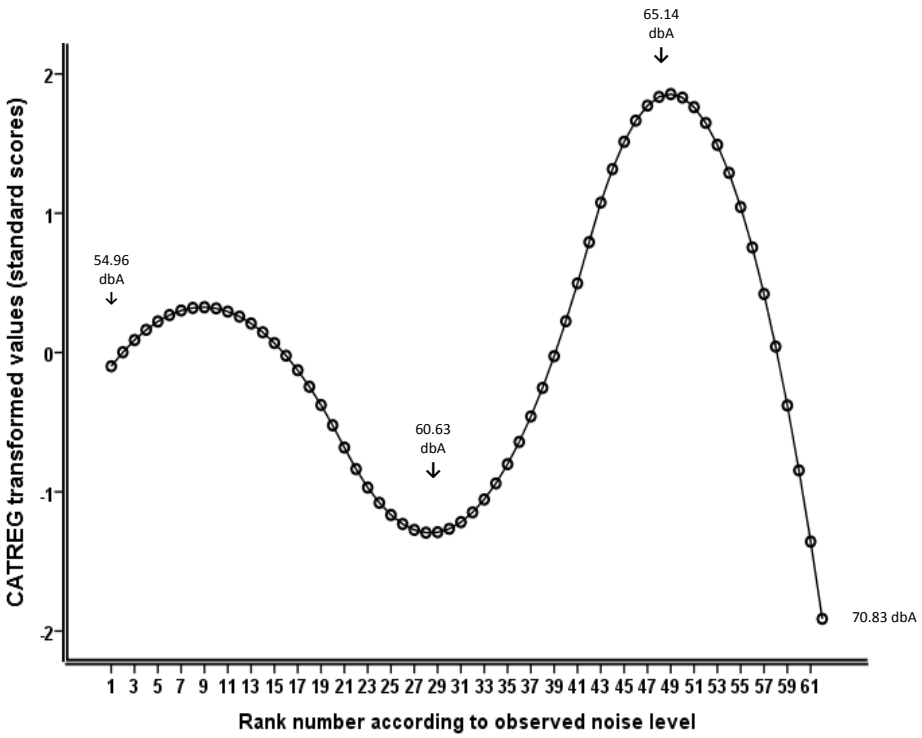


Figure 1. Plot of nonlinear regression analysis (spline nominal transformation) for child emotional wellbeing predicted by average noise levels. Note: β = .36, *p* = <.001

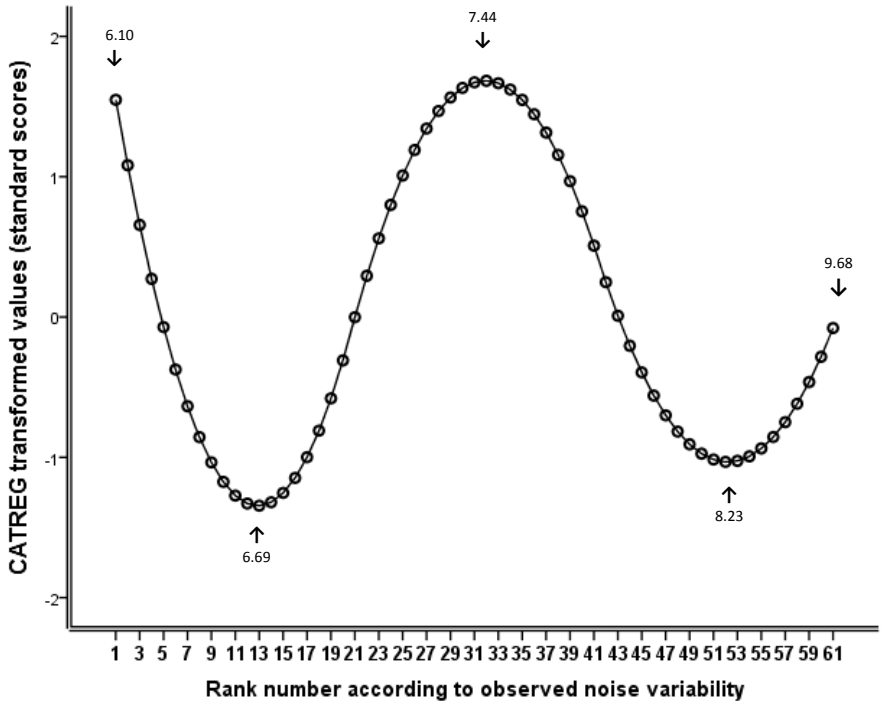


Figure 2. Plot of nonlinear regression analysis (spline nominal transformation) for child emotional wellbeing predicted by noise variability. Note: β = .33, *p* <. 001.

Average noise levels. The interpretation of the positive relation between transformed noise and wellbeing *(β* = .36*, p* < .001) becomes clear in Figure 1. The first part of the plot shows that as noise levels increased, child wellbeing levels (slightly) decreased. After a certain noise level (about 60.63 dbA) was reached, the association changed: Wellbeing increased with an increase of noise. Finally, beyond a noise level of about 65.14 dbA, an increase in noise level was related to a decrease in child wellbeing, again. There seemed to be a minimum amount of noise required to reach optimal wellbeing levels, but we also found a maximum. If noise levels fell below the first or above the latter threshold, child wellbeing decreased.

Noise variability. For noise variability (β = .33, *p* <.001) the pattern was comparable to the pattern for average noise levels (see Figure 2). Wellbeing decreased with increasing noise variability for the lower range of scores (approximately from 6.10 dbA to 6.69 dbA). Then, there was an increase of wellbeing up to a certain level of noise variability (around 7.44 dbA), after which there was a decrease in wellbeing with an increase of noise variability. In the last part of the figure (beyond 8.23 dbA), an increase of noise

variability seems to correspond with an increase of wellbeing, again. However, given the fact that only a relatively small number of observations (*n* = 9) are plotted in this part, this finding is less reliable.

Nonlinear model with covariates. In an additional analysis, we entered covariates child hours in care and child age as numerical variables to the model to see whether noise indicators would remain significant predictors (see Table 4). Child age proved to be a significant predictor *(β* = .44*, p* < .001), indicating that older children had higher levels of wellbeing. The nonlinear model with covariates added 16% explained variance in wellbeing (R^2 = .40, *p* < .01). Average noise levels (β = .39, p <.001) and noise variability remained significant predictors (β = .31*, p* < .001) in this model.

Table 4
Standardized regression coefficients from CATREG for outcome variable child emotional wellbeing

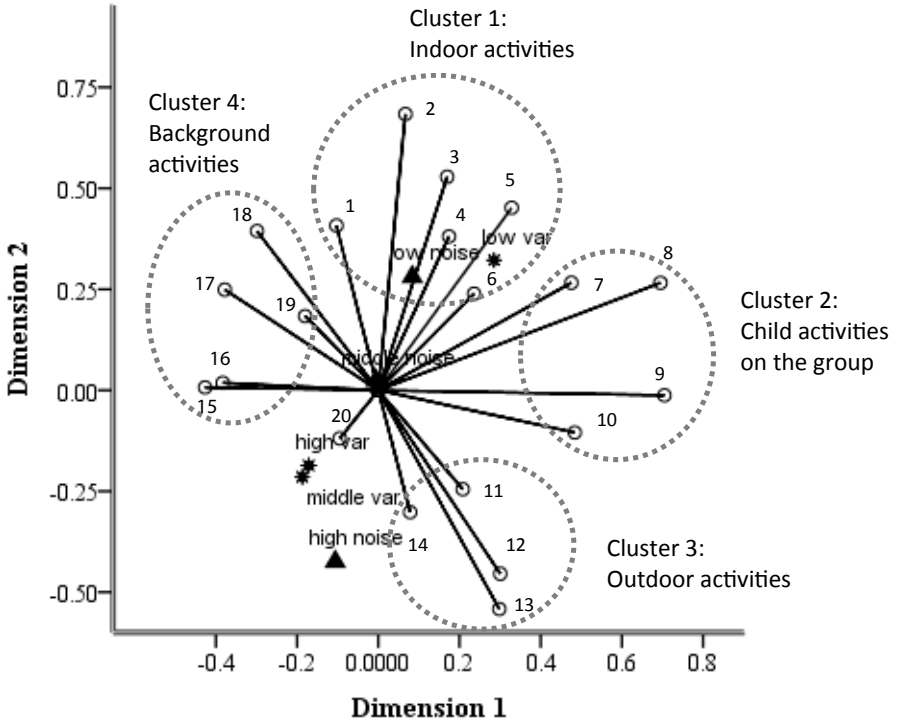
Predictor	Model A (R^2 = .24, *p* = .05)			Model B (R^2 = .40, *p* < .01)		
	β	*Bs SE*	*p*	β	*Bs SE*	*p*
Average noise levels[a]	.36	.10	<.001	.39	.10	<.001
Noise variability[a]	.33	.10	<.001	.31	.09	<.001
Child age[b]				.44	.10	<.001
Child hours in care[b]				.14	.13	.27

Note. Model A only includes noise indicators as predictors; Model B includes also covariates on the child level; *Bs SE*: bootstrap estimate of standard error; [a] Analyzed at a spline nominal level; [b] Analyzed at a numeric level

Nonlinear principal component analysis for sources of noise. We first explored a two-dimensional solution in CATPCA. There were no missing data and no outliers in the object plot (*N* = 64). The variance accounted for (VAF) by the two-dimensional solution was 25%, with eigenvalues of 2.51 (VAF= 12%) for the first dimension and 2.39 (VAF = 12%) for the second dimension. As the scree plot of the saved transformed variables in a linear PCA showed an 'elbow curve' after the second dimension and as the third dimension did not contribute much in terms of interpretation (Linting et al., 2007a), we decided to stay with the two-dimensional solution. In Figure 3 the results of the two-dimensional solution of the nonlinear PCA are shown.

The length of a variable vector represents how much of the variance in that variable is accounted for by the two dimensions. The categories for noise levels and noise variability are plotted as supplementary variables (i.e. they do not influence the solution), displayed by small triangles and small asterisks, respectively. In Appendix B, Table B1 component loadings of the variables are shown; the numbers of the variables in Figure 3 correspond to the numbers in Table B1.

Two dimensions were identified: Dimension 1 represents child activities on the group versus background activities (i.e. cluster 2 variables versus cluster 4 variables, respectively) and Dimension 2 represents indoor versus outdoor activities (i.e. cluster 1 variables versus cluster 3 variables, respectively). The clusters of variables are indicated by the dotted circles in Figure 3. Noise levels and noise variability seemed to be related most clearly to the second dimension: Sources of noise from the indoor activities in cluster 1 were related to low noise and low noise variability, whereas sources of noise from the outdoor activities in cluster 3 were related to high noise levels. The sources of noise on both ends of the first dimension in cluster 2 (child activities on the group) and cluster 4 (background activities) appeared to be less strongly associated with noise levels and noise variability.



Note: noise= average noise levels; var = noise variability

Figure 3. Plot of two-dimensional CATPCA for 20 sources of noise variables (*N* = 64). The variables of average noise levels and noise variability (categorized as 'low', 'middle', and 'high') are plotted in the solution as supplementary variables. Note: See Table B1 for correspondence of numbers with variables.

Discussion

Environmental chaos in center-based child care was the focus of this study. We showed that noise, a major aspect of environmental chaos, had adverse outcomes on child wellbeing. The association between child emotional wellbeing and noise was nonlinear: Extremely high as well as extremely low noise intensity and noise variability levels were associated with lower levels of child emotional wellbeing, over and above associations with child, caregiver, and child care center characteristics. These thresholds for noise indicate that certain maximal but also certain minimal levels of noise are required for optimal child wellbeing. Our findings are consistent with the results of Linting et al. (2013), who found evidence for a threshold at the upper end of the noise spectrum in home-based care.

As expected, average noise levels in center child care found in this study (*M* = 61.69 dbA; range 54.95 dbA to 70.83 dbA) were generally higher than average noise levels in home-based child care (*M* = 56.54 dbA, range 48.45 to 64.56 dbA; Groeneveld et al., 2010; Linting et al., 2013). This difference in noise intensity is substantial and cause for concern, considering that an increase in 10 dbA represents noise that is perceived twice as loud. Moreover, it should be noted that the use of average noise levels, aggregated over three observation sessions, fails to reveal that sound level meters in our sample registered peaks exceeding 95 dbA in some centers, albeit only for seconds. Noise levels above 90 dbA are comparable to high way noise and can cause hearing damage when these levels are sustained (Manlove et al., 2001). Moreover, groups with higher average noise levels in our study showed less variance, which implies that children in these groups were continuously exposed to intense noise levels. This finding is similar to the situation in home-based child care (Linting et al., 2013). Noise variability levels in our study (*M* = 7.39 dbA, range 6.10 to 9.68) were also comparable to levels in home-based care (*M* = 7.85 dbA, range 4.44 dbA to 10.66 dbA; Linting et al., 2013).

Apart from noise, group size and child-caregiver ratio (indicators of crowding) did not predict child wellbeing, which is in accordance with results from De Schipper et al. (2004). We therefore suggest that noise should be viewed as an important aspect of the child care environment on its own. Still, group size and child-caregiver ratio are important factors to take into account, because of the positive association with average noise levels and the negative association with noise variability that we found. Being in a larger group may not be related to less child wellbeing per se, but in larger groups of children more noise is produced with less noise variance. In addition, large group size and unfavorable child-caregiver ratios have been associated with more negative caregiver-child interaction in previous research (De Schipper, Riksen-Walraven, & Geurts, 2006).

The effect sizes for noise levels and noise variability as predictors of child wellbeing in the nonlinear analysis were substantial, given the high percentage of explained variance

and the moderate to large values of the regression coefficients. This indicates that noise should be acknowledged as an important factor in child care contributing to child emotional wellbeing.

Child, caregiver, and care characteristics

The average scores for observed emotional wellbeing indicate that children in our sample were generally feeling quite well in their child care center. Child emotional wellbeing could be predicted from general child care quality, but, surprisingly, not from caregiver sensitivity. Insufficient variance in caregiver sensitivity may explain why we were unable to find a significant relation with wellbeing: Caregivers scored relatively high. In contrast, there was substantial variance in general quality of care, which allowed us to find significant relations with wellbeing. Still, the associations between caregiver sensitivity and quality of care in our study imply that lower child care quality, predictive of lower child wellbeing, often coincides with less sensitive caregiving. It should be noted that the effect size for quality of care as a predictor of child emotional wellbeing was quite small. Another significant predictor for child wellbeing was child age: Within the age range of 0 to 4 years relatively older children showed higher levels of wellbeing, although it should be noted that the effect size was small. Child gender, hours in care, caregiver's years of experience and working hours did not predict child wellbeing.

The correlational nature of this study does not allow us to draw conclusions about cause and effect, for instance between noise levels and child wellbeing or caregiving quality. Still, we tried to shed some light on the causes of high and low noise levels by looking in detail to the sources of noise. The aim of this explorative analysis was to gain insight as to which kind of situations may result in very high or very low noise levels, and thus negatively affect child emotional wellbeing. Sources of noise in this study could be distinguished along two different dimensions. The first dimension, representing child activities on the group versus background activities, was not clearly related to noise levels: Both types of noise sources seem to be related to high and low noise levels. On the second dimension however, outdoor activities were clearly related to high noise levels, whereas indoor activities were related to low noise levels. The latter seem to consider relatively quiet indoor activities, such as children playing with soft toys like dolls and small cars, and caregivers moving about in the room. With respect to child wellbeing, we may speculate that outdoor activities with high noise levels are related to more rough play, which may in turn be related to lower child emotional wellbeing because of accidents with gross motor equipment or peers. On the other hand, very quiet or even dull indoor situations with few or no challenges may result in low wellbeing as well. We can also cautiously relate the sources of noise to the caregiver sensitivity levels, because sensitivity was rated from the same observations as the ones used for the coding of noise sources. Bivariate correlations in this study showed that higher average noise levels and

lower noise variability were related to higher caregiver sensitivity levels. Possibly, higher average noise levels (with less noise fluctuation) are observed in outdoor situations in which caregivers are especially focused on supervising and responding to the children, anticipating children's more rough outdoor play or accidents that may occur when children are riding bicycles and climbing glides. In contrast, in more neutral situations of quiet indoor play caregivers may be occupied by other tasks ('neutral sounds' and 'handling objects' may for instance include handing out sandwiches or clearing away toys) in which they are less focused on the children. It seems that some situations and activities may facilitate or evoke more sensitive caregiving than others. This context-specificity can be taken into account when designing intervention programs aimed at improving caregiver sensitivity and child wellbeing.

Implications for research and practice

This study adds to the scientific literature of environmental chaos theory in center-based child care. We showed that in child care centers noise levels stemming from both social and non-social sources can exceed a threshold after which noise levels become detrimental to child emotional wellbeing. On the other hand, very low levels of noise do not seem beneficial to child wellbeing either, potentially because these noise levels occur when not enough activities or play material are provided. The effects of noise were substantial, so that we conclude that child wellbeing may be improved substantially by taking into account noise levels as an indicator of child care quality.

Our study is, in part, a replication of the investigation by Linting et al. (2013) in a different setting: Home-based care versus center care. Identical procedures, measurements and data analysis were applied which allowed for comparisons of noise levels and variability across settings. It has become apparent that, compared to home-based child care, center child care is a different ecological niche where higher average levels of noise are experienced on a daily basis and where other sources of noise play a role. Still, with this study we further supported the theory that a threshold for noise in relation to child emotional wellbeing is apparent, not only in home-based child care (Linting et al., 2013), but also in center child care. Noise levels form a major indicators of the quality of care that should be integrated in other measures of child care quality. An implication of our study is that noise levels in child care centers should be regulated for optimal child wellbeing. To lower detrimental sound levels in noisy child care centers, interventions to improve acoustics could be applied, such as the placing of absorbent plates (Kishimoto, 2012; Maxwell & Evans, 2000). However, adequate noise levels on the lower end should be anchored as well: Adequate activities and caregiving should provide a minimum level of stimulation. Thus, caregivers might be trained to be sensitive to the lower and upper boundaries of noise level and variability that should be taken into account to optimize child wellbeing. Experiments are needed to evaluate the

effectiveness of these kinds of interventions and to provide more insight in the causal mechanisms between noise levels and child emotional wellbeing. Ultimately, effective interventions may help to protect professionals and children in child care from harmful levels of noise.

Appendix A: Coding system sources of noise

Category	Definitions and examples
1. Outdoor	Traffic / Other noise outside the building
2. Background	Sounds in target group on the background in target group (music from CD player, telephone ringing, doorbell)
	Sounds outside target group (human conversation in hallway, corridor, other groups)
	Other (e.g. front door slamming)
3. Children	Positive vocalizations (talking, laughing, singing)
	Negative vocalizations (crying, screaming, fighting)
	Neutral sounds (clapping, coughing, eating, drinking)
	Throwing/ pushing/ slapping/ moving objects (plates, cups, chairs)
	Moving around (running, crawling, jumping)
	Other *(please comment)*
4. Adults	Positive vocalizations (talking, laughing, singing)
	Negative vocalizations (scolding, shouting, irritated/angry vocalizations)
	Neutral sounds (clapping, coughing, eating, walking)
	Moving objects (plates, cups, chairs, toys, kitchen tools)
	Other *(please comment)*
5. Toys	Large motor toys: e.g. bicycles
	Mechanical / electrical toys (e.g. cars)
	Blocks, Lego, balls
	Musical instruments / musical toys
	Other *(please comment) e.g. plastic toys, play kitchen*

Appendix B: Procedures in nonlinear PCA

Before conducting the CATPCA, we took the steps explained by Linting and Van der Kooij (2012) to ensure the right method and procedures in the analysis. For instance, stability of the CATPCA can be at risk when categories have small marginal frequencies and strongly affect the solution (Linting, Meulman, Groenen, & Van der Kooij, 2007b). We dealt with this issue of stability by requiring a minimum of five observations in each category: When a category had less than five observations it was merged with the closest category by recoding the center's scores on this variable to the adjoining category score. This resulted in 14 dichotomous variables ('Children's positive vocalizations' and 'Adults' positive vocalizations' had scores in category 2 and 3 only; 12 variables had scores in category 0 and 1 only) and five variables with three categories ('Children moving objects', 'Children moving around', 'Adults' neutral sounds', 'Adults moving objects', and 'Toys –other'). One variable ('Background noise: Other') retained the four original categories and as this variable's quantification plot showed a straight line, we treated this variable as numerical. We treated the 14 dichotomous variables as numerical, too; for the remaining five variables we applied an ordinal level (ordinal quantification, with discretizing option 'ranking'). After conducting nonlinear PCA, we checked for outliers in the object plot. We explored two-dimensional and three-dimensional solutions. We considered variables to be fitting adequately in a solution when their component loading on at least one of the dimensions was higher than .300. Component loadings for sources of noise variables for the two-dimensional solution are presented in Table B1.

Table B1
Component loadings of sources of noise variables for the two-dimensional solution in CATPCA (values >.300 in bold)

Source of noise variable	Component loadings	
	Dimension	
	1	2
Cluster 1		
1. Children moving, slapping, throwing objects	-.103	**.408**
2. Adults moving objects (kitchen tools, material, toys)	.067	**.684**
3. Adult neutral sounds (e.g. clapping, coughing, walking)	.170	**.528**
4. Toys sounds: Mechanical toys (e.g. small cars)	.174	**.381**
5. Toys sounds: Other (e.g. plastic kitchen tools)	.329	**.452**
6. Toys: Musical toys*	.236	.239
Cluster 2		
7. Toys sounds: Blocks (e.g. Lego)	**.475**	.266
8. Children moving around (e.g. running, climbing, jumping)	**.695**	.266
9. Children's positive vocalizations	**.705**	-.012
10. Children's negative vocalizations	**.484**	-.105
Cluster 3		
11. Adult negative vocalizations*	.208	-.245
12. Children's neutral sounds (clapping, coughing)	.300	**-.454**
13. Toys sounds: motor development toys (bicycles etc.)	.298	**-.542**
14. Background sounds from outdoor (e.g. traffic passing)	.078	**-.302**
Cluster 4		
15. Children's other sounds	**-.427**	.007
16. Adults' other sounds	**-.383**	.018
17. Background sounds: Media (e.g. CD-player, doorbell)	**-.378**	.249
18. Background sounds : Human conversation outside group	-.299	**.394**
19. Background sounds: Other*	-.180	.184
Other		
20. Adult positive vocalizations*	-.095	-.119

Note. *Component loadings of < .300 on both dimensions: Variable does not seem to fit well in the two-dimensional solution.

Appendix C: Moderator analysis of cross-level interactions

Table C1
Moderator analysis of cross-level interactions (N =245)

Cross-level interaction						
Variable level 1	Variable level 2	Moderator variable	*Beta*	*SE*	*t*	*p*
Wellbeing	Average noise levels					
		child gender[1]	-.01	.02	-.64	.53
		child hours in care[2]	-.00	.00	-.31	.76
Wellbeing	Noise variability					
		child gender	-.02	.07	-.30	.77
		child hours in care	-.00	.00	-.18	.86
Wellbeing	Quality of care					
		child gender	-.08	.07	-1.03	.30
		child hours in care	.00	.01	.49	.62
Wellbeing	Caregiver sensitivity					
		child gender	.08	.06	1.25	.21
		child hours in care	.01	.00	1.53	.13

[1]boy = 0, girl =1; [2] child hours in care was entered as continuous variable

CHAPTER 5

General discussion

General discussion

This dissertation focused on the effectiveness of programs to improve child care quality and on the environmental chaos theory in association with child wellbeing in center care. Our first aim was to provide an overview of existing programs designed to improve child care quality and meta-analytically test their effectiveness. We were particularly interested in narrow-focus programs targeting caregiver behavior and child social-emotional wellbeing through caregiver training, distinguishing them from the broader Early Childhood Education programs (ECE) that have a cognitive focus. Second, we evaluated the effectiveness of the Video-feedback Intervention to promote Positive Parenting in Child Care (VIPP-CC; Groeneveld, Vermeer, Van IJzendoorn, & Linting, 2011; Juffer, Bakermans-Kranenburg, & Van IJzendoorn, 2008a) in child care centers. Finally, we investigated how indicators of the environmental chaos theory (Evans & Wachs, 2010) and noise in particular, may be combined to provide a more comprehensive model for processes involved in child care quality. In this chapter, the results of the three studies are integrated and discussed. In addition, recommendations for child care practice and future research are presented.

Effectiveness of narrow-focus interventions in child care

The meta-analysis in Chapter 2 included 19 randomized controlled trials on programs that targeted the caregiver to improve caregiver-child interaction and child social-emotional wellbeing. The overall effect size was moderate but rather robust (Hedges' *g* = 0.35), indicating that this kind of programs can be implemented to enhance child care quality in a cost-effective manner. The programs were most effective to improve quality at the caregiver level, and to a lesser extent at the classroom level and the child level.

Moderator analysis was restricted because of the relative small number of studies and we used primarily dichotomous moderator variables. We showed that program effectiveness did not differ for type of care, that is, center-based care or home-based child care; or for programs that were implemented within or outside the subsidized Head Start settings. Program intensity and duration, the focus of the intervention, the use of a specific child curriculum with weekly activities, and the use of video did not moderate program effectiveness. Two program characteristics were found to increase effectiveness: the use of a placebo intervention for the control group as compared to care-as-usual and the presence of an individual training component as compared to group-training-only. The increased effectiveness of programs with a placebo intervention may reflect higher quality investigations with more sophisticated measures and solid design. The finding that an individual training component is more effective than group-training-only should be considered preliminary and in need of further consolidation from other trials. It is conceivable that individual caregiver training-on-the-job is more effective than training

caregivers in group contexts through lectures, workshops and video examples, because during individual sessions caregivers can reflect on their own attitudes and behavior more specifically.

It should be noted that the underlying theoretical backgrounds of the programs varied widely and that methods were not always reported clearly. As a result, this meta-analysis could only shed limited light on what makes certain programs more effective than others. Another observation is that most programs included in the meta-analysis targeted children from low SES for whom those programs were sometimes even embedded within ECE programs. We found that only few intervention programs are available for child care services that provide care for children from middle to high SES backgrounds. It may also be that that these programs have not been tested in randomized controlled trials.

Drawing on the VIPP-CC that was tested in home-based care (Groeneveld et al., 2011), we examined the effectiveness of the VIPP-CC in center-based child care, focusing on children from moderate to high SES.

Effectiveness of the VIPP-CC in center child care

Using a randomized controlled trial we tested the effectiveness of the VIPP-CC, a short-term attachment-based intervention program that was adapted for use in child care centers (Groeneveld et al., 2011; Juffer et al., 2008a). Chapter 3 showed that the VIPP-CC was effective ($\eta^2 = 0.07$, i.e. $d = 0.55$) in enhancing caregiver sensitivity: after the intervention, observed caregiver sensitivity increased in the intervention group, but not in the control group. The effect size of the VIPP-CC in center care enhancing observed caregiver sensitivity takes a middle position with regard to effect sizes of the VIPP in other samples targeting observed caregiver sensitivity: effect sizes range from small ($d = 0.33$) for mothers with eating disorders (Stein et al., 2006) to large ($d = 0.78$) for mothers with insecure attachment representations (Klein Velderman, Bakermans-Kranenburg, Juffer, & Van IJzendoorn, 2006). In home-based care the VIPP-CC did not result in increased caregiver sensitivity, but there was an effect for general quality (Groeneveld et al., 2011). The moderate to large effect size of the VIPP-CC in center care is comparable to the effect size for the increase in general quality in home-based care ($d = 0.63$). It should be noted that in our trial on the VIPP-CC the increase in caregiver sensitivity was especially evident in structured play situations, when caregivers interacted with smaller groups of children. This may be explained by the fact that the structured play situations were more comparable to the training situations than, for instance, free play settings. Observed caregiver sensitivity was moderate to high at pretest, yet there was enough variance to prevent us encountering a ceiling effect. It is promising that even for caregivers who show moderate to high levels of sensitive caregiving there is room for improvement with the VIPP-CC.

In our study a small increase in general quality over time was observed in both the intervention and control group, which may reflect heightened awareness of classroom processes and materials through participation in the research project. At the start of the program, we observed a wide range in general quality of the centers in the intervention and control group, including centers of 'inadequate' to 'good' general quality. The considerable non-response in our study might have led to a bias in favor of well-functioning child care centers, despite great efforts to include centers from a wide variety of organizations and neighborhoods.

In addition to the effect for observed sensitivity, we found an intervention effect for caregiver attitudes towards sensitive caregiving and limit setting, similar to findings for home-based child care that were reported by Groeneveld et al. (2011). We reason that positive changes in attitudes may precede behavioral changes, so that at a later stage more positive caregiver behavior may be observed in a wider range of caregiving situations (Susman-Stillman, Pluess, & England, 2013). Finally, caregivers in the intervention group evaluated the program very positively: they reported that the training was very informative, interesting and useful to them. Remarkably, the placebo treatment that consisted of phone calls to talk about general development of children on the group was evaluated quite positively as well by the control group, although significantly less than in the experimental group. The positive evaluations may be speculated to be an important signal from the caregivers reflecting their need for frequent individual professional support.

Noise in center child care

Another goal of this dissertation was to investigate the quality of center child care by combining traditional indicators of quality as described by Riksen-Walraven (2004) with relatively new indicators of quality from the environmental chaos theory as developed by Evans and Wachs (2010). In the correlational study presented in Chapter 4 we predicted child emotional wellbeing from a combination of traditional and new indicators for child care quality.

Traditional indicators of child care quality refer to the distal and proximal factors described in the well-established model of Riksen-Walraven (2004). Distal factors are the more or less fixed aspects of the care setting, such as group size, play materials, and caregiver education level, whilst proximal factors include caregiver-child interactions, child peer interactions and the interaction of the child with the physical environment. According to this model, optimal child care quality can be reached when distal factors support the proximal factors. We combined this model by including indicators of child care quality that refer to environmental chaos, that is, noise, noise variability, crowding, and chaos (Evans and Wachs, 2010). Empirical studies have demonstrated that negative influences from the environment can have adverse consequences for children's health

and development (Evans & Wachs, 2010), for instance high noise levels may impair child cognitive processes and affect child stress regulation (Evans, 2006) and chaotic, unordered caregiving practices were associated with more negative interactions (Evans & Wachs, 2010).

In the reported empirical study, using pretest data from the intervention study described in Chapter 3, we focused on three specific indicators of environmental chaos: noise intensity, noise variability and crowding, as indicated by group size. Our results showed that both noise intensity and variability were strong predictors (Beta = .36 and Beta = .33, respectively) for child emotional wellbeing. Noise levels were higher when group sizes were larger and when children and caregivers were engaged in outdoor activities. In addition, we confirmed a threshold for noise with regard to observed child emotional wellbeing for center-based child care, in line with findings by Linting, Groeneveld, Vermeer and Van IJzendoorn (2013) who first reported a threshold for home-base child care. Nonlinear analyses in our study revealed that lower levels of child wellbeing were related to very high *and* very low levels of noise intensity and variability. This indicates that too much noise is not favorable for young children, but neither is a very quiet child care environment. The threshold for noise may reflect a threshold for stimulation: If stimulation levels are too low, children may get bored and express lower wellbeing. On the other hand, if stimulation levels are too high, children may not be able cope with the child care environment, resulting in lower observed wellbeing. The fact that noise is related with distal care factors such as group size and with proximal processes such as children in interaction with their physical environment strengthens our argument that noise should be viewed as an additional indicator for child care quality. We therefore propose combining the model of Riksen-Walraven (2004) with indicators of environmental chaos (see Figure 1).

Comparing noise in center-based care and home-based child care

This dissertation provides new information on the average noise levels and noise variability in center-based care and adds to the literature on environmental chaos in child care, following previous studies (Groeneveld, Vermeer, Van IJzendoorn, & Linting, 2010; Linting, Groeneveld, Vermeer, & Van IJzendoorn, 2013). Groeneveld and colleagues (2010) conducted empirical studies in two different types of child care and concluded that in center child care observed child emotional wellbeing was lower, caregiver sensitivity was lower, and average noise levels were higher than in home-based care. With regard to the latter finding, we confirmed that in center care average noise levels are significantly higher than in home-based care (d = 1.53). Average noise levels in our study had a mean of 61.69 decibel (SD = 3.30) and ranged from 54.95 decibel to 70.83 decibel. Taking into consideration that an increase of 10 decibel represents noise that is perceived twice as loud, these levels are much higher compared to the levels that were reported for home-

based care, with a mean of 56.54 decibel (*SD* = 3.43) and ranging from 48.45 decibel to 64.56 decibel (Linting et al., 2013). Moreover, in our study noise variability, reflecting the peaks and lows in noise, was on average 7.39 decibel (*SD* = 0.81) and ranged from 6.10 to 9.68 decibel. For home-based care Linting et al. (2013) reported a mean noise variability of 7.85 decibel (*SD* = 1.00) and a range from 4.44 to 10.66 decibel. Although these levels may seem comparable for center-based care and home-based care (*d* = - 0.51), it implies that for children in center care with higher average noise levels absolute peaks are much higher and may even reach 91 decibel, comparable to high way noise. Children in center care are thus more often exposed to high and harmful average noise levels on a daily basis than children in home-based care. Finally, in center child care thresholds for noise in relation to child emotional wellbeing were observed, comparable to the situation in home-based care (Linting et al., 2013). In our study, child emotional wellbeing decreased with increase of noise, below and beyond certain noise levels. It should be noted that the thresholds were observed at different average noise levels for center-based care compared to home-based care.

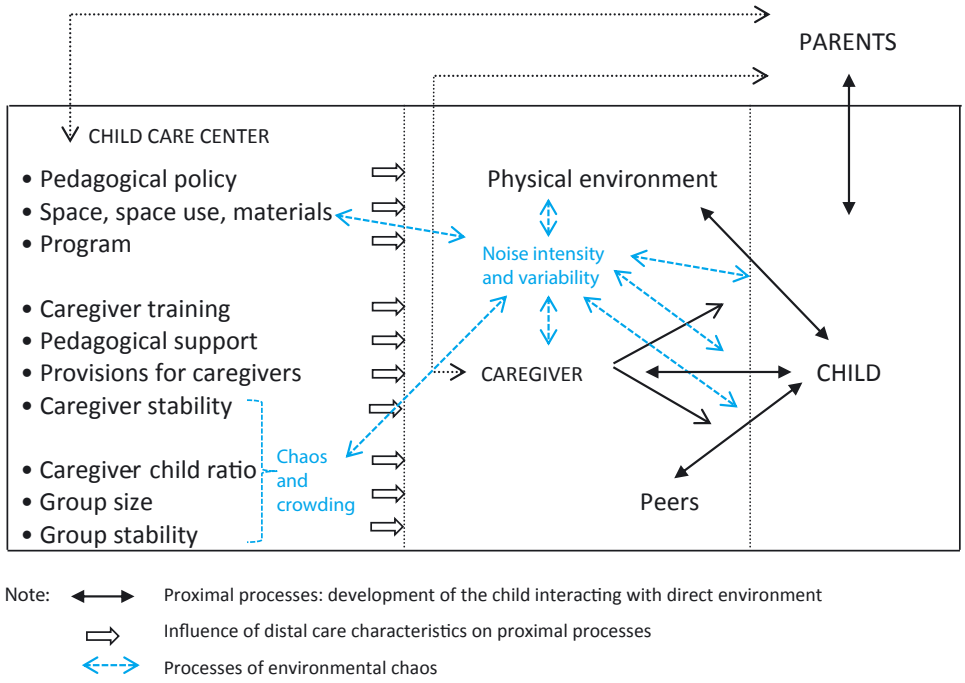


Figure 1. Model of Riksen-Walraven (2004) combined with indicators of environmental chaos

Implications for child care practice and research

Our findings are highly relevant for child care practice and parents, given the fact that 30% to 70% of children under age five attend center based child care in Europe and the U.S.A. (Organization for Economic Co-operation and Development; OECD, 2013). General quality and the caregiver-child relationships need constant monitoring by researchers and authorities, especially when for-profit centers are tempted to compromise caregiver-child ratios in economic low tides. Authorities in the U.S.A. and in the European Union recognize the importance of high quality care and education for children in the preschool age and by subscribing to the Starting Strong II plan (OECD, 2006) they have obliged themselves to address child care quality and the formal education for professional caregivers. Effective narrow-focus intervention programs that target the caregiver could be a start for improving quality. These programs can be implemented at an early stage in the professional career or they can become part of the employee curriculum of centers as an annual training.

In light of the above, the VIPP-CC has the potential to be implemented as a cost-effective way to increase caregiver quality because of the standardized program that has a relatively short duration, that is, six two-hour visits. The VIPP-CC could be implemented as additional training-on-the-job, or as suggested by many participating caregivers, at an earlier stage during internships for caregivers who are still in training. Another suggestion for further monitoring and improving child care quality is that in the future traditional indicators of general child care quality should be combined with indicators of environmental chaos to complete the model of proximal and distal processes that contribute to child wellbeing. Noise can be administered in a relatively easy and objective manner by using digital data loggers. Regulations for child care as monitored by local authorities should include these aspects, so that the wellbeing of children and staff can be ascertained. Moreover, more research is needed regarding interventions to regulate detrimental noise levels and other indicators of environmental chaos in child care.

The choice of parents between center child care and home-based care may depend on specific characteristics of the centers or caregivers in their neighborhood and what is more, on child characteristics such as the child's age, temperament, and susceptibility to the social environment. We stress that continuity of care is most important for young children to establish attachment relationships with non-parental caregivers, which is a fundamental requirement for optimal social-emotional and cognitive development.

Limitations and future directions

Some limitations should be mentioned. The modest sample size in the meta-analysis resulted in restricted moderator analysis and consequently, it remains unclear how the intensity (i.e. the number of sessions) and program duration contribute to effectiveness and which children benefit most or least. According to the differential susceptibility

hypothesis some children are more strongly affected by both negative and positive environmental influences (Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2007; Ellis, Boyce, Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2011) implying that in child care, highly susceptible children are more negatively affected by low quality child care, but also benefit more from high quality care (Pluess & Belsky, 2009). The hypothesis could not be tested in our studies. In our meta-analysis insufficient information was provided on child samples and in our randomized controlled trial of the VIPP-CC the hypothesis was not tested because of a principal focus on caregiver outcomes. Another limitation of our randomized controlled trial is that we only carried out the intervention with one caregiver, although in center child care two or three caregivers are responsible for the group of children. Benefits for children may be higher when caregiving quality in the group is enhanced by simultaneously or successively carrying out the intervention with all caregivers. In a similar vein, Ramchandani and Iles (2013) are carrying out a pilot study to apply the VIPP-SD to both parents in one family. Another limitation is that the sessions and scripts in our program were generally built around activities with three or four children on the group, whilst most of the time caregivers are together with one or two other caregivers responsible for the care of more than ten children on a daily base. Finally, in our study it was not feasible to include a follow-up assessment because of costs and planning of the study. It is valuable to know whether the intervention effects are retained in the long run.

For future research we have some recommendations. More randomized controlled trials should be conducted to the effectiveness of narrow-focus intervention programs and more of these programs should target child care settings serving children from low to moderately high SES. As for the VIPP-CC, we recommend evaluating the effectiveness of the program when carried out with two or three caregivers from one child care group successively or simultaneously. Moreover, long-term effects and effects of the VIPP-CC program on child emotional wellbeing should be examined in the future. Finally, specific interventions should be designed to regulate noise levels in child care centers and their effectiveness should be tested in randomized controlled trials.

In The Netherlands, there is need for a more solid scientific base to be able to compare caregiving quality in center-based care and home-based care. Since the pioneering studies by Groeneveld et al. (2010, 2011), not much research has been conducted on the quality of home-based child care, so that empirical evidence is restricted. Moreover, investigating the quality of home-based care has not yet been conducted as systematically as has been done for center care (Fukkink et al., 2013; Vermeer et al., 2008) which is why trends in quality in home-based care have not yet been established. We suggest to systematically evaluate quality in home-based care over the years, so that quality of center-based child care and home-based care can be compared. In the future, outcomes from those assessments may facilitate parent's choice for the most adequate care for their child.

Conclusion

The quality of center child care is relevant for many children in the preschool age, given the fact that in Western-industrialized countries 30% to 70% of children are in center child care. We showed that beyond traditional indicators, child emotional wellbeing can be predicted from indicators from the environmental chaos theory. To improve child care quality narrow-focus intervention programs that target caregiver-child interaction and child social-emotional development through caregiver training are effective. For child care quality at the caregiver level in particular, the attachment-based short-term intervention VIPP-CC is effective to improve caregiver sensitivity in center child care.

References[1]

Ahnert, L., Pinquart, M., & Lamb, M. E. (2006). Security of children's relationships with nonparental care providers: A meta-analysis. *Child Development, 74*(3), 664-679. doi: 10.1111/j.1467-8624.2006.00896.x

Ainsworth, M. D. S., Bell, S. M., & Stayton, D. (1974). Infant-mother attachment and social development. In M. P. Richards (Ed.), *The introduction of the child onto a social World* (pp. 99-135). London, United Kingdom: Cambridge University Press.

Babisch, W., Schulz, C., Seiwert, M., & Conrad, A. (2012). Noise annoyance as reported by 8-to 14-year old children. *Environment and Behavior, 44*(1), 68-86. doi: 10.1177/0013916510387400

Badanes, L. S., Dmitrieva, J., & Watamura, S. E. (2012). Understanding cortisol reactivity across the day at child care: The potential buffering role of secure attachments to caregivers. *Early Childhood Research quarterly, 27*, 156-165. doi: 10.1016/j.ecresq.2011.05.005

*Baker-Henningham, H., Walker, S., Powell, C., & Gardner, J. M. (2009). A pilot study of the Incredible Years Teacher Training programme and a curriculum unit on social and emotional skills in community pre-schools in Jamaica. *Child: Care, Health and Development*, *35*(5), 624-31. doi: 10.1111/j.1365-2214.2009.00964.x

Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2003). *Vragenlijst voor kennis en attituden over de opvoeding* [Questionnaire concerning knowledge and attitudes towards parenting]. Unpublished manuscript, Leiden University, The Netherlands.

Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., & Juffer, F. (2003). Less is more: Meta-analysis of sensitivity and attachment interventions in early childhood. *Psychological Bulletin, 129*(2), 195-215. doi: 10.1037/0033-2909.129.2.195

Barnas, M. V., & Cummings, E. M. (1994). Caregiver stability and toddler's attachment-related behavior towards caregivers in daycare. *Infant Behavior and Development, 17*, 141-147. doi: 10.1016/0163-6383(94)90049-3

Barnett, W. S. (2011). Effectiveness of early educational intervention. *Science, 333*, 975-978. doi: 10.1126/science.1204534

*Barnett, W., Jung, K., Yarosz, D., Thomas, J., Hornbeck, A., Stechuk, R., & Burns, S. (2008). Educational effects of the Tools of the Mind curriculum: A randomized trial. *Early Childhood Research Quarterly*, *23*(3), 299-313. doi:10.1016/jecresq.2008.03.001

Belfield, C. R., Nores, M., Barnett, S., & Schweinhart, L. (2006). The High/Scope Perry Preschool Program cost-benefit analysis using data from the age-40 follow up. *Journal of Human Resources, 41*(1), 162-190.

Belsky, J. (2006). Early child care and early child development: Major findings of the NICHD study of early child care. *European Journal of Developmental Psychology*, *3*(1), 95-110.doi: 10.1080/17405620600557755

Belsky, J., Burchinal, M., McCartney, K., Lowe Vandell, D., Clarke-Stewart, K. A., & Tresch Owen, M. (2007). Are there long-term effects of early child care? *Child Development, 78*(2), 681-701. doi: 10.1111/j.1467-8624.2007.01021.x

Besharov, D. J., & Morrow, J. S. (2006). Rethinking child care research. *Evaluation Review*, *30*(5), 539-555. doi: 10.1177/0193841X06291522

Binns, C., & Lee, M.K. (2010). The use of probiotics to prevent diarrhea in young children attending child care centers: A review. *Journal of Experimental and Clinical Medicine*, *2*(6), 269-273. doi: 10.1016/j.jecm.2010.08.001.

[1] References marked with an asterix were included in the meta-analyses in Chapter 2

Blok, H., Fukkink, R. G., Gebhardt, E. C., & Leseman, P. P. M. (2005). The relevance of delivery mode and other programme characteristics for the effectiveness of early childhood intervention. *International Journal of Behavioral Development, 29* (1), 35-47. doi: 10.1080/01650250444000315

Bodrova, E., & Leong, D. J. (1996). *Tools of the mind: The Vygotskian approach to early childhood education.* Upper Saddle River, NJ: Prentice-Hall.

Borenstein, M., Rothstein, D., & Cohen, J. (2000). Comprehensive meta-analysis: A computer program for research synthesis [Computer software]. Englewood, NJ: Biostat.

Borenstein, M., Hedges, L.V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* United Kingdom: John Wiley & Sons Ltd.

Bowlby, J. (1969). *Attachment and loss* (Vol. 1). *Attachment*. New York: Penguin Books.

Bronfenbrenner, U. (1979). Contexts of child rearing: Problems and prospects. *American Psychologist, 34*(10), 844-850.

Burchinal, M. R., Cryer, D., Clifford, R. M. & Howes, C. (2002). Caregiver training and classroom quality in child care centers. *Applied Developmental Science*, *6*(1), 2-11. doi: 10.1207/S1532480XADS0601_01

Burchinal, M. R., Lowe Vandell, D., & Belsky, J. (2013, August 12). Is the prediction of adolescent outcomes from early child care moderated by later maternal sensitivity? Results from the NICHD study of Early Child Care and Youth Development. *Developmental Psychology*. Advance online publication. doi: 10.1037/a0033709

Burger, K. (2010). How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. *Early Childhood Research Quarterly, 25*, 140-165. doi:10.1016/j.esqr.2009.11.001

*Cain, D. W., Rudd, L. C., & Saxon, T.F. (2007). Effects of professional development training on joint attention engagement in low-quality childcare centers. *Early Child Development and Care*, *177*(2), 159-185. doi: 10.1080/03004430500375927

Campbell, F. A., Wasik, B. H., Pungello, E., Burchinal, M., Barbarin, O., Kainz, K., ... & Ramey, C. T. (2008). Young adult outcomes of the Abecedarian and CARE early childhood educational interventions. *Early Childhood Research Quarterly*, *23*(4), 452-466. doi: 10.1016/j.ecresq.2008.03.003

Campbell, F. A., Wasik, B. H., Pungello, E., Burchinal, M., Barbarin, O., Kainz, K., … Ramey, C.T. (2012). Adult outcomes as a function of an early childhood educational program: An Abecedarian project follow-up. *Developmental Psychology, 48*(4), 1033-1043. doi: 10.1037/a0026644

Centraal Planbureau (2011). *CPB Notitie Kinderopvang in Kaart.* Retrieved from http://www.cpb.nl/publicatie/kinderopvang-in-kaart/cpb-notitie-27okt2011-kinderopvang-kaart.pdf

Chambers, B., Cheung, A., Slavin, R. E., Smith, D, & Laurenzano, M. (2010*). Effective early childhood education programmes: A best-evidence synthesis*. Retrieved from Center for British Teachers Education Trust website: http://cdn.cfbt.com/~/media/ cfbtcorporate/ files/research/2010/r-early-childhood-programmes-synthesis-2010.pdf.

Clarke, S. H., Campbell, F. A. (1998). Can intervention early prevent crime later? The Abecedarian project compared with other programs. *Early Childhood Research Quarterly, 13*(2), 319-343. doi: 10.1016/S0885-2006(99)80042-8

Clarke-Stewart, K. A., Lowe Vandell, D., Burchinal, M., O'Brien, M., & McCartney, K. (2002). Do regulable features of child-care homes affect children's development? *Early Childhood Research Quarterly*, *17*(1), 52 -86. doi: 10.1016/S0885-2006(02)00133-3

Corapci, F. (2010). Child-care chaos and child development. In G.W. Evans & T.D. Wachs (Eds.). *Chaos and its influence on children's development: An ecological perspective* (pp. 67-82). Washington, DC: American Psychological Association.

Dearing, E., McCartney, K., & Taylor, B. A. (2009). Does higher quality early child care promote low-income children's math and reading achievement in middle childhood? *Child Development*, *80*(5), 1329-1349. doi: 10.1111/j.1467-8624.2009.01336.x

De Kruif, R. E. L., McWilliam, R. A., Maher Ridley, S., & Wakely, M.B. (2000). Classification of teachers' interaction behaviors in early childhood classrooms. *Early Childhood Research Quarterly, 15*(2), 247-268. doi: 10.1016/S0885-2006(00)00051-X

De Kruif, R. E. L., Vermeer, H. J. , Fukkink, R. G., Riksen-Walraven, J. M. A., Tavecchio, L. W. C., Van IJzendoorn, M. H., et al. (2007). *De nationale studie pedagogische kwaliteit kinderopvang: Eindrapport project 0 en 1*. [The national study on childcare quality: Final report project 0 and 1]. Amsterdam, The Netherlands: NCKO.

De Schipper, E. J., Riksen-Walraven, M.J., & Geurts, S. A. (2006). Effects of child–caregiver ratio on the interactions between caregivers and children in child-care centers: An experimental study. *Child Development, 77*(4), 861-874. doi: 10.1111/j.1467-8624.2006.00907.x

De Schipper, J. C., Tavecchio, L.W. C., & Van IJzendoorn, M. H. (2008). Children's attachment relationships with day care caregivers: associations with positive caregiving and the child's temperament. *Social Development, 17*(3), 454-470. doi: 10.1111/j.1467-9507.2007.00448.x

De Schipper, J. C., Tavecchio, L. W., Van IJzendoorn, M. H., & Linting, M. (2003). The relation of flexible child care to quality of center day care and children's socio-emotional functioning: A survey and observational study. *Infant Behavior and Development*, *26*(3), 300-325. doi: 10.1016/S0163-6383(03)00033-X

De Schipper, J. C., Tavecchio, L. W., Van IJzendoorn, M. H., & Van Zeijl, J. (2004). Goodness-of-fit in center day care: Relations of temperament, stability, and quality of care with the child's adjustment. *Early Childhood Research Quarterly*, *19*(2), 257-272. doi: 10.1016/j.ecresq.2004.04.004

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, *333*(6045), 959-964. doi: 10.1126/science.1204529

Driessen, G. W. J. M. (2004). A large-scale longitudinal study of the utilization and effects of early childhood education and care in The Netherlands. *Early Child Development and Care, 174* (7-8), 667-689, doi: 10.1080/0300443042000187158

*Domitrovich, C. E., Cortes, R. C., & Greenberg, M. T. (2007). Improving young children's social and emotional competence: A randomized trial of the preschool "PATHS" curriculum. *The Journal of Primary Prevention, 28*(2), 67-91.
doi: 10.1007/s10935-007-0081-0

*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. doi: 10.3102/0002831208328089

Domitrovich, C. E., Greenberg, M. T., Kusché, C., & Cortes, R. (2005). *The preschool PATHS curriculum*. South Deerfield, MA: Channing Bete.

*Driscoll, K. C., & Pianta, R. C. (2010). Banking Time in Head Start: Early efficacy of an intervention designed to promote supportive teacher–child relationships. *Early Education & Development*, *21* (1), 38-64. doi: 10.1080/10409280802657449

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56 (2), 455-463.

Enmarker, I., & Boman, E. (2004). Noise annoyance responses of middle school pupils and teachers. *Journal of Environmental Psychology*, *24*, 527-526. doi: 10.1016/j.envp.2004.90.005

Erickson, M. F., Sroufe, L. A., & Egeland, B. (1985). The relationship between quality of attachment and behavior problems in preschool in a high risk sample. In I. Bretherton and E. Waters (Eds.), *Child Development Monographs, 50*, 147-166.

Evans, G. W. (2006). Child development and the physical environment. *Annual Review of Psychology, 57*, 423-451. doi: 10.1146/annrev.psych.57.102904.190057

Evans, G. W., & Wachs, T. D. (Eds.) (2010). *Chaos and its influence on children's development: An ecological perspective*. Washington, DC: American Psychological Association.

Eysel-Gosepath, K., Daut, T., Pinger, A., Lehmacher, W., & Erren, T. (2012). Sound levels and their effects on children in German primary school. *European Archives of Oto-rhino-laryngology*, 269, 2475-2483. doi: 10.1007/s00405-011-1899-x

Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191. doi: 10.3758/BF03193146

Fukkink, R., & Lont, A. (2007). Does training matter? A meta-analysis and review of caregiver training studies. *Early Childhood Research Quarterly, 22*, 294-311. doi:10.1016.j/esqr.2007.04.005

*Fukkink, R.G., & Tavecchio, L. W. C. (2010). Effects of Video Interaction Guidance on early childhood teachers. *Teaching and Teacher Education*, *26*(8), 1652-1659. doi:10.1016/j.tate.2010.06.016

Fukkink, R. G., Gevers Deynoot-Schaub, M. J. J. M., Helmerhorst, K.O.W., Bollen I., & Riksen-Walraven, J. M. A. (2013). *Pedagogische kwaliteit van de kinderopvang voor 0-4 jarigen in Nederlandse kinderdagverblijven in 2012.* Amsterdam, The Netherlands: NCKO.

Gerber. E. B., Whitebook, M., & Weinstein, R. S. (2007). At the heart of child care: Predictors of teacher sensitivity in center-based child care. *Early Childhood Research Quarterly, 22*, 327-346. doi: 10.1016/j.ecresq.2006.12.003

*Girard, L. C., & Girolametto, L. (2011). Training early childhood educators to promote peer interactions: Effects on children's aggressive and prosocial behaviors. *Early Education & Development*, 22 (2), 305-323. doi:10.1080/10409281003668060

*Girolametto, L., Weitzman, E., & Greenberg, J. (2003). Training day care staff to facilitate children's language. *American Journal of Speech Language Pathology, 12*, 299-311. doi: 10.1044/1058-0360(2003/076)

*Girolametto, L., Weitzman, E., & Greenberg, J. (2004). The effects of verbal support strategies on small-group peer interactions. *Language, Speech and Hearing Services in Schools, 35*, 254-268. doi: 10.1044/0161-1461(2004/024)

Goldstein, H., & Woodhouse, G. (1996). *Multilevel models with missing data*. Paper presented at the 11th International workshop on statistical modeling, Orvieto, Italy.

Goossens, F. A., & Van IJzendoorn, M. H. (1990). Quality of infant's attachment to professional caregivers: Relation to infant-parent attachment and day-care characteristics. *Child Development, 61*, 832-837. doi: 10.1111/j.1467-8624.1990.tb02825.x

Groeneveld, M. G., Vermeer, H. J., Van IJzendoorn, M. H. & Linting, M. (2010). Children's wellbeing and cortisol levels in home-based and center-based child care. *Early Childhood Research Quarterly, 25*, 502-514. doi: 10.1016/ecresq.2009.12.004

*Groeneveld, M. G., Vermeer, H.J., Van IJzendoorn, M. H., & Linting, M. (2011). Enhancing home-based child care quality through video-feedback intervention: A randomized controlled trial. *Journal of Family Psychology, 25*(1), 86-96. doi: 10.1037/a0022451

Haines, M. M., Stansfeld, S. A., Brentnall, S., Head, J., Berry, B., Jiggings, M., & Hygge, S. (2001). The West London schools study: the effects of chronic aircraft noise exposure on child health. *Psychological Medicine, 31*, 1385-1396. doi:10.1017/S003329170100469X

Hambrick-Dixon, P. J. (1986). Effects of experimentally imposed noise on task performance of black children attending day care centers near elevated subway trains. *Developmental Psychology, 22,* 259-264. doi: 10.1037/0012-1649.22.2.259

Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early Childhood Environment Rating Scale-Revised*. New York, NY: Teachers College Press.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*, 1900-1902. doi: 10.1126/science.1128898

Helmerhorst, K.O. W., Riksen-Walraven, J. M., Vermeer, H. J., Fukkink, R. G., & Tavecchio, L.W.C. (2014). Measuring the interactive skills of caregivers in child care centers: Development and validation of the Caregiver Interaction Profile Scales. *Early Education & Development, 0*, 1-21. doi: 10.1080/10409289.2014.840482

Howes, C., Rodning, C., Galluzzo, D. C., & Myers, L. (1988). Attachment and child care: Relationships with mother and caregiver. *Early Childhood Research Quarterly*, *3*(4), 403-416.

Howes, C., & Smith, E. W. (1995). Relations among child care quality, teacher behavior, children's play activities, emotional security, and cognitive activity in child care. *Early Childhood Research Quarterly*, *10*(4), 381-404.

Howes, C., & Spieker S. (2008). Attachment relationships in the context of multiple caregivers. In J. Cassidy, & P.R. Shaver (Eds.), *Handbook of Attachment: Theory, research and clinical applications* (2nd ed.) (pp. 317- 332). New York, NY: Guilford.

Howes, C., Whitebook, M., & Philips, D. (1992a). Thresholds of quality: Implications for the social development for children in center-based child care. *Child Development, 63*, 449-460. doi: 10.1111/1467-8624.ep9207061023

Howes, C., Whitebook, M., & Philips, D. (1992b). Teacher characteristics and effective teaching in child care: Findings from the National Child Care Staffing study. *Child & Youth Care Forum, 21*(6), 399-414. doi: 10.1007/BF00757371

Hox, J. (2010). *Multilevel analysis: Techniques and applications*. New York, NY: Routledge.

Hungerford, A., & Cox, M. J. (2006). Family factors in child care research. *Evaluation Review, 30* (5), 631-655. doi: 10.1177/0193841X06291532

Izard, C. E. (2001). *The Emotions Course. Helping children understand and manage their feelings: An emotion-centered primary prevention program for Head Start. Teachers Manual.* Unpublished manuscript, Newark, DE: University of Delaware.

*Izard, C. E., Trentacosta, C. J., King, K.A., Mostow, A. J. (2004). An emotion-based prevention program for Head Start Children. *Early Education & Development, 15* (4), 407-422. doi: 10.1207/s15566935eed1504_4.

Janus, M., & Brinkman, S. (2010). Evaluating Early Childhood Education and Care programs. In E. Baker, B. McGaw, & P. Peterson (Eds.), *The International Encyclopedia of Education* (pp. 25-31). Elsevier.

Jones, S. M., Bub, K. L., & Raver, C. C. (2013). Unpacking the black box of the Chicago School Readiness Project intervention: The mediating roles of teacher–child relationship quality and self-regulation. *Early Education & Development, 24* (7), 1043-1064. doi: 10.1080/10409289.2013.825188

Juffer, F., Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2005). The importance of parenting in the development of disorganized attachment: Evidence from a preventive intervention study in adoptive families. *Journal of Child Psychology and Psychiatry, 46*(3), 263-274. doi: 10.1111/j.1469-7610.2004.00353.x

Juffer, F., Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2008a). *Promoting Positive Parenting: An attachment-based intervention.* New York, NY: Lawrence Erlbaum Associates.

Juffer, F., Bakermans-Kranenburg, M.J., & Van IJzendoorn, M. H. (2008b). Supporting adoptive families with video-feedback intervention. In F. Juffer, M. J. Bakermans-Kranenburg, & M. H. van IJzendoorn (Eds.), *Promoting Positive Parenting: An attachment-based intervention* (pp. 139-153). New York, NY: Lawrence Erlbaum Associates.

Kalinauskiene, L., Cekuoliene, D., Van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., Juffer, F., & Kusakovskaja, I. (2009). Supporting insensitive mothers: The Vilnius randomized control trial of video feedback intervention to promote maternal sensitivity and infant attachment. *Child: Care, Health & Development, 35,* 613-623. doi: 10.1111/j.1365-2214.2009.00962.x

Kishimoto, T. M. (2012). The integration of care and education: a case concerning the problem of noise. *European Early Childhood Education Research Journal, 20*(4), 493-503. doi: 10.1080/1350293x.2012.737240

Kjellberg, A., Landström, U., Tesarz, M., Söderberg, L., & Akerlund, E. (1996). The effects of nonphysical noise characteristics, ongoing task and noise sensitivity on annoyance and distraction due to noise at work. *Journal of Environmental Psychology, 16,* 123-136. doi: 10.1006/jevp.1996.0010

Klatte, M., Bergström, K., & Lachmann, T. (2013). Does noise affect learning? A short review of noise effects on cognitive performance in children. *Frontiers in Developmental Psychology, 4,* 1664-1078. doi: 10.3389/fpsyg.2013.00578

Klein Velderman, M. Bakermans-Kranenburg, M. J., Juffer, F., & Van IJzendoorn, M. H. (2006). Effects of attachment-based interventions on maternal sensitivity and infant attachment: Differential susceptibility of highly reactive infants. *Journal of Family Psychology, 20,* 266-274**.** doi: 10.1037/0893-3200.20.2.266

Larsson, N., Ward, D. S., Neelon, S. B., & Story, M., (2011). What role can child-care settings play in obesity prevention? A review of the evidence and call for research efforts. *Journal of the American Dietic Association, 111*(9), 1343-1362. doi: 10.1016/j.jada.2011.06.007

Lindstrom, F., Persson Waye, K., Sodersten, M., McAllister, A., & Ternstrom, S. (2011). Observations of the relationship between noise exposure and preschool teacher voice usage in day-care environments. *Journal of Voice, 25*(2), 166-172. doi: 10.1016/j.jvoice.2009.09.009

Linting, M., Groeneveld, M. G., Vermeer, H. J., & Van IJzendoorn, M. H. (2013). Threshold for noise in daycare: Noise level and noise variability are associated with child wellbeing in home-based childcare. *Early Childhood Research Quarterly, 28*, 960-971. doi: 10.1016/j.ecrq.2013.03.005.

Little, R. J. A. (1988). Missing-data adjustments in large surveys. *Journal of Business and Economic Statistics,6,* 287-296. doi: 10.1080/07350015.1988.10509663

Love, J. M., Harrison, L. Sagi-Schwartz, A., Van IJzendoorn, M. H., Ross, C., Ungerer, J. A., … & Chazan-Cohen, R. (2003). Child care quality matters: How conclusions may vary with context. *Child Development, 74*(4), 1021-1033. doi: 10.1007/BF00757371

Lowe Vandell, D., Belsky, J., Burchinal, M., Steinberg, L., & Vandergrift, N. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD Study of Early Child Care and Youth Development. *Child Development, 81*(3), 737-756. doi: 10.1111/j.1467-8624.2010.01431.x

Machado, J. (2003). *Early childhood experiences in the language arts* (7th ed.). New York, NY: Delmar.

Manlove, E. E., Frank, T., & Vernon-Feagans, L. (2001). Why should we care about noise in classrooms and daycare settings? *Child & Youth Care Forum, 30* (10), 55-64.

Matheny, A. P., Wachs, T. D., Ludwig, J. L., & Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the Confusion, Hubbub, and Order Scale. *Journal of Applied Developmental Psychology*, *16*(3), 429-444. doi: 10.1016/0193-3973(95)90028-4

Maxwell, L. E., & Evans, G.W. (2000). The effects of noise on preschool children's pre-reading skills. *Journal of Environmental Psychology, 20*, 91-97. doi:10.1006/jevp.19999.01444.

Maxwell, L. (2010). Chaos outside the home: the school environment. In G.W. Evans & T. D. Wachs (Eds.). *Chaos and its influence on children's development: An ecological perspective* (pp. 83-95). Washington, DC: American Psychological Association.

McAllister, A.M., Granqvist, S., Sjolander, P., Sundberg, J. (2009). Child voice and noise: a pilot study of noise in day cares and the effects on 10 children's voice quality according to perceptual evaluation. *Journal of Voice, 23* (5), 587-593. doi: 10.1016/j.jvoice.2007.10.017

McCartney, K., Burchinal, M., Clarke-Stewart, A., Bub, K. L., Owen, M. T., & Belsky, J. (2010). Testing a series of causal propositions relating time in child care to children's externalizing behavior. *Developmental Psychology*, *46*(1), 1-17.doi: 10.1037/a0017886

Meij, H., Mutsaers, K., & Pennings, T. (2009). *Effectiviteit van voor- en vroegschoolse programma's in Nederland*. The Netherlands: Nederlands Jeugd Instituut.

Mesman, J., Stolk, M.N., Van Zeijl, J., Alink, L.R.A., Juffer, F., Bakermans-Kranenburg, M.J., Van IJzendoorn, M.H., & Koot, H.M. (2008). Extending the video-feedback intervention to sensitive discipline: The early prevention of antisocial behavior. In F. Juffer, M. J. Bakermans-Kranenburg, & M. H. Van IJzendoorn (Eds.), *Promoting positive parenting: An attachment-based intervention* (pp. 171-191). New York, NY: Lawrence Erlbaum Associates.

Meulman, J. J., Heiser, W. J., & IBM SPSS (2012). *IBM SPSS Categories 21.0*. Chicago, IL: SPSS.

Moon, R.Y., Calabrese, T. & Aird, L. (2008). Reducing the risk of sudden infant death syndrome in child care and changing provider practices: Lessons learned from a demonstration Project. *Pediatrics, 122,* 788-798. doi: 10.1542/peds.2007-3010.

Nederlands Consortium Kinderopvang Onderzoek (2006). *Schalen leidstervaardigheden. Handleiding (verkort)*. [Caregiver scales. Manual (shortened version)]. Amsterdam, The Netherlands: NCKO.

Negrão, M., Pereira, M., Mesman, J., & Soares, I. (August, 2013*). Enhancing positive parent-child interactions in at-risk families: The VIPP-SD in Portugal*. Poster presented at the International Attachment Conference, Pavia, Italy.

Nelson, G., Westhues, A., & MacLeod, J. (2003). A meta-analysis of longitudinal research on preschool prevention programs for children. *Prevention & Treatment, 6*(1), 31a. doi:10.1037/1522-3736.6.1.631a

*Neuman, S.B., & Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal*, *46*(2), 532-566. doi: 10.3102/0002831208328088.

Noailly, J., & Visser, S.(2009). The impact of market forces on child care provision: Insights from the 2005 Child Care Act in the Netherlands. *Journal of Social Policy, 38*(3), 477-498. doi: 10.1017/S0047279409003109

Nores, M., Belfield, C.R., Barnett, W.S., & Schweinhart, L. (2005). Updating the economic impacts of the High/Scope Perry Preschool Program. *Educational Evaluation and Policy Analysis, 27*(3), 245-261. doi: 10.3102/01623737027003245

Organization for Economic Co-operation and Development (2006). *Starting Strong II: Early Childhood Education and Care*. OECD Publishing.

Organization for Economic Co-operation and Development, Family Database, Social Policy Division and Directorate of Employment (2013). Labour and Social Affairs *PF3.2: Enrolment in childcare and pre-schools* Retrieved from:http://www.oecd.org/els/family/ PF3.2%20Enrolment%20in%20childcare%20and%20preschools%20-%20290713.pdf

Patterson, G. R (1982). Coercive family processes. Eugene, OR: Castilia.

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes. C., Kagan, S. L., & Yazejian, N. (2001). The relation of preschool childcare quality to children's cognitive and social developmental trajectories through second grade. *Child Development, 72*(5), 1534-1553.

Perlman, M., Zellman, G. L., & Le, V. N. (2004). Examining the psychometric properties of the Early Childhood Environment Rating Scale-Revised (ECERS-R). *Early Childhood Research Quarterly*, *19,* 398-412. doi 10.1016/j.ecresq.2004.07.006

Pianta R. C., & Hamre, B. (2001). *Students, teachers, and relationship support (STARS).* Lutz, FL: Psychological Assessment Resources.

Ramey, C. T., & Landesman Ramey, S. (1998). Early intervention and early experience. *American Psychologist, 53* (2), 109-120. Doi: 10.1037/0003-066X.53.2.109.

*Raver, C. C., Jones, S. M., Li-Grining, C.P., Metzger, M., Champion, K. M., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly, 23*(1), 10-26. doi: 10.1016/j.ecresq. 2007.09.001

Raver, C. C., Li-Grining, C., Bub, K., Jones, S. M., Zhai, F., & Pressler, E. (2011). CSRP's impact on low-income preschoolers' preacademic skills: Self-regulation as a mediating mechanism. *Child Development, 82*(1), 362-378. doi:10.1111/j.1467-8624.2010.01561.x

Rijlaarsdam, J. (2014). *Social disadvantage and child emotional and behavioural problems: At HOME in the Netherlands.* Doctoral dissertation, Erasmus MC: University Medical Center Rotterdam.

Riksen-Walraven, M. (2004). Pedagogische kwaliteit in de kinderopvang: Doelstellingen en kwaliteitscriteria. In R. Van IJzendoorn, L. Tavecchio & M. Riksen-Walraven (Eds.), *De kwaliteit van de Nederlandse kinderopvang* (pp. 100-123). Amsterdam, The Netherlands: Uitgeverij Boom.

Rubin, D. B. (1986). Statistical matching using file concatenation with adjusted weights and multiple imputations. *Journal of Business and Economic Statistics, 4,* 87-94.

*Rusby, J. C., Smolkowski, K., Marquez, B., & Taylor, T. K. (2008). A small-scale randomized efficacy trial of Carescapes: Enhancing children's social development in child care homes. *Early Childhood Research Quarterly*, *23*(4), 527-546. doi:10.1016/j.ecresq.2008.08.003

*Rusby, J. C., Taylor, T. K., & Marquez, B. (2004). Promoting positive social development in family childcare settings. *Early Education & Development, 15*(4), 449-470. doi: 10.1080/10409289.2004.10532130

Rutten, S. (2009). *Voor-en Vroegschoolse educatie in een veranderende omgeving. Stand van zaken en een blik vooruit*. The Netherlands: Sardes.

Sala, E., Airo, E., Olkinuora, P., Simberg, S., Strom, U., Laine, A.,… & Suonpaa, J. (2002). Vocal loading among day care center teachers. *Logopedics Phoniatrics Vocology 27*, 21-28. doi: 10.1080/140154302760146943

Sardes (2012). *Intergrale kindcentra. Visie document*. The Netherlands: Sardes.

Shager, H. M., Schindler, H. S., Magnuson, K. A., Duncan, G. J., Yoshikawa, H., & Hart, C. M. D. (2013). Can research design explain variation in Head Start research results? A meta-analysis of cognitive and achievement outcomes. *Educational Evaluation and Policy Analysis, 35*(1), 76 -95. doi: 10.3102/0162373712462453

Shield, B. M., & Dockrell, J. E. (2003). The effects of noise on children at school: A review. *Building Acoustics, 10*(2), 97-116. doi: 10.1260/135101003768965960

*Snyder, J., Low, S., Schultz, T., Barner, S., Moreno, D., Garst, M., & Schrepferman, L. (2011). The impact of brief teacher training on classroom management and child behavior in at-risk preschool settings: Mediators and treatment utility. *Journal of Applied Developmental Psychology*, *32*(6), 336-345. doi:10.1016/j.appdev.2011.06.001

Spieker, S. J., Nelson, D. C., Petras, A., Jolley, S. N., & Barnard, K. E. (2003). Joint influence of child care and infant attachment security for cognitive and language outcomes of low-income toddlers. *Infant Behavior and Development*, *26*(3), 326-344. doi: 10.1016/S0163-6383(03)00034-1

Stansfeld, S. A., Clark, C., Cameron, R. M., Alfred, T, Head, J., Haines, M. M., Van Kamp, I., Van Kempen, E., & Lopez-Barrio, I. (2009). Aircraft and road traffic noise exposure and children's mental health. *Journal of Environmental Psychology, 29*, 203-207. doi: 10.1016/j.jenvp.2009.01.002

Stein, A., Woolley, H., Senior, R., Hertzmann, L., Lovel, M., Lee, J. …Fairburn, C.G. (2006). Treating disturbances in the relationship between mothers with bulimic eating disorders and their infants: A randomized, controlled trial of video feedback. *The American Journal of Psychiatry, 163,* 899-906. doi: 10.1176/appi.ajp.163.5.899

Susman-Stillman, A., Pleuss, J., & Englund, M. M. (2013). Attitudes and beliefs of family- and center-based child care providers predict differences in caregiving behavior over time. *Early Childhood Research Quarterly*, 28 (4), 905-917. doi: 10.1016/j.ecresq.2013.04.003

Sutton, A. J., Duval, S. J., Tweedie, R. L., Abrams, K. R., & Jones, D. R. (2000). Empirical assessment of effect of publication bias on meta-analyses. *British Medical Journal*, *320* (7249), 1574-1577. doi: 10.1136/bmj.320.7249.1574

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), 109-124. doi:10.1177/1476718x10387900

Torquati, J. C., Raikes, H., & Huddleston-Casa, A. (2007). Teacher education, motivation, compensation, workplace support, and links to quality of center-based child care and teachers' intention to stay in the early childhood profession. *Early Childhood Research Quarterly, 22*, 261-275. doi: 10.1016/j.ecresq.2007.03.004

Twisk, J. W. (2006). *Applied multilevel analysis: A practical guide*. Cambridge University Press.

Tziner, A., Haccoun , R. R. & Kadish, A. (1991). Personal and situational characteristics influencing the effectiveness of transfer of training improvement strategies. *Journal of Occupational Psychology, 64*, 167-177.

Van Buuren, S. (2010). Multiple imputation of multilevel data. *The Handbook of Advanced Multilevel Analysis*, 173-196. Books.Google.Com.

Vandell, D. L., Belsky, J., Burchinal, M., Steinberg, L., & Vandergrift, N. (2010). Do effects of early child care extend to age 15 years? Results from the NICHD study of early child care and youth development. *Child Development*, *81*(3), 737-756.

Van der Kooij, A. J. (2007). Prediction accuracy and stability of regression with optimal scaling transformations (Doctoral thesis, Leiden University, The Netherlands). Retrieved from https://openaccess.leidenuniv.nl/handle/1887/12096

Van Zeijl, J., Mesman, J., Van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., Juffer, F., Stolk, M.N., Koot, H. M., & Alink, L. R. A. (2006). Attachment-based intervention for enhancing sensitive discipline in mothers of one- to three-year-old children at risk for externalizing behavior problems. *Journal of Consulting and Clinical Psychology, 74(6),* 994-1005. doi: 10.1037/0022-006X.74.6.994

Vermeer, H. J. (July, 2012). *The ECERS-R Brief: Psychometric Properties*. Poster presented at the biennial meeting of the International Society for the Study of Behavioral Development, Alberta, Canada.

Vermeer, H. J., & Bakermans-Kranenburg, M. J. (2008). Attachment to mother and non-maternal care: Bridging the gap. *Attachment & Human Development, 10*(3), 263-273. doi: 10.1080/14616730802113588

Vermeer, H. J., Van IJzendoorn, M. H., De Kruif, R. E. L., Fukkink, R. G., Tavecchio, L. W. C., Riksen-Walraven, J. M., & Van Zeijl, J. (2008). Child care in The Netherlands: Trends in quality over the years 1995-2005. *The Journal of Genetic Psychology, 169*(4), 360-385.

Vesely, C. K., Brown, E. L., & Mahatmya, D. (2013). It takes two: Sensitive caregiving across contexts and children's social, emotional, and academic outcomes. *Early Education & Development*, *24*(7), 960-978. doi: 10.1080/10409289.2013.825185

Votruba-Drzal, E., Levine Coley, R., Maldonado-Carreno, C., Li-Grining, C.P., & Chase-Lansdale, L. P. (2010). Child care and the development of behavior problems among economically disadvantaged children in middle childhood. *Child Development, 81*(5), 1460-1474. doi: 10.1111/j.1467-8624.2010.01485.x

Ward, D.S., Vaughn, A., McWilliams, C., & Hales, D. (2010). Interventions for increasing physical activity at child care. *Medicine & Science in Sports & Exercise, 42*(3), 526-534.doi: 10.1249/MSS.0b013e3181cea406

Watamura, S. E., Phillips, D. A., Morrissey, T. W., McCartney, K., & Bub, K. (2011). Double jeopardy: Poorer social-emotional outcomes for children in the NICHD SECCYD experiencing home and child-care environments that confer risk. *Child Development*, *82*(1), 48-65. doi: 10.1111/j.1467-8624.2010.01540.x

Webster-Stratton, C. (2004). *The Incredible Years Teacher Training Series*. Seattle, WA: Incredible Years.

Weikart, D.P. (1998). Changing early childhood development through educational intervention. *Preventive Medicine, 27*, 233-237. doi: 10.1006/pmed.1998.0280

Weitzman, E. (1994). The Hanen program for early childhood educators: In-service training for child care providers on how to facilitate children's social, language and literacy development. *Infant-Toddler Intervention: The Transdisciplinary Journal, 4*, 173-202.

Woolley, H., Hertzmann, L., & Stein, A. (2008). Video-feedback intervention with mothers with postnatal eating disorders and their infants. In F. Juffer, M .J. Bakermans-Kranenburg, & M. H. Van IJzendoorn (Eds.), *Promoting Positive Parenting: An attachment-based intervention* (pp. 111-138). New York, NY: Lawrence Erlbaum Associates.

*Zhai, F., Raver, C.C., & Li-Grining, C. (2011). Classroom-based interventions and teachers' perceived job stressors and confidence: Evidence from a randomized trial in Head Start settings. *Early Childhood Research Quarterly*, *26*(4), 442-452. doi: 10.1016/j.ecresq. 2011.03.003

Nederlandse samenvatting (Summary in Dutch)

Voor een groot aantal werkende ouders in westerse landen is kinderopvang een onmisbare vorm van ondersteuning in de opvoeding (Organization for Economic Co-operation and Development, 2013). Een belangrijke vraag is hoe het gesteld is met de kwaliteit van de kinderopvang en hoe deze kwaliteit verbeterd kan worden. Uit een grootschalig longitudinaal onderzoek in de Verenigde Staten is niet alleen gebleken dat opvang van slechte kwaliteit nadelige uitkomsten heeft voor de sociaal-emotionele ontwikkeling maar ook dat hoge kwaliteit van kinderopvang kan zorgen voor een betere cognitieve ontwikkeling (Vandell, Belsky, Burchinal, Steinberg, & Vandergrift, 2010). Kwaliteit van kinderopvang kan worden vastgesteld aan de hand van distale of structurele aspecten, zoals de aanwezigheid van bepaald ontwikkelings- en speelmateriaal, groepsgrootte en afspraken met betrekking tot veiligheid en hygiëne (Riksen-Walraven, 2004). De belangrijkste kwaliteitsaspecten zijn echter de proximale kenmerken: de omgang tussen de pedagogisch medewerkers en de kinderen, en tussen kinderen onderling (Riksen-Walraven, 2004). Vanuit de gehechtheidstheorie is de rol die de pedagogisch medewerker kan vervullen in de zorg voor de kinderen essentieel (Ahnert, Pinquart, & Lamb, 2006; De Schipper, Tavecchio, & Van IJzendoorn, 2008). Gebleken is dat afgezien van de ouders ook pedagogisch medewerkers belangrijke gehechtheidsfiguren kunnen zijn voor kinderen in hun groep, dat wil zeggen, vertrouwde personen naar wie de kinderen toe kunnen gaan als ze verdrietig, moe of bang zijn en bij wie ze veiligheid en troost kunnen zoeken (Bowlby, 1969; Goossens & Van IJzendoorn, 1990). Voor het opbouwen van deze band is niet alleen continuïteit van dezelfde pedagogisch medewerkers op een groep essentieel, maar ook dat zij sensitief zijn, wat betekent dat zij signalen van de kinderen herkennen en vervolgens tijdig en met een adequate reactie op het gedrag reageren (Ainsworth, Bell, & Stayton, 1974).

In dit proefschrift zijn drie studies (een meta-analyse, een experimenteel onderzoek met gerandomiseerde onderzoeksopzet en een correlationeel onderzoek) beschreven waarin de proximale kwaliteitskenmerken in de kinderopvang centraal staan. We onderzoeken daarbij ook de samenhang met de distale kenmerken, zoals beschreven in het model van Riksen-Walraven (2004). Vernieuwend is dat we ook aspecten uit de theorie van omgevingschaos (Evans & Wachs, 2010) aan het model toevoegen als indicatoren van kwaliteit van kinderopvang. Volgens deze theorie is een hoge mate van chaos in de omgeving nadelig voor het emotionele en fysiek welbevinden (Evans & Wachs, 2010). Er is sprake van hoge mate van omgevingschaos in het geval van hoge geluidsniveaus, als er (te)veel mensen in één ruimte zijn, en bij het ontbreken van routines, planning en orde (Matheny, Wachs, Ludwig, & Philips, 1995).

De Nederlandse kinderopvang

Twee empirische studies in dit proefschrift betreffen de opvang van kinderen in kinderdagverblijven in Nederland. In kinderdagverblijven worden kinderen tot vier jaar opgevangen en verzorgd in groepen tot 16 kinderen door twee tot drie pedagogisch medewerkers per groep. Tot de jaren '90 behoorde de pedagogische kwaliteit van de Nederlandse kinderopvang in internationaal opzicht tot één van de beste. Sinds 1995 is deze kwaliteit echter sterk afgenomen (Fukkink, Gevers Deynoot-Schaub, Helmerhorst, Bollen, & Riksen-Walraven, 2013; Vermeer et al., 2008). Mogelijk is deze daling in kwaliteit het gevolg van de marktwerking in de kinderopvang (www.fnv.nl; Noailly & Visser, 2009). Momenteel wordt door ouders en experts gevreesd dat in de huidige economische crisis, waarbij sprake is van veel vraaguitval, de pedagogische kwaliteit in veel centra verder moet inboeten om bedrijven winstgevend te houden. Ondanks de zorgen over de matige kwaliteit van de kinderopvang in Nederland, worden er nog maar weinig programma's ingezet om de kwaliteit te verbeteren. Bestaande programma's zijn vooral gericht op de cognitieve ontwikkeling van de kinderen, zoals de Vroeg en Voorschoolse Educatie (VVE) programma's (Driessen, 2003). Bovendien zijn er in Nederland, maar ook internationaal gezien, überhaupt zeer weinig trainingsprogramma's voor de kinderopvang wetenschappelijk getoetst op effectiviteit (Meij, Mutsaers, & Penning, 2009; Rutten, 2009). Een specifiek aandachtspunt van dit proefschrift is daarom de effectiviteit van interventieprogramma's, in het bijzonder vaardigheidstrainingen voor pedagogisch medewerkers.

Meta-analyse naar de effectiviteit van vaardigheidstrainingen

Hoofdstuk 2 biedt een overzicht van de huidige stand van zaken naar de effectiviteit van vaardigheidstrainingen voor pedagogisch medewerkers. De meta-analyse in dit hoofdstuk omvat 19 experimentele studies waarbij gebruik is gemaakt van randomisatie en een controlegroep. De onderzochte trainingsprogramma's waren relatief kortdurend en kenmerkten zich door een duidelijke focus, namelijk het verbeteren van de algemene kwaliteit, de interactie tussen pedagogisch medewerker en kinderen en het sociaal-emotioneel welbevinden van de kinderen (aspecten van de proximale kwaliteit). De gecombineerde effectgrootte van deze studies was middelgroot (Hedges' g = 0.35) en vrij robuust, wat aangeeft dat dit type programma ingezet kan worden om de kwaliteit van de kinderopvang op een efficiënte manier te verbeteren. De trainingen waren het meest effectief in het verbeteren van de interactievaardigheden van de pedagogisch medewerkers en in mindere mate in het verbeteren van de algemene kwaliteit van de groep en het sociaal-emotioneel welbevinden van de kinderen. Uit onze moderatoranalyse bleek dat effectiviteit van de vaardigheidstrainingen niet verschilde voor type opvang (kinderopvang versus gastouderopvang). Programma's waren effectiever als er in de studie gebruik werd gemaakt van een placebotraining

voor de controlegroep vergeleken met studies die geen placebotraining hadden ingezet. Een mogelijke verklaring hiervoor is dat studies met een placeboprogramma over het algemeen van hogere kwaliteit zijn met betrouwbaardere meetinstrumenten en een meer solide opzet. Verder is een voorzichtige conclusie dat trainingen met een individuele trainingscomponent effectiever zijn dan trainingen met alleen groepssessies. Deze bevinding zal verder bevestigd moeten worden in toekomstige experimenten en meta-analyses. Het is aannemelijk dat vaardigheidstrainingen met een individuele component effectiever zijn dan trainingen die alleen in groepsverband plaats vinden, omdat deelnemers tijdens individuele sessies in het bijzonder reflecteren op hun eigen houding en gedrag. Opmerkelijk was dat voor veel studies gold dat essentiële details van de studie-opzet en de vaardigheidstrainingen voor pedagogisch medewerkers slecht gerapporteerd waren en dat de meeste trainingen zich uitsluitend richtten op kinderen en kinderdagverblijven in achterstandssituaties. Vanuit wetenschappelijk oogpunt is er behoefte aan meer gerandomiseerde experimenten naar vaardigheidstrainingen in alle vormen van kinderopvang.

Effectiviteit van de VIPP-Child Care in kinderdagverblijven

In hoofdstuk 3 staat de effectiviteitsstudie naar de aangepaste versie van de *Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline* (VIPP-SD; Juffer, Bakermans-Kranenburg, & Van IJzendoorn, 2008) in kinderdagverblijven centraal. De studie is uitgevoerd in een gerandomiseerde opzet met voor- en nameting bij een interventiegroep (n = 34) en een controle groep (n = 30). De VIPP-SD integreert op een unieke manier twee theorieën: de gehechtheidstheorie (Ainsworth, Bell & Stayton, 1974; Bowlby, 1969) en de 'coercion' theorie van Patterson (1982), die beschrijft hoe negatieve interactiepatronen tussen ouders en kind in stand gehouden worden door ineffectieve manieren van straffen en belonen. De VIPP-SD methode heeft als doel de sensitiviteit van ouders te verhogen en het sensitief grenzen stellen te bevorderen.

Uit diverse effectiviteitsstudies in gezinnen is gebleken dat ouders die deze training hadden gevolgd na afloop sensitiever waren (Juffer, Bakermans-Kranenburg, & Van IJzendoorn, 2008). De training is in een recent onderzoek aangepast voor gebruik in de gastouderopvang en deze nieuwe variant, VIPP-Child Care (VIPP-CC) bleek effectief in het verbeteren van de algemene kwaliteit van de gastouderopvang en de opvattingen van gastouders ten opzichte van sensitieve opvoeding (Groeneveld, Vermeer, Van IJzendoorn, & Linting, 2011). De VIPP-CC is voor de huidige studie verder aangepast voor gebruik in kinderdagverblijven. De methode bestaat uit zes tweewekelijkse trainingsbezoeken van ongeveer anderhalf uur. Tijdens ieder bezoek werd de pedagogisch medewerker gefilmd in verschillende speel- of verzorgingssituaties met vier kinderen op de groep. Na afloop van het filmen werden de opnames van het vorige bezoek (één of twee weken eerder) besproken. Er werd gewerkt vanuit een gestandaardiseerde handleiding door opgeleide trainers.

Uit onze studie blijkt dat de VIPP-CC effectief is in kinderdagverblijven. Pedagogisch medewerkers in de interventiegroep scoorden na afloop hoger op geobserveerde sensitiviteit ten opzichte van de voormeting dan pedagogisch medewerkers in de controlegroep. Dit bleek vooral zo te zijn in voorgestructureerde situaties (boekje lezen, puzzel maken en puzzel opruimen) met een kleine groep kinderen. Mogelijk worden de hogere scores tijdens de voorgestructureerde taakjes veroorzaakt door een gelijkenis met de trainingssituaties en een kleinere groep kinderen, waarbij het makkelijker lijkt om geleerde vaardigheden en inzichten toe te passen dan in een grote groep. Het is bovendien over het algemeen makkelijker om signalen van kinderen te zien en er op in te spelen in kleinere groepen (Ahnert et al., 2006). De VIPP-CC is ook effectief gebleken wat betreft het verbeteren van de opvattingen over sensitief opvoeden. Pedagogisch medewerkers in de interventiegroep waren het na afloop van de training meer eens met de stellingen die uitgaan van sensitieve benadering van de opvoeding dan pedagogisch medewerkers in de controlegroep. Het is aannemelijk dat er eerst een verandering in opvattingen plaatsvindt, voordat er een verandering in gedrag plaatsvindt (Susman-Stillman, Pleuss, & Englund, 2013). Dit zou kunnen betekenen dat de pedagogisch medewerkers in de interventiegroep de verbeterde sensitiviteit verder zullen generaliseren in hun gedrag (Tziner, Haccoun, & Kadish, 1991). Uit de evaluaties bleek dat de VIPP-CC door de meeste deelnemers werd ervaren als leuk, informatief, interessant en nuttig. Er werd aangegeven dat de training vooral bruikbaar zou kunnen zijn voor pedagogisch medewerkers met weinig ervaring of voor de opleiding tot pedagogisch medewerker in het middelbaar beroepsonderwijs. Er is geen effect gevonden van de VIPP-CC op de algemene kwaliteit van kinderopvang, zoals vastgesteld met een wereldwijd gebruikt meetinstrument waarbij zowel distale als proximale kenmerken worden geobserveerd. Dit resultaat is niet geheel onverwacht, want op de meeste van de met dit instrument gemeten aspecten, zoals de inrichting van de ruimte, hygiëne beleid, dagelijkse routines, en speelmateriaal is in de training geen nadruk gelegd. Bovendien zijn het meestal niet de pedagogisch medewerkers die beslissen over deze aspecten van algemene kwaliteit, maar managers of beleidsmakers. De gemiddelde algemene kwaliteit voor alle dagverblijven gezamenlijk was voldoende.

Geluid als kwaliteitsindicator in de kinderopvang

In de correlationele studie beschreven in hoofdstuk 4 wordt de theorie van omgevingschaos onderzocht in de kinderopvang. Het emotioneel welbevinden van de kinderen werd onderzocht in samenhang met de traditionele indicatoren van kwaliteit van opvang zoals beschreven in het model van Riksen-Walraven (2004) en de relatief nieuwe indicatoren voor kwaliteit uit de omgevingschaostheorie van Evans en Wachs (2010). Traditionele indicatoren van kwaliteit zijn groepsgrootte, de aanwezigheid en het gebruik van ontwikkelings- en speelmateriaal en interacties tussen pedagogisch

medewerker en kinderen. De belangrijkste indicatoren uit de omgevingschaostheorie die we getoetst hebben zijn geluidsniveaus en geluidsvariabiliteit, door middel van digitale geluidsmeters Onze resultaten laten zien dat geluidsniveaus en geluidsvariabiliteit op een niet-lineaire manier sterk geassocieerd waren met emotioneel welbevinden van de kinderen (respectievelijk *Bèta* = 0.36 en *Bèta* = 0.33). De geluidsniveaus in de kinderopvangcentra waren gemiddeld 61.69 decibel (*SD* = 3.30, uiteenlopend van 54.95 tot 70.83 decibel) wat significant hoger is (*d* = 1.53) dan de geluidsniveaus die eerder in de gastouderopvang zijn gemeten (*M* = 56.54 decibel, *SD* = 3.43, uiteenlopend van 48.45 tot 64.56 decibel) (Linting, Groeneveld, Vermeer, & Van IJzendoorn, 2013). Geluidsvariabiliteit, of wel de uitschieters in geluid boven en onder het gemiddelde, verschilde niet significant tussen gastouderopvang en kinderopvang (*d* = - 0.51). Voor kinderdagverblijven was de variabiliteit gemiddeld 7.39 decibel (*SD* = 0.81, uiteenlopend van 6.10 tot 9.68 decibel), voor gastouderopvang was het gemiddelde 7.85 decibel (*SD* = 1.00, uiteenlopend van 4.44 tot 10.66 decibel). Dit betekent dat kinderen in dagverblijven niet alleen worden blootgesteld aan hogere gemiddelde (constante) geluidsniveaus, maar ook dat de absolute bovengrens veel hoger ligt dan voor kinderen in gastouderopvang. Geluidsniveaus waren gemiddeld hoger bij grotere groepen en als er sprake was van buitenactiviteiten. Bovendien hebben we een drempel kunnen vaststellen voor geluidsniveaus, vergelijkbaar met bevindingen van Linting en collega's (2013). Uit onze studie bleek dat het welbevinden minder hoog was bij zowel zeer lage als zeer hoge geluidsniveaus en geluidsvariabiliteit. Dit impliceert dat het niet bevorderlijk is voor het welbevinden van jonge kinderen om veel rumoer te ervaren, maar dat een te stille omgeving ook niet optimaal is. Dit kan betekenen dat er voor mate van stimulering een ondergrens en een bovengrens bestaat. Als er te weinig activiteit is en de kinderen in de groep te weinig gestimuleerd worden laten zij wellicht tekenen van verveling en onbehagen zien. Aan de andere kant kunnen te hoge stimuleringsniveaus met veel activiteit ervoor zorgen dat de kinderen niet goed met alle prikkels om kunnen gaan en daardoor een lager welbevinden laten zien.

Conclusie

Omdat een groot aantal kinderen gebruik maakt van kinderopvang en lage kwaliteit langdurige, negatieve gevolgen voor de ontwikkeling kan hebben, is het waarborgen van deze kwaliteit essentieel. In dit proefschrift hebben we laten zien dat bij het toetsen van kwaliteit niet alleen de gebruikelijke indicatoren maar ook indicatoren uit de theorie van omgevingschaos en daarbij geluid in het bijzonder, in acht zouden moeten worden genomen. Geluidsniveaus en geluidsvariabiliteit zijn belangrijke voorspellers voor het welbevinden van kinderen in de kinderopvang, waarbij zowel zeer lage als zeer hoge geluidsniveaus nadelig zijn. Verder is gebleken dat kortdurende, vaardigheidstrainingen voor pedagogisch medewerkers in de kinderopvang effectief zijn en dat zij zouden

kunnen worden ingezet om de interacties tussen pedagogisch medewerker en kinderen en de sociaal-emotionele ontwikkeling van kinderen te verbeteren. In het bijzonder is gebleken dat de VIPP- Child Care, een kortdurende video-feedback interventie die gebaseerd is op de gehechtheidstheorie en de theorie van sociaal leren, effectief is in het verhogen van sensitiviteit van pedagogisch medewerkers in de kinderopvang.

Dankwoord

Deze bladzijde wil ik graag gebruiken om een aantal mensen te bedanken. Ten eerste veel dank aan de managers van de dagverblijven, de pedagogisch medewerkers en hun collega's, de kinderen en hun ouders die hebben meegewerkt aan het onderzoek. De deelname aan onderzoek naar de kwaliteit van de kinderopvang vergt in tijden waarin de sector onder druk staat veel goede moed en inzet. Ik wil graag alle 18 studenten die hebben meegewerkt aan het project bedanken, zonder wie het onderzoek niet mogelijk was geweest. Jullie hebben honderden kilometers afgelegd en veel uren gemaakt om alle metingen en trainingsbezoeken uit te voeren, veel dank daarvoor. Ik ben verder dankbaar voor de steun, het vertrouwen en de interesse van al mijn collega's en de prettige vorm van samenwerken. In het bijzonder wil ik bedanken: Madelon, Rianne, Anne, Marije, Mariëlle, Sandra, Sophie, Estelle, Marieke, Christie, Senguel, Anja; bedankt voor jullie gezelligheid, humor en optimisme; Rodrigo gracias por tu amistad y por llevar un poco de Latinamerica a Leiden, lo necitamos! Gea en Esther, bedankt voor alle fijne ondersteuning. Mijn familie wil ik bedanken voor de steun en de interesse voor mijn onderzoeksproject. Lieve vriendinnetjes: bedankt voor alles wat jullie voor mij betekenen. De afgelopen vijf jaar hebben een onvergetelijke periode gevormd. Bedankt voor jullie vriendschap in mooie én moeilijke tijden. Lieve Mir, mijn grote zus en mijn geweten, dankjewel voor wie je bent. Mir en Joep, jullie zijn onmisbaar voor mij; fijn dat jullie mij steunden tijdens dat vreemde 'empirische' onderzoek. Kleine lieve Siri, dankjewel dat ik jouw tante mag zijn en zo dicht bij jou mag staan. Jij leert mij alles opnieuw. Lieve mama, de waarde van je eigen ouders leer je pas waarderen als je denkt klaar te zijn met opgroeien. Dankjewel voor alles dat je mij hebt meegegeven en voor wat ik bedoeld of onbedoeld van jou en papa meegekregen heb. Lieve papa, dit proefschrift is voor jou. Ik weet hoe trots je op mij zou zijn. Liefste René, ik ben zo blij met jou en met wie jij bent, dankjewel!

Curriculum Vitae

Claudia Werner werd geboren op 21 juli 1982 te Leiderdorp. In 2000 behaalde zij haar gymnasiumdiploma aan het Stedelijk Gymnasium te Leiden. Aansluitend verbleef Claudia een jaar in Veracruz, Mexico, in het kader van het Rotary Youth Exchange Program. In 2001 startte zij met de opleiding Psychologie aan de Universiteit Utrecht die zij in 2007 afrondde met als doctoraalspecialisatie Ontwikkelingspsychologie. Tijdens haar studie deed zij werkervaring op als ambulant begeleider voor kinderen met ontwikkelingsstoornissen. Claudia schreef haar afstudeerscriptie aan de Universiteit van Glasgow, Schotland, over de relatie tussen hulpgerelateerd gedrag en gehechtheidsclassificaties bij kinderen met een reactieve hechtingsstoornis. Ook was zij daar betrokken bij de uitvoering van de *Mellow Parenting* interventie voor deze doelgroep. Claudia voerde aansluitend in Nederland haar klinische stage uit bij De Jutters, centrum voor kinder- en jeugdpsychiatrie in Den Haag, waar zij haar basisaantekening diagnostiek behaalde. Na haar afstuderen werkte zij twee jaar bij De Jutters als psycholoog in het 0-4 team en voor het Centrum voor Jeugd en Gezin. In juni 2009 begon Claudia als promovenda aan de Universiteit Leiden op de afdeling Algemene en Gezinspedagogiek, waar zij onderzoek deed naar de effectiviteit van interventies voor pedagogisch medewerkers in de kinderopvang. De resultaten van haar onderzoek zijn beschreven in dit proefschrift. Claudia is verder intensief betrokken geweest bij het onderzoeksproject naar stressregulatie bij moeders met tekortschietend ouderschap dat werd uitgevoerd in samenwerking met gezinspsychiatrisch centrum Yulius te Barendrecht. Naast haar aanstelling als promovenda heeft Claudia één dag per week gewerkt als docent bij de afdeling Algemene en Gezinspedagogiek. Momenteel is Claudia vanuit deze afdeling werkzaam als VIPP-SD trainer en als docent bij de nieuwe Masterspecialisatie Forensische Gezinspedagogiek.