Effects of preschool education in mixed and targeted classrooms

Annika de Haan

Cover Esther Ris, Proefschriftomslag.nl
Layout Renate Siebes, Proefschrift.nu
Printed by Ridderprint, Ridderkerk
ISBN 978-90-393-6428-4

Effects of preschool education in mixed and targeted classrooms

Effecten van voor- en vroegschoolse educatie in gemengde groepen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht op gezag van de rector magnificus,
prof.dr. G.J. van der Zwaan, ingevolge het besluit van het
college voor promoties in het openbaar te verdedigen
op woensdag 28 oktober 2015 des middags te 2.30 uur

door

Annika Klasiena Engela de Haan

geboren op 30 maart 1982
te Assen

Promotoren: Prof. dr. P.P.M. Leseman
Prof. dr. E.P.J.M. Elbers

This thesis was accomplished with financial support from the Municipality of Utrecht.

Contents

Chapter 1

General introduction

Aim of this dissertation

One of the remaining questions in the field of early childhood education is whether efforts aiming to promote the school readiness skills of young learners should be targeted specifically at children at risk for delays in these basic skills or whether efforts should be aimed at all young children (e.g., Barnett, 2011). In the city of Utrecht, The Netherlands, an experiment was conducted in which preschool education was deliberately aimed at all children, regardless of background. The present dissertation concerns an evaluation of this experiment and reports the effects of two characteristics of early childhood education on participating children's development in emergent academic skills. First, the implementation of an education program, which yielded equivocal results in earlier studies in The Netherlands. Second, the classroom composition with respect to disadvantaged children (targeted vs. mixed classrooms). This dissertation expands the research on early childhood education by taking actual child experiences into account in examining the effects of an education program and the classroom composition on children's development.

The achievement gap of disadvantaged children

The achievement gap in language, emergent literacy, and math skills between children raised in low socioeconomic status (SES) or immigrant families and their peers from other backgrounds is already apparent at the start of formal schooling (Jordan & Levine, 2009; Korat, 2005; Magnuson & Waldfogel, 2005; Smith & Dixon, 1995). In The Netherlands this disadvantaged group of children concerns about 13 percent of all children in primary school (Inspectie van het Onderwijs, 2013). Since these children already suffer from educational disadvantages at the start of formal schooling and given the importance of emergent academic skills for later school success (for a review, see La Paro & Pianta, 2000), early education programs aiming to foster disadvantaged children's development in these skills have been, as in most industrialized countries, implemented in The Netherlands (OECD, 2006). Evidence of positive short- and long-term effects of such programs, based on several studies, is accumulating (Anderson et al., 2003; Barnett, 1995; Blok, Fukkink, Gebhardt, & Leseman, 2005; Burger, 2010; Campbell, Pungello, Miller-Johnson, Burchinal, & Ramey, 2001; Gorey, 2001; Leseman, 2009; Melhuish, 2004; Nelson, Westhues, & MacLeod, 2003; Yoshikawa et al., 2013). In The Netherlands family background factors such as parental educational attainment level and parental cultural background are taken into account to determine eligibility of children for participation in early education programs for 2- to 4-year-olds. Early education programs are continued in kindergarten, an integral part of primary school, that starts at age four and lasts two to two-and-a-half years. The group of

disadvantaged children is diverse and consists of native Dutch children from low educated families; children from migrants from the former Dutch colonies; children from so-called 'guest workers' from the Mediterranean countries that came to The Netherlands in the 1960s (e.g., Italy, Spain, Portugal, Greece, Yugoslavia, Turkey and Morocco); and children from refugee families from diverse areas of conflict. Children stemming from the former group of 'guest workers', especially of Moroccan and Turkish origin, are most prevalent (about 35 percent) among the group of disadvantaged children (Driessen, 2000).

Driessen and Merry (2014) recently published a review and cross-cohort analyses of three comprehensive cohort studies conducted in The Netherlands in the period 1988–2008 (LEO, PRIMA, and COOL). With regard to *starting differences* in the second year of kindergarten (at age six) their conclusion was that, although initial differences were somewhat reduced over the years, especially for the category of heavily disadvantaged minority children (at least one parent attended no schooling after primary school), the gap between disadvantaged children and their peers is still substantial. With regard to the *development* of children over the course of primary school, Driessen and Merry (2014) concluded that, although in the first cohorts hardly any decline in the achievement gap was observed, later cohorts showed that the gap in language achievement, although still substantial, has decreased and that differences in math achievement nearly disappeared at the end of primary school (at age 12).

In sum, although disadvantaged minority children continue to have the lowest language scores in the last year of primary school, differences between these children and their non-disadvantaged peers were reduced over the years, which suggests that the achievement gap is closing. The National Bureau of Statistics also concluded in the last comprehensive national report on the national end of primary school tests, used by about 80 percent of schools, that the gap is closing (Centraal Bureau voor de Statistiek, 2012). However, although disadvantaged minority children's achievements are improving, there is still a long way to go, since the group of disadvantaged minority children still lags behind native Dutch children from low SES families.

The start of early childhood education programs from an international perspective

Although in the 1960s some children from middle and high SES families attended some sort of preschool, formal schooling for young children was uncommon in most countries. The majority of children stayed at home until they started in school, which starts in The Netherlands since the 1950s in kindergarten at four years of age. In the United States however,

the majority of children did not receive any schooling until first grade when they were six or seven years old. In the 1960s the potential of compensatory education programs in combating educational disadvantages became widely acknowledged in many industrialized countries. In the United States, for example, many initiatives were employed under the flag of the federal Head Start program. Initially, the Head Start program conceived as a catch-up summer program of a few weeks offered to disadvantaged preschool aged children before they entered primary school. Soon, the Head Start program expanded to center-based education and care. Since the start, Head Start has acknowledged the importance of parents for child development. Therefore, the Head Start program included activities explicitly aiming at parents' role in fostering their child's development. Head Start programs still encompass a set of comprehensive services (education, health, nutrition, and social services) all aiming to diminish the deteriorating effects of living in adverse circumstances, but there is large variety in the available services. The Head Start program has certain performance standards that govern the services every Head Start center must provide, but communities may adapt the content and implemented program to local needs, desires and resources. In the late 1980s, when the importance of early education for later school success was rediscovered, the number of children served by Head Start showed a steep increase (Zigler, Gilliam, & Jones, 2006b). Today, approximately 42 percent of eligible children are served (Schmit, November 6, 2014).

Next to Head Start programs several model programs were developed by universities and charitable organizations in the 1960s. Famous examples are the High Scope/Perry Preschool project, the Abecedarian project, and the Chicago Child-Parent Centers.

Research on Head Start and model programs

The main finding of a first comprehensive review of Head Start studies was that participation in a Head Start program had sizeable effects on children's cognitive and social-emotional development after one year of attendance (McKey et al., 1985). The Head Start Impact Study (U.S. Department of Health and Human Services, Administration for Children and Families, 2010), a large scale randomized control study, showed that both 3- and 4-year-old Head Start children showed better abilities in a range of developmental domains than children in the control group after nine months of Head Start participation. Another large scale study, the Head Start Family and Child Experiences Survey (FACES), found similar results; although Head Start children lagged behind peers at the start, they progressed at comparable or even better than average rates during the program year and, therefore, could partially compensate the initial delay (especially in letter-word knowledge) (Aikens et al., 2010; Aikens, Kopack Klein, Tarullo, & West, 2013; Malone, Hulsey, Aikens, West, & Tarullo, 2010; Zill et al., 2003). Both the Head Start Impact Study and FACES found that Head Start participation had greater

impact on children who started as 3-year-olds compared to children who started at age four. In addition, both studies showed that children who started at age three showed better social skills and fewer problem behavior at program exit than children who started at age four.

The model programs developed by universities and charitable organizations, such as the High Scope/Perry Preschool project, the Abecedarian project, and the Chicago Child-Parent Centers were evaluated with strong research designs. All programs showed short- and long-term benefits on the cognitive and social-emotional development of participating children from low income families (Barnett, 1993; Campbell & Ramey, 1994; Campbell et al., 2001; Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002; Ramey & Campbell, 1984; Ramey & Ramey, 2004; Reynolds, Temple, Robertson, & Mann, 2001; Reynolds, Temple, Ou, Arteaga, & White, 2011).

Several meta-analyses including both model and large-scale programs such as Head Start showed positive short- and long-term results on child development (Anderson et al., 2003; Barnett, 1995; Barnett, 1998; Blok et al., 2005; Burger, 2010; Camilli, Vargas, Ryan, & Barnett, 2010; Gorey, 2001; Nelson et al., 2003; Royce, Darlington, & Murray, 1983). Generally, results of model programs are more positive than results of Head Start, especially on the long-term (including fewer delinquency, reduced rates of grade retention, reduced enrollment in special education and higher economic independency) (see, for example, Barnett, 1995; Reynolds & Temple, 2008). One explanation is that model programs had larger budgets, better paid and higher educated teachers and more favorable teacher to child ratios.

The target population of early education programs

Since the importance of early learning is widely agreed upon, one of the remaining questions facing the field of early childhood education today is whether efforts should be targeted especially at children at risk for delays in school readiness skills or whether all young children should be able to profit from early childhood education (e.g., Barnett, 2011). Over the last decade an increase in universal preschool programs for 4- or even 3-year-olds, deliberately aimed at all children, regardless of their ethnic or socioeconomic background, is observed in the United States. One of the main reasons dividing proponents of either a targeted or universal approach are the higher costs associated with universal provisions (Barnett, 2011; Rolnick & Grunewald, 2011). Generally, it seems most cost-effective to target early childhood education programs exclusively at disadvantaged children, because then resources are concentrated on fewer children (for a review, see Rolnick & Grunewald, 2011).

Presumably, the most important advantage of universal programs over targeted programs is that children at risk for developmental delays can profit from attendance of

higher skilled peers. The presence of higher skilled peers may have positive effects on the language learning environment in mixed classrooms through three pathways. First, so-called direct peer effects concern the language skills of peers as an important resource for children's language learning; most of children's language input in preschools comes from peers (Dickinson & Tabors, 2001). Given the importance of language input for language development (e.g., Hoff & Naigles, 2002; Hoff, 2003; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Naigles & Hoff-Ginsberg, 1998; Zucker, Solari, Landry, & Swank, 2013), the language learning environment in mixed preschools serving both disadvantaged children and their non-disadvantaged peers might be more beneficial compared to targeted preschools. Second, a heterogeneous mix of children in terms of academic skills may influence teacher practices, both in terms of the quality of interactions and the provision of certain types of learning experiences for children. A third pathway mentioned in the literature, not taken into account in the current study, concerns parental involvement, which is generally likely to be stronger in mixed classrooms (see, for example, Reid & Ready, 2013; Shager, 2012).

One of the central themes of this dissertation concerns the effects of mixed, or universal, preschool and kindergarten classrooms on disadvantaged children's development in emergent academic skills. Regarding these effects, there is some evidence that disadvantaged children's cognitive development is positively affected by attendance of a universal preschool classroom (see, for example, Fitzpatrick, 2008; Gormley Jr., Gayer, Phillips, & Dawson, 2005; Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2004). However, few studies compared disadvantaged children's development in universal preschool with their peers' development in targeted provisions. A first exception is a study by Henry, Gordon, and Rickman (2006), which found that disadvantaged children showed better abilities on a range of school readiness skills after a year in universal preschool than their peers in targeted Head Start programs. Another study however, found only one difference between children in universal and targeted preschool; disadvantaged children in universal preschool showed better expressive language abilities (Dotterer, Burchinal, Bryant, Early, & Pianta, 2013). Other evidence of peer effects comes from studies into effects of the composition of the classroom with regard to 1) socioeconomic and ethnic-cultural background (Biedinger, Becker, & Rohling, 2008; Dotterer et al., 2013; Henry et al., 2006; Henry & Rickman, 2007; Lee, Loeb, & Lubeck, 1998; Mayo & Leseman, 2008; Reid & Ready, 2013; Schechter & Bye, 2007); 2) age (Bell, Greenfield, & Bulotsky-Shearer, 2013; Guo, Tompkins, Justice, & Petscher, 2014; Moller, Forbes-Jones, & Hightower, 2008); and 3) actual peer ability levels (Henry & Rickman, 2007; Justice, Petscher, Schatschneider, & Mashburn, 2011; Lee et al., 1998; Mashburn, Justice, Downer, & Pianta, 2009; Shager, 2012). These studies showed different, though mainly positive, effects of mixed classrooms on children's development.

More evidence on the effects of universal as opposed to targeted preschool provisions is needed. In this dissertation classroom observations and in-depth interaction analyses are used to gain more insight into processes underlying the effects of universal and targeted provisions.

The importance of quality

Initially, the focus in research on early childhood education programs was mainly on effectiveness: does or doesn't the program show positive impacts on participating children's development (Shager, 2012). Over the years attention has shifted to factors that might moderate effectiveness, such as the quality of the program. By now, the importance of quality is well documented; children who attended high quality education and care showed better outcomes than peers in lower quality programs (Bryant, Burchinal, Lau, & Sparling, 1994; Burchinal, Peisner-Feinberg, Bryant, & Clifford, 2000; Burchinal & Cryer, 2003; Howes et al., 2008; Mashburn et al., 2008; Peisner-Feinberg et al., 2001; Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2011).

Structural and process quality are the two often distinguished dimensions of quality (see, for example, Howes et al., 2008). The former concerns aspects such as group size, adult to child ratio and teacher's education, characteristics that are mostly determined by policy regulations. These structural characteristics are presumed to exert their influence on child development indirectly through process quality which refers to the daily interaction processes in the classroom and includes the social-emotional and instructional support provided by teachers (Howes et al., 2008). However, since the association between these aspects of classroom quality and child outcomes is generally weak and not consistent across studies and contexts (Burchinal et al., 2008 as cited in Chien et al., 2010; Early et al., 2010; Justice, Mashburn, Hamre, & Pianta, 2008b; Slot, Mulder, Verhagen, & Leseman, 2014), other characteristics of child experiences in the classroom such as the type of activities children are engaged in and the role the teacher plays during these activities, may be more important in predicting child outcomes (see also Chien et al., 2010; Early et al., 2010). In few studies however, actual classroom activities were related to child outcomes. This is surprising given the long-standing debate on the best learning approach for young children centering around the question ". . . where along the continuum of 'child-centered' to 'teacher-directed' it is best to define the role of the teacher for optimizing children's [development] . . ." (Winsler & Carlton, 2003, p. 156). In this dissertation actual child experiences in preschool and kindergarten classrooms are related to child development.

Early childhood education programs in The Netherlands

As mentioned before, in the second half of the 20th century many so called 'guest workers' from Mediterranean countries, including foremost Morocco and Turkey, immigrated to The Netherlands. At first, the government anticipated return to their home countries, but in the second part of the 1970s and especially the 1980s it became apparent that many immigrants were to stay and reunite with their families in The Netherlands. Because immigrant children often experienced difficulties in education, partly due to another than Dutch home language and low parental educational attainment level, several initiatives were developed and introduced to ameliorate these children's position in education (Driessen, 2000). Over the years the awareness that disadvantages needed to be tackled as early as possible gained ground, which resulted at first mainly in home-based education programs that were often conducted in close collaboration with neighborhood centers and primary schools. An example is the 'Opstap' project, based on the Israeli Home Intervention Program for Preschool Youngsters (HIPPY), which yielded, after initial disappointing results (see Eldering & Vedder, 1999), positive effects for especially Turkish children's cognitive development and emergent math skills (Van Tuijl, Leseman, & Rispens, 2001).

In the 1990s the Dutch government concluded that the efforts were not effective enough; the educational disadvantages of minority children did not diminish over the years. New policy entailed the awareness that efforts should be focused even more than before on preschool activities and more specifically on center-based programs. As a result, several programs were developed and their small-scale implementation was closely evaluated. Positive evaluations (Veen, Roeleveld, & Leseman, 2000) led in the year 2000 to the 'Regeling Voor- en Vroegschoolse Educatie' [Act Early Childhood Education] by the Ministry of Education and Science. This act marked the start of large-scale implementation of early education programs in The Netherlands. The act entailed an important precondition considering the intensity of programs; early childhood education programs should start in preschool and follow through to kindergarten, thereby spanning the age range of 2;6 (in some municipalities 2;0 years) to about 6 years of age.

The Dutch early childhood education and care system

In The Netherlands there are different provisions of early education and care for children under and over four years of age. Although kindergarten is compulsory from five years of age onwards, children are entitled to enter from their fourth birthday, which the vast majority (98%) of children does. Kindergarten lasts two to two-and-a-half years and all children attend about 25 hours per week divided over five days. The kindergarten department is free of charge

and is integrated in the primary school system since 1985. Since all children, regardless of background, are eligible, the Dutch kindergarten system is an example of a universal early childhood education provision. However, while the system is universal in nature, especially in neighborhoods with a dense population of minority and low socioeconomic status families, some schools are attended almost exclusively by disadvantaged children.

The landscape of center-based provisions for children under four years of age is scattered. Since the 1960s there are two types of center-based provisions: center-based day care ('kinderdagverblijf') and play groups/preschools ('peuterspeelzaal/voorschool'). Play groups originated in the 1960s and were often established by parents who deemed it important for their children to play with other children. Over time the emphasis on educational activities has grown and play groups became increasingly acknowledged as a means to foster disadvantaged children's (language) development and school readiness skills. In the first decade of the 21st century – since the 'Regeling Voor- en Vroegschoolse Educatie' [Act Early Childhood Education] – this has led, especially in cities and neighborhoods with a high concentration of disadvantaged children to the implementation of 'educational' preschools with targeted programs. The main differences between preschools and play groups concern the implementation of an education program, more frequent attendance of children (at least ten hours, generally over four half days as opposed to two half days), two instead of one teacher on a group of maximum 16 children and an obligatory intensive collaboration with a nearby primary school. Preschools are often targeted exclusively at disadvantaged children, indicated by family background factors such as parental educational attainment level and parental cultural background. Generally, children start preschool between the age of 2 to 2;6. The scattered landscape of early childhood education and care for 0- to 4-year-olds creates a separation, because preschool is mostly used by low SES families, whereas children from middle and high SES families often visit day care. This separation is related to the higher costs of day care, but also to the fact that employed parents generally need whole day care, which is only provided by day care provisions (preschool only offers half day care). In 2013 about 39 percent (284,000 children) of 0- to 4-year-olds attended day care, whereas about 27 percent (99,000 children) of 2- and 3-year-olds attended preschool (Buitenhek Management & Consult, 2014; Centraal Bureau voor de Statistiek, 2014).

Programs in early childhood education

The four most used comprehensive programs in The Netherlands have in common that they address holistic development of the child but emphasize language development (see Table 1.1 for a national overview (Jepma, Rutten, & Beekhoven)). In addition, all programs emphasize the importance of an alternation of playful teacher-managed and free child-

Table 1.1 National overview of implemented center-based early childhood education programs in The Netherlands

Type of program	Program	Percentage of centers
Comprehensive programs	Piramide	33
	Startblokken/Basisontwikkeling	16
	Ko Totaal	21
	Kaleidoscoop	7
	Speelplezier	2
	Local variations	8
Domain-specific programs	Doe meer met Bas	8
	Taallijn VVE	3
	Boekenpret	3
Other programs (local adaptations)		8
Unknown		1

Note. Percentages do not add up to 100 due to the given that these are weighed percentages, rounding differences, and double use of programs (about five percent of provisions used two programs) (Jepma et al., 2009).

managed activities. In addition to these comprehensive programs, domain-specific 'programs' focusing on a specific area of development are available. These domain-specific programs are additions to comprehensive programs (Jepma et al., 2009). All programs are approved by the National Youth Institute (Nederlands Jeugd Instituut, NJI) as a theoretically sound intervention, which means that these programs were judged as potentially effective, but research unequivocally demonstrating positive effects is still lacking.

With regard to the four most used comprehensive programs, the main difference between Piramide and Ko Totaal on the one hand and Startblokken/Basisontwikkeling and Kaleidoscoop, the Dutch version of the international High Scope program, on the other hand concerns the availability of a manual. The former programs offer a manual including detailed pre-designed activities, organized around themes and 'learning standards', for example the specific words children should learn during a theme. Startblokken/Basisontwikkeling and Kaleidoscoop do not provide pre-designed themes or activities, but rather a framework of child development to which activities should be aligned. For an extensive description of different education programs in The Netherlands, Dutch readers are referred to Nap-Kolhoff et al. (2008).

Research in The Netherlands

The implementation of three education programs, Piramide, Kaleidoscoop and Startblokken/Basisontwikkeling, was closely monitored in quasi-experimental research (see Veen et al., 2000). The evaluation of Kaleidoscoop and Piramide revealed suboptimal implementation of these programs. Teachers were for example due to frequent staff changes often not certified to work with the program. Nevertheless, the results of the evaluation study were encouraging; Kaleidoscoop and Piramide showed positive, small to medium-sized effects on children's cognitive and language development. The initial promising results for Kaleidoscoop and Piramide were, especially for Piramide sustained in later studies (de Goede & Reezigt, 2001; Veen, Derriks, & Roeleveld, 2002). However, some effects varied across (age) groups, years and outcome measures (de Goede & Reezigt, 2001; Veen et al., 2002).

The evaluation study of Startblokken/Basisontwikkeling showed no positive effects on children's cognitive and language development (to the contrary, a small to medium-sized negative effect was found for language development). Positive effects were found on aspects of children's social-emotional development (Veen, Fukkink, & Roeleveld, 2006).

Although several later evaluation studies of education programs in general (regardless of the specific program that was used), showed some promising results for specific subgroups of children or schools (Nap-Kolhoff et al., 2008; Veen, Roeleveld, & van Daalen, 2008), no unequivocal positive results on child development were found (Karssen, van der Veen, Veen, van Daalen, & Roeleveld, 2013). However, these and other studies (e.g., Bruggers, Driessen, & Gesthuizen, 2014; van Schooten & Sleegers, 2009) suffer from important methodological limitations of which a lack of control group and/or sound baseline measurement of children's skills are among the most important. For example, some of these studies (e.g., Karssen et al., 2013; Nap-Kolhoff et al., 2008) used a non (quasi-)experimental retrospective design. In the Karssen study data on classroom quality were for example gathered two years after children attended preschool. The lack of a baseline measure of children's skills is problematic since a recent study showed that the average skill level in preschools is far below average right at the start in preschool. This may lead to an underestimation of the possible effects of an education program (Slot, 2014).

To conclude, evaluation studies of education programs in The Netherlands report small to medium effects on a limited number of outcome variables. So far, positive results were not unequivocally demonstrated (Driessen, 2012), but it should be noted that many studies suffered from methodological limitations (e.g., Bruggers et al., 2014; van Schooten & Sleegers, 2009). Although in some studies lack of convincing effects might be attributed (in part) to suboptimal implementation, another reason may be found in the global measures that are used, aimed at limited areas of development (Nap-Kolhoff et al., 2008). In addition,

actual child experiences in the classroom were not taken into account in evaluating the effects of education programs. A recent study in The Netherlands however, suggested that the use of an education program was associated with better emotional and educational process quality (Slot, Leseman, Verhagen, & Mulder, 2014), aspects that were in turn to a small extent related to children's development in respectively vocabulary and attention skills (Slot, Mulder et al., 2014). Although the associations between these aspects of process quality and child development were small, these results do suggest that it is important to gain a thorough insight into child experiences in the classroom. However, most evaluation studies did not include measures of actual activities and interactions children experience in the classroom, nor the association between these experiences and child outcomes (in some studies even actual child assessment is lacking and teacher evaluations are used instead (e.g., Karssen et al., 2013)). In short, there is no conclusive evidence on the positive effects of early childhood education in The Netherlands, but many studies reporting null effects lack sound methodology to support this claim.

Early childhood education in Utrecht

In the current study children's development in preschools and kindergartens working with an education program aiming at fostering disadvantaged children's development was examined. The development of disadvantaged children was compared to their peers' development in classrooms without such a program. The education program involved was *Taalrijk* ('Enriched Language'), which is a local adaptation of the program Ko Totaal. Ko Totaal/ Taalrijk is especially suited for both targeted and mixed classrooms due to its structure with an alternation of teacher-managed activities in large and small groups. Teacher-managed activities in small groups enable intensive teacher-child interaction and are considered as an important means in accelerating children's development. Especially in mixed kindergarten classrooms, where often no second teacher is appointed, this enables teachers to provide disadvantaged children in their classrooms with extra attention. As other early childhood education programs, Taalrijk offers a preschool and kindergarten program. The program has a teacher-managed thematic approach; themes last about three to four weeks and the manual includes pre-designed activities with guidelines on organization, instruction language, and words and concepts that have to be incorporated in the activity. Examples of themes are seasons of the year, animals, food, health and fairy tales. Before certification teachers completed a training that encompassed eight half days of training and four coaching on the job sessions over a two year period.

Mixed classrooms

In neighborhoods with a high concentration of disadvantaged children, many play groups in the city of Utrecht were – as in other large cities in The Netherlands – transformed into educational preschools during the first decade of the 21st century. Distinctive features of educational preschools are the presence of two teachers and the implementation of an education program, such as Taalrijk. At first, in the city of Utrecht only disadvantaged children could be enrolled in educational preschools. However, as mentioned before, the accumulation of disadvantaged children has several drawbacks, which may reduce the positive impact of preschool on enrolled children.

In order to overcome these drawbacks, the city of Utrecht deliberately aimed to implement 'mixed' preschools in neighborhoods with a mixed population. Whereas groups in regular preschools were composed of 15 disadvantaged children, groups in mixed preschools were composed of 10 disadvantaged children who attended preschool for four half days (as in regular preschools) and two groups of five non-disadvantaged children who both attended preschool for two half days (as in regular play groups). Because parents of kindergarten children are free in their choice for a primary school, the composition of kindergartens and thus primary schools is dependent mainly on the composition of the neighborhood. Generally, this leads to 'targeted' kindergartens in neighborhoods with a dense population of disadvantaged children and to more mixed kindergartens where the population of the neighborhood is relatively mixed.

The current study

The studies of education programs that are carried out in The Netherlands so far, focused on the mere fact of using an education program. Internationally, many studies however, have shown that implementation fidelity is an important factor in the success of a program. Surprisingly, no study in The Netherlands of which we are aware took into account what actually happens in the classroom; what activities are done and how they are done. This might be more important than the mere reported use of a program (see also Chien et al., 2010; Early et al., 2010). In this dissertation, measures of actual child experiences in the classroom were taken into account in evaluating the effects of the use of an education program on children's development. In addition, these measures were used to answer one of the remaining questions in the field of early childhood education: Should education programs aiming to promote the school readiness skills of young learners be targeted specifically at children at risk for delays in these basic skills or should these efforts be aimed at all young children (e.g., Barnett, 2011)? In short, in the present dissertation measures of actual child experiences

in the preschool and kindergarten classroom were taken into account in evaluating the effects of an education program and the classroom composition on participating children's development in emergent academic skills.

The studies in this dissertation report on three components of our field work. First, children were tested three times over a two year period on a wide range of academic outcomes. Second, intensive classroom observations were done to gain an in-depth insight into children's everyday activities and interactions in preschools and kindergartens. Third, video-observations were conducted, to unravel interaction processes in more detail (see Figure 1.1 for an overview of research activities).

Outline of this dissertation

Chapter 2 reports a study of teacher- and child-managed academic activities. After a description of children's everyday activities in preschool and kindergarten, we will examine the relation between teacher- and child-managed academic activities and children's gains in emergent literacy and math skills over a one year period.

Chapter 3 reports a longitudinal cohort-sequential latent growth modeling approach that was used to determine whether disadvantaged children show enhanced development in emergent academic skills over a two year period in 1) classrooms where the education program Taalrijk was implemented, 2) classrooms where teachers engaged frequently in teacher-managed academic activities, and 3) mixed classrooms. In this study we focus on disadvantaged children, since they are the target group of education programs.

Chapter 4 reports a study focusing on peer interactions. In this chapter we attempt to further disentangle the effects of mixed classrooms, as reported in the previous chapter. We will first provide descriptive statistics on interaction opportunities in relation to classroom contexts. Furthermore, we will determine the proportion of interaction that children have with their teacher respectively with peers, and of peer interaction we will subsequently determine the share of interaction with native Dutch peers and whether this differs between children in mixed and targeted classrooms. Finally, we will determine whether interactions with native Dutch peers are associated with children's gains in emergent academic skills.

Chapter 5 reports a study of video-observations of large and small group teacher-managed activities. After an overview of preschool and kindergarten teachers' types of utterances and the cognitive demands of questions and statements, differences between two contexts in the types of questions and the cognitive demands of the most frequently observed types of utterances are explored. The first context concerns the share of disadvantaged children in the classroom (targeted vs. mixed classrooms) and the second context concerns

2008	2008	2008	2008	2009	2009	2009	2009	2010	2010	2010
Febr/ March	April/ May	Summer	Autumn	Febr/ March	April/ May	Summer	Autumn	Febr/ March	Spring	Summer
Preschool cohort										
Preschool			Preschool				Kindergarten year 1			
3 years			4 years				5 years			
Test 1	Obs		Video	Test 2			Test 3			
Kindergarten cohort										
Year 1			Year 2				Grade 1			
4½ years			5½ years				6½ years			
Test 1	Obs		Video	Test 2			Test 3			

Figure 1.1 An overview of field work that is reported on in this dissertation.
Note. Test = Individual child assessment; Obs = Classroom observations, Video = Video-observations.

the number of children involved in the activity (small vs. large groups). In addition, we will also examine whether children contribute verbally to the activity more frequently during small group activities than during large group activities. Finally, characteristics of teachers' language input are related to children's immediate language output in order to determine the effect of different types of teacher utterances on children's language output.

In *Chapter 6*, the general discussion, the results of the four empirical chapters are discussed.

Chapter 2

Teacher- and child-managed academic activities in preschool and kindergarten and their influence on children's gains in emergent academic skills

de Haan, A. K. E., Elbers, E., & Leseman, P. P. M. (2014). Teacher- and child-managed academic activities in preschool and kindergarten and their influence on children's gains in emergent academic skills. *Journal of Research in Childhood Education*, *28*, 43-58.

Abstract

The aim of this study was to assess whether children's development benefited from teacher- and child-managed academic activities in the preschool and kindergarten classroom. Extensive systematic observations during four half-days in preschool (N = 8) and kindergarten classrooms (N = 8) revealed that classrooms differed in the amount of time spent on teacher-managed academic activities. The time teachers spent on math activities was remarkably low. Teacher-managed activities accelerated children's development, particularly for preschool children (N = 47); they showed larger gains in language, literacy, and math skills if their teacher devoted a relatively large proportion of classroom time to these topics. With regard to kindergarten children (N = 45) only their math skills seemed influenced by their teacher's engagement in academic activities; these children showed larger gains in math skills if their teacher initiated many language and literacy activities. Children's language and literacy development benefited from child-managed language and literacy activities. In contrast, child-managed math activities were not related to children's gains in math skills. The study provides insight into children's daily experiences in the preschool classroom and yields important implications for the professional development of teachers in early childhood education.

Introduction

At the start of formal schooling, children raised in low income or immigrant families often lag behind peers in language, emergent literacy, and math skills (Jordan & Levine, 2009; Korat, 2005; Magnuson & Waldfogel, 2005; Smith & Dixon, 1995). Since these emergent academic skills are strong predictors of later school success (for a review, see La Paro & Pianta, 2000), early education programs aiming to accelerate at-risk children's development in these domains have been implemented in most industrialized countries (OECD, 2006). Generally, review studies showed positive short- and long-term effects of these programs, provided that the intensity and duration are sufficient (Barnett, 1998; Burger, 2010; Gorey, 2001; Leseman, 2009; Nelson et al., 2003; Ramey & Ramey, 2004). In addition to the importance of the intensity and duration of the program, there is a growing body of evidence indicating that children who received care or preschool education of high quality enter school with better developed language, emergent literacy and math skills (Burchinal et al., 2000; Burchinal & Cryer, 2003). In studies assessing the relation between classroom quality and child outcomes, quality is often assessed with measures focusing on the classroom environment and the sensitivity and responsiveness of the teacher. The strength of the association between these aspects of classroom quality and child outcomes in such studies is rather weak (Burchinal et al., 2008 as cited in Chien et al., 2010; Early et al., 2010; Justice et al., 2008b). A possible explanation for the modest size of the association is that quality as such may not be decisive;

the effectiveness of a program may also depend on what activities are done and how they are done (see also Chien et al., 2010; Early et al., 2010). So far, few studies opened 'the black box' of a typical preschool day, let alone the contribution of actual classroom activities to children's development. This study provides a description of children's everyday activities in the public preschool and kindergarten system of The Netherlands, and examines the relationship between teacher- and child-managed academic activities and children's gains in language-literacy and math skills.

The scarcity of research on actual classroom activities and their association with child outcomes is rather surprising considering the long-standing debate on the best learning approach for young children. In short, the debate centers around the question ". . . where along the continuum of 'child-centered' to 'teacher-directed' it is best to define the role of the teacher for optimizing children's [development] . . ." (Winsler & Carlton, 2003, p. 156). Walsh (1989) noted that the growing academic emphasis in preschool and kindergarten, which often goes along with an increase in teacher-directed activities, is driven by pressure on schools to become more effective. The rising expectations of early childhood education as an accelerator of children's development, and thus as a means to diminish the education gap, is accompanied by higher expectations of teachers and children and results in a stronger emphasis on accountability. Consequently, early childhood teachers, at least in the United States, devote increasingly more time to direct instruction of basic literacy and math skills (Graue, 2008; Gullo & Hughes, 2011a; Gullo & Hughes, 2011b; Stipek, Feiler, Daniels, & Milburn, 1995). There is a lack of conclusive evidence regarding the debate (Van Horn, Karlin, Ramey, Aldridge, & Snyder, 2005) and proposals to bridge the controversy have been put forward (Graue, 2008; Neuman, Copple, & Bredekamp, 2000). Although a focus on academic content has been found to be associated with negative outcomes on social-motivational measures and social-emotional quality in some studies (Hyson, 1991; Schweinhart & Weikart, 1988; Schweinhart & Weikart, 1997; Stipek et al., 1995; Stipek et al., 1998), the question has been raised whether this negative association is really inevitable (Neuman et al., 2000; Stipek et al., 1998). New approaches to introducing academic content in early childhood care and education settings, based in the research tradition of emergent literacy and emergent numeracy, have shown that child-following playful and authentic activities in literacy and numeracy centers in the classroom, with allowing initiative to children, can be used effectively to introduce children into academic subjects, without decreasing social-emotional quality (Bodrova, 2008; Bus, Leseman, & Neuman, 2012; Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003; Neuman et al., 2000).

Some studies examined everyday activities of children in the classroom and the teacher's role in shaping these activities (Early et al., 2010; Elicker & Mathur, 1997; Guimarães & McSherry, 2002; Phillips, Gormley Jr., & Lowenstein, 2007; Phillips, Gormley Jr., & Lowenstein,

2009; Tonyan & Howes, 2003). Only in a few studies however, classroom activities were related to child outcomes. For example, in a study by Chien et al. (2010), observations were used to classify children in one of four profiles: free play, individual instruction, group instruction or scaffolded learning. Children in the free play profile, in which children have most free choice time, made the smallest gains in a number of language, literacy and math outcomes. With regard to teacher activities, several studies revealed large differences between early childhood teachers in providing activities as intended, while working within the same program or with the same concept, and having received the same training (Connor, Morrison, & Slominski, 2006; Early et al., 2005; Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006; Meyer, Wardrop, Hastings, & Linn, 1993). It is important to note that these teacher differences are related to child outcomes. For instance, Connor et al. (2006) found that when preschool teachers working within the same comprehensive program spent comparatively more time to language and emergent literacy activities, their students gained more in alphabet knowledge, letter-word recognition, and vocabulary. Especially academic activities during which the teacher was in lead and, to a lesser extent, child-managed academic activities, were positively associated with growth in these emergent school skills. In addition, Justice, Mashburn, Pence, and Wiggins (2008a) found that children participating in a comprehensive language curriculum, with relatively much attention for language learning activities, showed accelerated growth in expressive language skills in pre-kindergarten compared to children in a control group. Similarly, Klibanoff et al. (2006) found that the amount of math-related talk teachers initiated in the classroom, in which large differences were observed, was strongly related to children's growth in pre-mathematical knowledge and skills. In sum, focusing on the teacher, on what she or he does in the classroom, may shed new light on the issue of quality of early childhood education and may suggest new starting points for the implementation and quality monitoring of early education for low income and immigrant children.

Current study

Together, the studies reviewed above indicate that children may benefit from language, literacy and math activities carried out in the preschool and kindergarten classroom. The current study focuses on children's everyday activities in the preschool and kindergarten classroom and the contribution of teacher- and child-managed academic activities on children's academic school readiness. Pre-primary school education in The Netherlands consists of two interconnected systems, spanning the age range of 2;6 to 6 years of age. Children are admitted to preschools from age 2;6. There is a small, income dependent fee and about 60% of the eligible children participates for two to four half days per week (about 2.5 hours per day). Children from low income families and from immigrant families are especially encouraged to participate.

However, preschool is voluntary, parents can also choose to enroll their children at child care centers, with a less clear educational component than in preschool. At age 4, the vast majority of children, over 98%, start in the kindergarten departments of primary school for five days per week (about five hours per day). There is no fee and attendance is compulsory from age 5 onwards. Both preschools and kindergartens implement a basic developmental, child following approach with ample room for free play, fine and gross motor activities, and creative and expressive work. Although it is largely unknown how strong and how effective the academic emphasis is (Doolaard & Leseman, 2008), there is no tradition of didactic teaching or teacher-directed drill and practice of school skills. In view of the controversies regarding early childhood education, we will, in addition to a description of children's everyday activities in the classroom, attempt to answer two main questions in the remainder: Are teacher differences in academic focus related to young children's school readiness? And, second, can child-managed academic activities also contribute to school readiness?

Method

Participants

In total 92 children were involved in the present study, 47 (27 boys; 57%) attending one of eight participating preschools and 45 (22 boys; 49%) attending one of eight participating kindergarten classrooms, all located in a middle-sized city in the western part of The Netherlands. The mean age of preschool children at the first measurement occasion was 3.00 years (*SD* = 0.30). Eleven preschool children were of native Dutch origin (23%), fifteen children had parents of Moroccan origin (32%), five children had parents from Turkish origin (11%), and ten children had parents with several different cultural backgrounds (21%). In addition, six children had parents from mixed cultural backgrounds (13%). Parents' educational level ranged from 1 (no education) to 7 (university degree), the mean was 3.41 (*SD* = 1.73). The mean age of the kindergarten children was 4.36 years (*SD* = 0.24) at the first measurement occasion. Fourteen kindergarten children were of native Dutch origin (31%), fourteen children had parents of Moroccan origin (31%), three children had parents of Turkish origin (7%), and seven children had parents with another cultural background (16%). Furthermore, seven children had one native Dutch parent and one parent with another cultural background (16%). The mean level of parental education was 3.24 (*SD* = 2.11) and ranged from 1 (no education) to 7 (university degree).

Parents of children who recently enrolled in preschool or kindergarten (within the last six months) were informed by a letter and brochure about the study and asked for their

active informed consent resulting in participation of the majority of eligible children (84% of preschoolers and 79% of kindergarten children).

Assessments and measures

Child assessment

All children were tested twice, first shortly after entrance in preschool or kindergarten and one year later. At both measurement occasions, testing was divided over two sessions of approximately 30 minutes. Children were tested individually by trained research assistants, using laptop computers, in a quiet room at preschool or kindergarten.

The Dutch version of the PIPS; Performance Indicators in Primary Schools (Tymms, 2001), a computerized adaptive assessment of a wide range of school readiness skills, was used to assess emergent language and literacy and math skills. Previous research with the PIPS in The Netherlands showed the PIPS to be a culturally fair test, with a good predictive validity and high reliability (Van der Hoeven-van Doornum, 2005).

Language and emergent literacy skills

Language and emergent literacy skills consisted of six subtasks: receptive vocabulary, writing, ideas about reading, letter identification, reading words, and reading sentences. The maximum number of items is 180, but due to the adaptive character of the task (i.e., after a certain number of errors the subtest at hand is automatically ended and the program continues to the next subtest), not all items were administered with all children. At both measurement occasions some missing scores, all missing at random, were imputed separately for the preschool and kindergarten cohort using regression analyses. The percentage of missing scores ranged between 2.2 and 14.9[1].

Emergent math skills

Emergent math skills consisted of seven subtasks: ideas about math, counting, informal sums, digit identification, shapes, adding and subtracting, and advanced sums. The maximum number of items is 69, but again due to the adaptive character of the task not all items were administered with all children. Also for emergent math skills some missing scores were imputed separately for the preschool and kindergarten cohort using regression analyses. With regard to math skills, the percentage of missing scores ranged between 2.2 and 17.01.

1 There are various reasons for missing scores. At the first measurement occasion scores were mainly missing due to unwillingness of children to cooperate or repeated absence of children when the testing session was scheduled. At the second measurement occasion scores were mainly missing due to relocation of the child's family.

Classroom observations

Observations of six randomly selected children in each of the 16 classrooms were conducted to gain a comprehensive insight into children's everyday activities in preschool and kindergarten classrooms. These observation data enabled us to quantify opportunities for engagement in a number of activities deemed relevant for school readiness development. In each of the 16 groups the target children were observed during five half-days of 2.5 to 3 hours, which were scheduled within two weeks during spring. The first half-day served to familiarize the children and the teacher with the presence of the observer and was therefore not included in the analyses. Teachers were instructed to carry out their usual schedule during the observations.

A cyclic interval coding approach was used to observe the target children. These cycles started with a ten seconds observation period of an individual child. After these ten seconds, the activity the observed child was engaged in during that interval was coded on a number of dimensions. After coding all dimensions for a particular interval, the observer started observing the next child for ten seconds. These cycles of observing and coding were repeated from the start until the end of the morning or afternoon. The time for coding was not restricted, but assistants were strongly urged to code as fast as possible. The coding scheme was programmed in E-prime (Schneider, Eschman, & Zuccolotto, 2002) and installed on portable laptop computers. This procedure resulted in preschools in a mean of 209 intervals (*SD* = 32.24) in the mornings and 128 (*SD* = 19.56) observed and coded intervals in the afternoons, which lasted about one hour less than the mornings. In kindergartens, where children spent about one hour longer in the morning and thirty minutes longer in the afternoon than in preschool, a mean of 349 intervals (*SD* = 57.46) were coded in the morning and a mean of 209 intervals (*SD* = 40.63) were observed and coded in the afternoon.

Three trained research assistants and the first author conducted the observations. To assess inter-rater reliability assistants coded written realistic classroom situations which were also programmed in E-prime (Schneider et al., 2002). The coding procedure of the hypothetical situations was identical to the coding procedure in real time classroom situations. Assistants spent approximately 60 minutes coding hypothetical situations, during which they coded on average 35 situations. The inter-rater reliability for the activity aspect of the hypothetical situations was good with all coders' Cohen's kappa over .80.

Activity

After every ten-second observation interval the activity the child was engaged in during that interval was coded. If two activities were observed during one interval, the activity the child was engaged in longest was coded. The activity coding scale was adapted from Howes and Smith (1995) and, after pilot research, designed in such a way that virtually all play

and educational activities carried out in preschool and kindergarten classrooms could be uniquely coded in mutually exclusive categories. Typical examples of activities were making music and singing, creative and craft activities, and free play. Also activities with an academic focus were included, such as book reading, solving geometrical shapes puzzles, working with seriation material, et cetera. In addition, the categories outdoor play (mainly gross-motor activities, but not further differentiated), eating a snack and drinking, and not being engaged in any activity (waiting, transition to the next activity) and a rest category were included (the complete coding scheme is available from the first author upon request). In addition to the type of activity, we also coded the social context of the activity. The social context enabled us to examine who initiated or was mainly in charge of a particular activity; the teachers or the observed children. Composite aggregated scores at the classroom level, combining the type of activity with the initiator (or main agent) of that activity, were computed representing the mean percentages of time observed for teacher-managed language and literacy activities, teacher-managed math activities, child-managed language and literacy activities, and child-managed math activities respectively, based on four half days observation and six target children per classroom. Several observation categories were taken together in these broader categories. The aggregated classroom scores were used as factor in subsequent ANOVAs.

Teacher-managed language and literacy activities

Teacher-managed sharing time and teacher-managed language and literacy activities such as book reading, storytelling, language games (examples are thinking of words or names starting with a certain letter and guessing a word after the teacher's description of certain characterizing features of that word), rhyming, and introducing and explaining new words were coded as teacher-managed language and literacy activities.

Teacher-managed math activities

This category consisted of activities, managed by the teacher, involving counting, ranking, talking about (birth) dates, calendar use, naming shapes, ordering, comparing and estimating quantities.

Child-managed language and literacy activities

Two types of child-managed activities were aggregated into this category: looking at or reading books, and emergent literacy activities such as writing names, playing letter games and functional literacy activities in pretend play such as pretending to read a menu or writing down an order when playing restaurant. Activities could have been planned or initiated by the teacher, but if at the time of observation the teacher was not directly involved, they were considered child-managed.

Child-managed math activities

This category consisted of several activities initiated or mainly controlled by children involving counting, naming shapes, ordering according to colour, size, or shape, pattern completion and other math related skills. Most frequently observed were activities such as solving jig-saw puzzles, playing board games involving numbers and dice, and tasks for which patterns needed to be replicated by children, such as making string beads. Again, although the teachers could have stimulated the children to carry out these activities, if the children were mainly in charge they were considered child-managed.

Results

Children's everyday activities

Observations were conducted in eight preschool and eight kindergarten classrooms. In preschool, the total number of observed intervals ranged from 396 to 783 per classroom (M = 625.88, SD = 124.58). In kindergartens the number of observed intervals ranged from 732 to 1,299 per classroom (M = 1086.63, SD = 171.35). In Table 2.1 an overview of the mean time allocation in preschools and kindergartens is shown in percents of the total observation time. The 'other' category includes primarily activities that did not fit in any of the categories of the coding scheme (preschool: 8%; kindergarten: 13%). In addition, computer work (kindergarten: 13%), gymnastics and dancing (preschool: 2%; kindergarten:

Table 2.1 Time allocation in percentages of the observed intervals in preschools and kindergartens

	Preschool		Kindergarten	
	M % (*SD*)	Range %	*M* % (*SD*)	Range %
Teacher-managed literacy activities	7.22 (5.85)	0.93–16.29	14.96 (5.54)	8.06–26.71
Teacher-managed math activities	0.20 (0.29)	0.00–0.76	1.67 (1.10)	0.00–2.83
Child-managed literacy activities	2.11 (1.33)	0.75–4.36	1.59 (1.60)	0.00–4.37
Child-managed math activities	2.02 (2.33)	0.00–7.03	4.82 (6.03)	1.03–19.32
Creative activities	8.60 (6.05)	0.55–18.69	6.47 (3.51)	0.00–11.09
Free play	15.38 (6.21)	10.12–28.22	6.73 (3.23)	2.22–12.16
Snack time	11.42 (2.72)	5.81–14.80	5.35 (0.77)	3.90–6.45
Outside play	15.68 (9.92)	5.05–27.46	17.60 (7.24)	7.66–26.79
Other activity	20.57 (9.76)	9.58–39.44	22.43 (8.00)	10.24–32.27
No activity/transition	16.21 (7.02)	6.10–24.44	18.05 (8.66)	6.93–31.37

3%), music and singing activities (preschool: 6%; kindergarten: 2%) and other fine motor activities (preschool: 4%; kindergarten: 4%) were aggregated in the rest category.

To determine the effects of teacher- and child-managed academic activities the nested structure of the data had to be taken into account. Applying multilevel regression analysis was considered not feasible due to the small number of observations on the classroom level (n = 16 cf. Maas & Hox, 2005). Therefore, as alternative, within-subjects repeated measures ANOVAs were conducted with the aggregated classroom scores as factor. Main effects of time, although significant in all analyses, were not the main focus of this study and are therefore not reported.

Teacher-managed academic activities

With regard to teacher-managed language-literacy and math activities, a median split based on the number of children was used to create a factor with two levels; preschools respectively kindergartens where relatively few versus relatively many observation intervals were coded as teacher-managed activities. Since hardly any interval was coded as teacher-managed math activity in preschools, these activities were not further taken into account in the analyses (the percentage ranged from 0 to 0.76, M = 0.20, SD = 0.29). The descriptive statistics of teacher-managed language-literacy and math activities are for both groups shown in Table 2.2.

Table 2.2 Descriptive statistics of preschools and kindergartens low in teacher-managed language-literacy and math activities versus preschools and kindergartens high in these activities (n is the number of children)

			n	M % (SD)	Range %
Preschool	Literacy activities	Low	22	3.23 (2.17)	0.93–6.31
		High	24	13.87 (2.12)	12.37–16.29
Kindergarten	Math activities	Low	22	0.99 (0.75)	0.00–1.96
		High	23	2.80 (0.03)	2.77–2.83
	Literacy activities	Low	20	11.53 (2.50)	8.06–13.56
		High	25	18.38 (5.54)	13.98–26.71

To assess whether children in classrooms where relatively many teacher-managed activities were observed showed larger gains in related school readiness skills than children in classrooms where fewer of these activities were carried out, repeated measures ANOVAs were performed with the level of teacher-managed activities (based on the median split) as between subjects factor. For both the group scoring low and the group scoring high on teacher-managed language and literacy activities, the descriptive statistics of the outcome

measure are shown in Table 2.3. The repeated measures ANOVAs showed that preschool children in classrooms with a relatively high proportion of teacher-managed language and literacy activities, showed larger gains in literacy skills than children in preschools where fewer of these activities were carried out (F (1, 44) = 3.69, η^2 = .08, p = .06). With regard to children in kindergarten, the analysis showed that their language and literacy development was not accelerated by participating in a kindergarten classroom where the teacher initiated relatively many literacy activities (F (1, 43) = .16, η^2 = .00, p = .69).

As mentioned before, the occurrence of teacher-managed math activities in preschools was so scarce that it did not allow for a distinction in preschool classrooms where comparatively many versus few of these activities were carried out. In kindergarten, however, children whose teachers initiated and directed relatively many math activities showed larger gains than children in classes with fewer teacher-managed math activities (see Table 2.4 for descriptive statistics). However, this small effect was not significant (F (1, 43) = 2.45, η^2 = .05, p = .13).

Given the fact that math and language skills are highly intertwined, we also assessed whether children's math skills were benefited by teacher-managed language and literacy activities. The analyses showed that preschool children in classrooms with a comparatively high proportion of teacher-managed language and literacy activities showed larger gains

Table 2.3 Children's development in language-literacy skills in classrooms low in teacher-managed language and literacy activities versus in classrooms high in these activities

		Children's literacy skills					
			Time 1		Time 2		T2 - T1
	Teacher literacy activities	*n*	*M*	*SD*	*M*	*SD*	*M*
Preschool	Low	22	5.36	5.10	16.49	8.98	11.13
	High	24	6.22	6.64	23.66	15.34	17.44
Kindergarten	Low	20	28.40	11.53	49.65	17.05	21.25
	High	25	18.59	10.95	38.46	14.37	19.87

Table 2.4 Kindergarten children's development in math skills in classrooms low in teacher-managed math activities versus in classrooms high in these activities

		Children's math skills					
			Time 1		Time 2		T2 - T1
	Teacher math activities	*n*	*M*	*SD*	*M*	*SD*	*M*
Kindergarten	Low	22	23.07	8.20	36.27	6.24	13.20
	High	23	16.52	8.59	32.18	6.96	15.66

in math skills than children in preschools where fewer of these activities were observed (F (1, 45) = 9.01, η^2 = .17, p = .00). In contrast to language and literacy development, children in kindergarten showed larger gains in math skills when they were in classrooms where relatively many teacher-managed language and literacy activities were carried out (F (1, 43) = 3.91, η^2 = .08, p = .05) (see Table 2.5 for descriptive statistics).

Table 2.5 Children's development in math skills in classrooms low in teacher-managed language and literacy activities versus in classrooms high in these activities

		Children's math skills					
			Time 1		Time 2		T2 - T1
	Teacher literacy activities	*n*	*M*	*SD*	*M*	*SD*	*M*
Preschool	Low	23	5.09	4.25	11.19	6.16	6.10
	High	24	4.20	4.77	14.61	7.96	10.41
Kindergarten	Low	20	24.10	8.29	36.85	6.45	12.75
	High	25	16.22	7.96	32.05	6.54	15.83

Child-managed language-literacy and math activities

In addition to teacher-managed language-literacy and math activities, the effect of child-managed language-literacy and math activities on school readiness development was also examined. Again the observed preschools and kindergartens were split in two groups; preschool classrooms respectively kindergarten classrooms where children were observed to initiate or manage relatively few versus relatively many of these activities (see Table 2.6 for descriptive statistics).

Table 2.6 Descriptive statistics of preschools and kindergartens low in child-managed language-literacy and math activities versus preschools and kindergartens high in these activities (*n* is the number of children)

			n	*M* % (*SD*)	Range %
Preschool	Math activities	Low	23	0.46 (0.62)	0.00–1.31
		High	24	3.59 (2.40)	1.50–7.03
	Literacy activities	Low	24	1.14 (0.46)	0.75–1.79
		High	22	3.08 (1.17)	1.92–4.36
Kindergarten	Math activities	Low	21	1.75 (0.52)	1.03–2.33
		High	24	7.89 (7.70)	2.61–19.32
	Literacy activities	Low	26	0.36 (0.41)	0.00–0.75
		High	19	2.82 (1.33)	1.31–4.37

With regard to language and literacy skills, the proportion of child-managed language and literacy activities in the classroom revealed a medium effect on language and literacy development; preschool and kindergarten children attending classrooms with a relatively high degree of child-managed language and literacy activities, showed larger gains in language and literacy development than children in classrooms that were relatively low in these activities, though not statistically significant at $p = .05$ (preschool: $F(1, 44) = 2.71$, $\eta^2 = .06$, $p = .11$; kindergarten: $F(1, 43) = 3.34$, $\eta^2 = .07$, $p = .08$) (see Table 2.7 for descriptive statistics).

For children in both preschool and kindergarten frequent engagement in child-managed math activities was not associated with larger gains in math skills (preschool: $F(1, 45) = 1.52$, $\eta^2 = .03$, $p = .22$; kindergarten: $F(1, 43) = 0.10$, $\eta^2 = .00$, $p = .76$) (see Table 2.8 for descriptive statistics).

Table 2.7 Children's development in language-literacy skills in classrooms low in child-managed language and literacy activities versus in classrooms high in these activities

		Children's literacy skills					
			Time 1		Time 2		T2 - T1
	Child literacy activities	*n*	*M*	*SD*	*M*	*SD*	*M*
Preschool	Low	24	5.29	4.70	17.09	11.44	11.80
	High	22	6.38	7.05	23.65	14.13	17.27
Kindergarten	Low	26	19.99	10.46	37.91	12.72	17.92
	High	19	27.00	13.33	51.00	18.18	24.00

Table 2.8 Children's development in math skills in classrooms low in child-managed math activities versus in classrooms high in these activities

		Children's math skills					
			Time 1		Time 2		T2 - T1
	Child math activities	*n*	*M*	*SD*	*M*	*SD*	*M*
Preschool	Low	23	4.82	3.56	12.14	5.79	7.32
	High	24	4.47	5.31	13.70	8.50	9.23
Kindergarten	Low	21	21.48	9.59	35.67	6.73	14.19
	High	24	18.19	8.22	32.88	6.86	14.69

Discussion

In this study we examined children's everyday activities in Dutch preschool and kindergarten classrooms. Subsequently, the amount of teacher- and child-managed language-literacy and math activities was related to children's gains in related domains of school readiness skills over a one year period. With regard to teacher-managed academic activities there are two main findings. First, the occurrence of teacher-managed academic activities was quite low. Especially math activities were rarely initiated by teachers. Second, despite their relative low occurrence, our findings showed that teacher-managed activities have the potential to benefit children's development. Preschool children's language-literacy and math development was positively associated with the amount of teacher-managed language and literacy activities in the classroom. Kindergarten children showed larger gains in math skills in classrooms where teachers initiated relatively many language-literacy and math activities. In contrast, an effect of teacher-managed language and literacy activities on kindergarten children's language and literacy development was not found.

The low occurrence of teacher-managed math activities is not surprising given previous research showing that young children's teachers tend to focus more on language and literacy than on math activities (Graham, Nash, & Paul, 1997; Howes et al., 2008; Phillips et al., 2007; Phillips et al., 2009). Although we cannot entirely rule out the possibility that the teachers in this study incorporated math activities in subtle ways in other activities of their program that were not detected by the observers, the presently studied preschool and kindergarten classrooms seem to be lacking a clear focus on math. The relative position of math activities in the present sample seems even less favorable compared to findings in other studies involving U.S. preschools that revealed a more balanced focus on language, literacy and math in early childhood education settings (Chien et al., 2010; Connor et al., 2006; Early et al., 2010; Elicker & Mathur, 1997; Howes et al., 2008; Phillips et al., 2007; Phillips et al., 2009; Tonyan & Howes, 2003). Moreover, also similar to other studies, if teacher-managed math activities were coded in the observed classrooms, they were hardly requiring higher order mathematical thinking (Graham et al., 1997; Tudge & Doucet, 2004), but rather involved counting or shape naming.

In comparison with math activities, teacher-managed language and literacy activities were more common, however, in preschool, on average, these activities still belonged to the least prominent ones. Similar to other studies, our observations further revealed that in both preschool and kindergarten much time was spent on transition activities such as gathering material to start an activity, tidying up after an activity, and awaiting one's turn when choosing an activity on a plan board (Connor et al., 2006; Early et al., 2005; Elicker & Mathur, 1997). Although part of the transition and waiting activities might have contributed

to children's development as well (e.g., the informal sorting of play materials during tidying up might have promoted emergent math skills), we think that, on average, too much time is lost. According to our informal impressions (not specifically coded) most of the transition time boils down to waiting time and does not contribute at all to the developmental domains under investigation in this study. As reported in other studies, much time was also spent on daily care routines such as eating snacks and drinking (Chien et al., 2010; Early et al., 2005; Early et al., 2010; Elicker & Mathur, 1997). Furthermore, a remarkable predominant activity in terms of time spending in both preschools and kindergartens in our sample was outdoor play (17 respectively 18% of the total time), leaving less time for academic activities. Although outdoor play provides many opportunities for promoting children's development, informal observations revealed that teachers rarely took advantage of these opportunities.

The contribution of teacher-managed language and literacy activities to preschool and kindergarten children's math development might seem counterintuitive at first sight. However, an improvement in language skills probably leads to better math skills as well, because improved language skills may help the child to better understand mathematical concepts. Furthermore, teacher-managed language and literacy activities are presumably interlarded with count words, comparison words and other mathematical concepts as well. Klibanoff et al. (2006) for instance, describe in their study on math talk that teachers' references to mathematical concepts frequently occurred outside the context of planned mathematical instruction, for example during creative activities and book reading. In addition, Van den Heuvel-Panhuizen et al. (Van den Heuvel-Panhuizen & van den Boogaard, 2008; Van den Heuvel-Panhuizen, van den Boogaard, & Doig, 2009) found that picture books written without the intention to teach children about mathematics evoke mathematics-related thinking in kindergarten children; about half of children's spontaneous utterances in reaction to the picture books were mathematics-related.

There is a contradictory and unexpected finding in the results: on the one hand there is an association between teacher-managed language and literacy activities and children's development of mathematical skills; on the other hand, these teacher-managed activities do not appear to promote children's language and literacy development, at least in the kindergarten group. A possible explanation is that the nature of teacher-managed language and literacy activities is relatively global and, as a result, less specifically adapted to individual children's needs for language learning. Taken together, these factors might have made it more difficult to isolate effects of teacher-managed language and literacy activities on kindergarten children's development in these skills. In addition, another observation study (de Haan, Elbers, & Leseman, 2014a) revealed differences between preschools and kindergartens that might have had an impact on the children's language and literacy development; partly caused by the higher teacher to child ratio in preschool, kindergarten children had less

interaction with their teacher than preschool children. Overall, preschool children might rely more on their teacher as a source of development, which might explain the positive association between teacher-managed language and literacy activities and preschool children's development in language and literacy skills. Still, the result is puzzling, given the fact that language and literacy activities in both kindergarten and preschool did affect children's mathematical skills. Doing qualitative analyses of interactions between teachers and children, while comparing the interactions in preschool and kindergarten classes, might shed light on these unexpected results in the future.

With regard to child-managed activities we found that activities like picture book reading (leafing through the pages, looking at pictures) and copying words with letter stamps and similar activities involving print, were related to children's gains in these skills. In contrast, child-managed math activities were not associated with children's development in math. A possible explanation for these mixed findings is that the most frequently observed child-managed math activities, jigsaw puzzle solving and pattern completion activities, are less obviously related to the outcome measure than the child-managed language and literacy activities were related to that particular outcome measure. Virtually all observed teacher-managed math activities, of which we did find an effect on kindergarten children's math skills, were concerned with basic emergent math skills such as counting, naming shapes and comparing quantities, what might explain their contribution to children's development in math skills as assessed with our outcome measure (cf. Klibanoff et al., 2006).

Although our study provides insight into the role of academic activities in preschools and kindergartens with a predominant developmental approach and in how these activities may contribute to children's school readiness, the possibility that other activities than taken into account in our analyses contributed to children's language-literacy and math development cannot be excluded. A future observation study, in which more fine-grained activity categories will be used, should further disentangle the contribution of different activities to children's development. To complement our quantitative results, future qualitative studies should detail the content of the activities and, even more important, the instruction and interaction quality of the activities. Furthermore, the present sample was quite small, limiting statistical generalization. Future studies should not only include a larger sample size, but ideally also samples from different geographical areas and ethnic communities to examine the generality of the findings of the current study. Despite these limitations we conclude that teacher-managed academic activities are infrequent, overall, in the preschools and kindergartens involved in this study. Nevertheless, if present, these activities can have a considerable effect on children's school readiness development, especially for the youngest children.

We found a large range regarding the time teachers spent on academic activities in the preschool and kindergarten classrooms in this study (preschool: 1–16%; kindergarten:

11–27%), indicating that some teachers were apparently quite able to invest a considerable amount of time in academic content, whereas other teachers, working in the same system with the same pedagogical approach, hardly provided any academic content during the four half days of observation. Similar findings regarding teacher differences were reported in previous studies (Connor et al., 2006; Early et al., 2005; Meyer et al., 1993). Our study contributes to the growing body of evidence that academic content in early childhood education programs is important for children's school readiness development, especially for children from disadvantaged backgrounds (Bus et al., 2012). Therefore, teachers should be urged to reconsider their daily schedules, to incorporate academic content activities in their program and to minimize the loss of time due to mere waiting, transitions between activities, and lack of interesting activities. Improving classroom management skills and exchange of best practices among teachers might propel this process. Furthermore, teachers should become aware of all opportunities to stimulate children's development. In our study, some teachers proved to be quite successful in incorporating stimulating activities during transition time. For example, they asked children to arrange themselves in a row from the smallest to the tallest child before going outside for play. Furthermore, teachers should be encouraged to overcome their often reported discomfort in carrying out math-related activities (Ginsburg, Lee, & Boyd, 2008). Although the teachers in this study, on average, spent very little classroom time on math activities, there was a small marginally significant effect on kindergarten children's math development. More math activities might lead to larger gains in children's math skills (Guarino, Hamilton, Lockwood, Rathbun, & Hausken, 2006).

An important final point to highlight is that nearly all teacher-managed activities were carried out in whole group settings. It is likely that a whole group setting limits interaction opportunities and may therefore be less effective in stimulating development; the teacher has to divide his or her attention between many children and usually one child at a time is allowed to make a contribution to the ongoing activity (Powell, Burchinal, File, & Kontos, 2008). Research findings support this claim by showing that children in whole group settings are twice as likely to be passively listening or merely watching than talking and acting. Furthermore, the effect of small group instruction on skill development is reported to be 10 times greater than instruction in whole group settings (Connor et al., 2006). In sum, these studies suggest that teacher-managed academic activities carried out in small group settings might be more effective in accelerating children's development than similar activities in whole group settings. Nevertheless, teachers rarely seem to initiate teacher-led small group activities and children spend much time in less effective whole group settings (Chien et al., 2010; Elicker & Mathur, 1997). An important condition for effective early childhood education is the optimal allocation of time, with an alternation of teacher-managed

academic activities and child-managed educational and play activities while loss of time is restricted to a minimum. It is important to stress that rather than arguing in favor of a highly academic skills oriented or rote learning approach, we suggest to embed academic content in a mixture of meaningful teacher-managed academic and child-managed play activities.

Chapter 3

Targeted versus mixed preschools and kindergartens: Effects of class composition and teacher-managed activities on disadvantaged children's emergent academic skills

de Haan, A. K. E., Elbers, E., Hoofs, H., & Leseman, P. P. M. (2013). Targeted versus mixed preschools and kindergartens: Effects of class composition and teacher-managed activities on disadvantaged children's emergent academic skills. *School Effectiveness and School Improvement*, *24*, 177-194.

Abstract

In this study, longitudinal cohort-sequential latent growth modeling was used to determine the effects of a) socioeconomically mixed preschool and kindergarten classrooms, b) the implementation of an education program aiming to accelerate socioeconomically disadvantaged children's emergent academic skills and c) the amount of teacher-managed academic activities on three to six-year-old disadvantaged children's emergent literacy and math skills (n = 91). The results indicate that disadvantaged children in mixed preschool and kindergarten classrooms gained more in literacy and math than disadvantaged children in targeted classrooms. The results also indicate that the use of a special education program to promote disadvantaged children's emergent literacy and math was not effective, probably because of a lack of implementation fidelity. However, the extent to which teachers engaged in the kind of activities that were intended by the education program, in particular frequent initiation and guidance of language, literacy and math activities, was significantly related to disadvantaged children's outcomes.

Introduction

In face of the persistent achievement gap, most industrialized countries provide early childhood education programs to promote the cognitive, language, literacy and numeracy skills of young children from low income and minority groups in order to provide them with a fair start in primary school (OECD, 2006). Numerous studies have shown that children who attended preschool or educational day care, enter primary school with higher levels of cognitive, language and math skills, and better behavioural adjustment (Anderson et al., 2003; Barnett, 1998; Blok et al., 2005; Burger, 2010; Campbell et al., 2001; Gorey, 2001; Nelson et al., 2003; Ramey & Ramey, 2004; Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2010). In addition, children seem to show better outcomes in programs of high quality (Bryant et al., 1994; Burchinal et al., 2000; Howes et al., 2008; Mashburn et al., 2008; Peisner-Feinberg et al., 2001). Commonly, two aspects of quality are distinguished (Howes et al., 2008). The first aspect concerns the structural characteristics of the classroom, such as teacher's education, classroom group size, and the adult to child ratio. The second aspect concerns the interaction processes in the classroom, including the social-emotional and instructional support teachers provide to children. Structural quality is considered an essential precondition for process quality, but thought to affect child outcomes only indirectly (Howes et al., 2008). Although the results of high quality preschool programs for educationally disadvantaged children are generally positive, several questions remain regarding the ways in which early childhood education for disadvantaged children should be provided.

The present study addresses two issues that seem currently relevant given the increased policy focus in several countries on extending early childhood education provisions: should we concentrate efforts on children most in need or adopt a more universal approach, and, related to this, should we implement special education programs in targeted provisions or focus on improving the quality of early childhood teachers working in universal systems? We report the results of a small-scale intervention study conducted in The Netherlands evaluating the effects of 1) targeted versus universal mixed educational arrangements, 2) the use of a special education program and 3) actual instruction activities by the teachers regardless of the use of a special program on socioeconomically disadvantaged preschool children's emergent literacy and math skills.

Considering early childhood education, two main types of intervention can be distinguished: universal and targeted interventions. Universal interventions do not concentrate on a specific subgroup, but instead provide care and education for all children regardless of whether they belong to an educationally disadvantaged group or have special educational needs, presumably to benefit all children. Targeted interventions, in contrast, are specifically provided to eligible groups only. In the Dutch context these targeted groups concern families with non-western backgrounds and/or weak educational backgrounds. Targeted interventions seem most cost-effective and are thought to yield substantial economic benefits in the long-term (Heckman, 2006). Indeed, insofar as universal interventions are to the benefit of all children, it is only through targeted interventions that society can compensate the early educational disadvantages caused by poverty and immigration (Magnuson & Waldfogel, 2005).

However, there may be a major drawback of targeted interventions, not taken into account in the aforementioned argument. Due to the concentration of problems related to poverty and immigration, the educational effectiveness of early childhood education programs serving mainly educationally disadvantaged children may be reduced. Only few studies to date have addressed the effects of socioeconomic, racial, and ethnic-cultural classroom composition in early childhood education programs on cognitive and social-emotional outcomes. Lee et al. (1998) found medium to strong negative effects of a high concentration of children from poor minority families in preschool care and education settings. Schechter and Bye (2007) compared targeted and mixed preschools and reported a clear advantage for the mixed arrangement. In mixed arrangements the early delay in disadvantaged children's language development relative to age norms gradually decreased, whereas in the targeted arrangement the delay further increased. Henry and Rickman (2007) did not find effects of classroom socioeconomic composition as such on disadvantaged children's cognitive and language development, however they found that these children profited from interaction with more able peers. Mashburn et al. (2009) found similar effects

of interaction with more able peers on children's language development, but this effect seemed to be limited to children who were already functioning on a high level of language skill. Gormley Jr. and Phillips (2005) showed that disadvantaged children benefited more from universal pre-kindergarten programs with a mixed composition than from targeted programs, presumably because of the opportunity to interact with more able peers. Findings in Dutch research suggest that a small share of disadvantaged children in preschool classrooms helps these children to catch-up, whereas a high concentration enhances the early education gap (Mayo & Leseman, 2008).

One of the reasons for the suboptimal results of education programs in decreasing the achievement gap in both short- and long-term, is the lack of implementation fidelity. Several studies reveal large differences between early childhood teachers in providing activities as intended, while working within the same program and having received the same training (Connor et al., 2006; Early et al., 2005; Klibanoff et al., 2006; Meyer et al., 1993). Moreover, teacher differences are related to child outcomes. For instance, Connor et al. (2006) found that only when preschool teachers who were working within the same comprehensive program for disadvantaged children spent comparatively more time to language and emergent literacy activities, their children gained in alphabet knowledge, letter-word recognition and vocabulary. In addition, Justice et al. (2008a) found children participating in classrooms with relatively much time spent on language learning activities to show accelerated language growth in pre-kindergarten compared to children in a control group. Similarly, Klibanoff et al. (2006) found large differences between teachers in the same preschool system in the amount of math-related talk they initiated in the classroom, with the amount of math talk being strongly related to children's growth in pre-mathematical knowledge and skills. In sum, focusing on the teacher, on what she or he does in the classroom, may provide starting points for improving the quality and effectiveness of early childhood education.

The Dutch preschool system

Pre-primary school education in The Netherlands consists of two interconnected systems, spanning the age range of 2;6 to 6 years of age. Children are admitted to preschools from age 2;6. There is a small, income dependent fee and about 60% of the eligible children participates for two to four half days per week (about 2.5 hours per day). Educationally disadvantaged children, according to the national educational priority policy defined as coming from families with a non-western and/or weak educational background, are especially encouraged to participate. At age 4, the vast majority of children, over 98%, start in the kindergarten departments of primary school for five days per week (about five hours per day). There is no fee and attendance is compulsory from age five onwards. Both preschools and kindergartens

implement a basic developmental, child following approach with ample room for free play, fine and gross motor activities, and creative and expressive work. Especially in inner city areas, in view of the educational disadvantages of part of the participating children, preschools and kindergarten often work with a special education program to promote disadvantaged children's language, literacy and math development. In the preschool system this development has led to the implementation of 'educational preschools' intended exclusively for disadvantaged children, indicated by their parents' educational attainment and cultural background. Since only disadvantaged children are eligible for participation, these educational preschools are an example of the targeted approach. Although the Dutch kindergarten system is an example of a universal childhood education provision, since all children are eligible for participation, some kindergartens are attended almost exclusively by disadvantaged children, mainly because of the location of the school in neighborhoods with high concentrations of minority and low social-economic status families.

The current study

The current study was conducted as an evaluation of the effects of mixing within preschool and kindergarten classrooms serving educationally disadvantaged children. Part of the participating preschool and kindergarten sample implemented a structured education program. Use of the program and creating mixed classrooms varied independently. In addition, detailed observations of teachers' initiation, guiding and directing of language, literacy and math activities were obtained in all classrooms. We will answer three questions in this article. First, we will determine whether enrollment in a mixed preschool or kindergarten yields more favorable results in terms of emergent academic skills development than participation in a targeted program. Second, we will assess whether children in preschools and kindergartens working with a special education program show larger gains in emergent academic skills than children in provisions without such a program. Third, related to this, we will assess whether teachers using a special education program initiate, guide and direct more stimulating educational activities in the areas of literacy and math than teachers without such program, and determine whether teachers' initiating, guiding and directing of literacy and math activities promotes children's academic skills development.

Method

Participants

A sample of 91 children was involved in the present study, 48 children (28 boys; 58%) attending one of the fourteen participating preschool classrooms and 43 children (17 boys; 40%) attending one of the twelve participating kindergarten classrooms, all located in a middle-sized town in the western part of The Netherlands (see Table 3.1 for descriptive statistics).

Only children eligible for extra support within the framework of the national educational priority policy were selected in these classrooms. The present sample, therefore, consists exclusively of children considered educationally disadvantaged, that is, they came from ethnic-minority families where at least one of the parents had a weak educational background or from Dutch families where both parents had a weak educational background (not higher than the junior vocational training level). The mean educational attainment level of parents (preschool cohort: 3.4 (*SD* = 1.7); kindergarten cohort: 3.2 (*SD* = 2.1)) on a seven point scale ranging from 1 (only a few years of primary education at most) to 7 (university degree), indicates an average educational attainment below junior vocational training level.

Parents of children who met our age criteria were informed about the study and asked for their active informed consent resulting in participation of the majority of eligible children (84% of the preschool and 79% of the kindergarten children).

Table 3.1 Mean age, sex, mean educational attainment level and cultural background of parents for the preschool and kindergarten cohort

		Preschool		Kindergarten	
N		48		43	
Mean age time 1 (*SD*)		3.0	(0.4)	4.3	(0.3)
Boys (%)		28	(58%)	17	(40%)
Mean educational level parents		3.4	(1.7)	3.2	(2.1)
Cultural background parents (%)	Dutch	12	(25%)	6	(14%)
	Moroccan	15	(31%)	19	(44%)
	Turkish	6	(13%)	7	(16%)
	Other/mixed	15	(31%)	11	(26%)

Note. The mean educational attainment level is measured on a seven point scale ranging from 1 (only a few years of primary education at most) to 7 (university degree).

Procedures

Child assessment

Children were tested three times, first upon entrance in preschool or kindergarten and then on two occasions one, respectively two, years later. At all measurement occasions, testing was divided over two sessions of approximately 30 minutes. Children were tested individually by trained research assistants, using laptop computers, in a quiet room at preschool or kindergarten.

The Dutch version of the PIPS; Performance Indicators in Primary Schools (Tymms, 2001), a computerized adaptive assessment of a wide range of emergent literacy and math skills, was used to assess emergent literacy and math skills. Previous research with the PIPS in The Netherlands showed good predictive validity and high reliability (Van der Hoeven-van Doornum, 2005).

Classroom observations

To assess the opportunities for engagement in activities deemed relevant for academic skills development, observations were conducted in all preschool and kindergarten classrooms focussing on six children per classroom who were randomly selected for the observations. Teachers were instructed to carry out their usual activities and routines. The observations by trained research assistants and the first author took four half-days of 2.5 to 3 hours in a two-week period and were conducted shortly before the second round of child assessment. A cyclic interval coding approach was used to observe the target children. A child was observed during ten seconds. After ten seconds, the activity the child was engaged in during the preceding interval was coded for type of activity and involvement of the teacher (and a number of other categories not included in the present study). When the coding of a particular interval was completed (which took on average about 40 seconds), a ten seconds observation of the next child started. The cycles of observing and coding were repeated from the start until the end of the morning or afternoon. The coding scheme was programmed in E-prime (Schneider et al., 2002) and installed on laptop computers. The procedure resulted for the preschool classrooms in a mean number of 209 observation intervals (*SD* = 32.2) in the mornings and 128 (*SD* = 19. 6) intervals in the afternoons, which lasted about one hour less than the mornings. In kindergarten classrooms, where children spent about one hour longer in the morning and thirty minutes longer in the afternoon, a mean number of 349 intervals (*SD* = 57.5) were observed and coded in the morning and a mean number of 209 intervals (*SD* = 40.6) in the afternoon.

To determine inter-rater reliability, assistants coded detailed transcriptions of realistic classroom situations, based on previously made video recordings, which were also programmed in E-prime (Schneider et al., 2002). The coding procedure of the situations was

similar to the coding procedure in real time classroom situations. Assistants coded on average 35 different situations. The inter-coder reliability was satisfactory, with a Cohen's kappa of .80.

Measurements

Emergent literacy

The emergent literacy skills test consisted of six subtests that were combined in a single score: receptive vocabulary, writing, ideas about reading, letter identification, reading words, and reading sentences. The maximum number of items is 180, but due to the adaptive nature of the test (i.e., after a certain number of errors the subtest at hand is automatically ended and the program continues to the next subtest) not all items were administered to all children. At both measurement occasions missing scores, all missing at random, were imputed separately for the preschool and kindergarten cohort using the EM method. The percentage of missing scores ranged between 2% and 27%.

Emergent math

Emergent math skills consisted of seven subtests: ideas about math, counting, informal sums, digit identification, shapes, adding and subtracting, and advanced sums. The maximum number of items is 69, and again due to the adaptive nature of the test not all items were administered to all children. Also for emergent math skills missing scores were imputed separately for the preschool and kindergarten cohort using the EM method. The percentage of missing scores ranged between 2% and 25%.

Teacher-managed activities

After every ten-second observation interval the activity the child was engaged in during the interval was coded. If two activities were observed during one interval, the activity the child was engaged in longest was coded. In addition, we also coded who initiated and/or was mainly in charge of a particular activity, the teacher, the observed child, or other children.

The activity coding scale was adapted from Howes and Smith (1995). Typical examples of activities were making music and singing, creative and craft activities, free play, outdoor play and snack time (the complete coding scheme is available from the first author upon request). In the present study, we focus specifically on activities with a presumed academic content. Composite scores were computed based on the six randomly selected children and aggregated to the classroom level representing the mean percentage of intervals of the total number of observed intervals with teacher-managed activities. The following measures were included in the current study.

Teacher-managed literacy activities

Language and literacy activities such as book reading, language games (examples are thinking of words or names starting with a certain letter), rhyming, introducing and explaining new words, and classroom talk about knowledge topics that were initiated, directed or guided by teacher were considered teacher-managed literacy activities.

Teacher-managed math activities

Activities involving counting, rank ordering, talk about numbers, quantities or birth dates, use of the calendar, naming geometrical shapes, playing with board games, measuring, comparing and estimating, that were initiated, directed or guided by the teacher were considered teacher-managed math activities.

Education program

Preschools and kindergartens working with a special education program to promote disadvantaged children's emergent academic skills were compared with preschools and kindergartens without such program, but providing regular care and education activities. All educational preschools and kindergartens in the current study that implemented a special education program, used the program *Taalrijk* ("Enriched Language"). The program provides a curriculum that starts in preschool and continues into kindergarten, using a thematic approach to foster children's language, emergent literacy and emergent math skills. Themes spanning about three to four weeks give concrete suggestions to the teachers for sharing time topics, story books to read and small group activities with fantasy play materials and special materials for literacy and math development. A core element of the program is the alternation of whole group and small group activities, following a cyclic schedule that recommends to organize small group activities at least twice a week. In addition, a manual gives examples of how to organize an activity, the kind of instruction language to be used and, especially, the list of words and concepts to be incorporated in the activity. Typical themes are the seasons of the year, animals in the zoo and at the farm, healthy food, injuries and illnesses, dwarfs and giants. Teachers working with the program received eight half days initial training focusing on a general sensitive interaction style and working with the manual and providing help with scheduling the activities and introducing ways of monitoring children's development. In addition, licensed trainers visited the classrooms for observation and consultation about four times in a two year period to ensure implementation fidelity. Preschools and kindergartens without this education program also provided a mix of child-centered and teacher-managed activities, implemented a regular curriculum of activities which partly overlapped with the Taalrijk activities, used occasionally themes as well, but

mostly in whole group settings and less extensively (for a description of regular preschool education in The Netherlands, see Van Tuijl & Leseman, 2007). Moreover, disadvantaged children attended regular preschools for two instead of three to four half days, that is, 5 versus 7.5 (occasionally) to 10 hours per week (mostly) in educational preschools. Attendance in kindergarten was the same for all children, about 25 hours per week. All preschools and kindergartens had standard equipment of fantasy and construction play materials, special educational materials (for instance, letter boxes, jig-saw puzzles, categorization materials), arts and craft materials, music instruments and outdoor space and equipment for gross motor activities. All preschools and kindergartens had at least one fully licensed teacher (intermediate vocational training level in the preschools, higher vocational bachelor level in the kindergartens). In all preschools and some kindergartens the teachers were assisted by another teacher or assistant-teacher for part of the day, but in preschools and kindergartens working with Taalrijk, these extra teachers were also trained in working with the program. For the present purposes, a dummy variable was created indicating use of the program Taalrijk (value 1) versus the regular curriculum (value 0).

Mixed classrooms

Two types of classroom arrangements were distinguished. The targeted arrangement was characterized by classrooms that consisted in majority of children who were considered disadvantaged according to official educational policy criteria, based on parents' educational attainment and ethnic-cultural background. In the preschools, targeted groups on average consisted of 83% (range 64% to 100%) disadvantaged children; in targeted kindergarten groups on average 61% (range 48% to 68%) of the children were labelled disadvantaged. In contrast, the mixed arrangement concerned preschool and kindergarten classrooms with a comparatively low representation of disadvantaged children. In preschools, the range was 17% to 57%, in kindergarten 10% to 38%. For the present purposes, a dummy variable was created indicating mixed (value 1) versus targeted (value 0) classroom composition. Some of the mixed preschools and kindergartens also used the special education program, whereas others provided the standard curriculum.

Analysis plan

The study was designed as a longitudinal cohort-sequential study involving three measurement times in each cohort, with two overlapping times for the preschool and kindergarten cohort (around age 4 and 5). Accelerated Latent Growth Modeling (LGM; Duncan, Duncan, & Strycker, 2006), with AMOS (version 17; Arbuckle, 2006), was used to

combine the two age cohorts in single models of children's literacy and math development, spanning the age range from 3 to 6 years, estimating the means and the variances of the intercepts and slopes for the dependent variables, and the effects on intercepts and slopes of the predictor variables *education program*, *mixed arrangement* and *teacher-managed academic activities* (for short: *teacher activity*). Given the fact that the children were not randomly assigned to program, arrangement and teacher, and may have differed in initial level of literacy and math skills due to selection, the analysis focused specifically on the slopes of children's development and on the effects of the predictor variables on the slopes. To evaluate model fit, the chi-square (χ^2), the root mean square error of approximation (RMSEA), and the comparative fit index (CFI) were used. As a rule of thumb, a non-significant χ^2 or a ratio of χ^2 and the number of degrees of freedom < 2 indicates good model fit, RSMEAs below .05 indicate good fit and below .08 reasonable fit, and CFI greater than .95 can be considered a good fit and values greater than .90 indicate an acceptable fit (Kline, 2005).

Two series of LGM-models were tested for emergent literacy and emergent math, respectively. First, growth models were fitted. Regression weights for the intercepts were fixed to 1, regression weights for the slopes were fixed to reflect the actual measurement times. Using the multiple group option, regression weights for the preschool cohort were fixed at 0, 1 and 2, and for the kindergarten cohort at 1, 2 and 3 respectively. Moreover, the measurement errors of the measures taken at the overlapping measurement times (T2 and T3 for the preschool cohort, T1 and T2 for the kindergarten cohort, respectively), the intercept means and variances, the slope means and variances, and the covariances between intercepts and slopes were constrained to be equal for both cohorts. Non-linear growth was examined by free estimation of the first measurement weight in the preschool cohort and the final measurement weight in the kindergarten cohort. The model for emergent literacy fitted significantly better when the slope regression weight reflected a quadratic trend. The model for emergent math fitted the linear model best. See Figure 3.1 for an illustration of the combined accelerated latent growth model. Second, the three predictor variables were added to the growth models, keeping the growth part of the models constant. Structural paths from the predictors to the intercepts and slopes, and covariances between the predictor variables were specified.

Results

The correlation matrix and the means and standard deviations of the dependent variables for each age cohort are shown in Table 3.2 (preschool cohort: above the diagonal; kindergarten cohort: below the diagonal). There were no outliers with a Mahalanobis distance > .10.

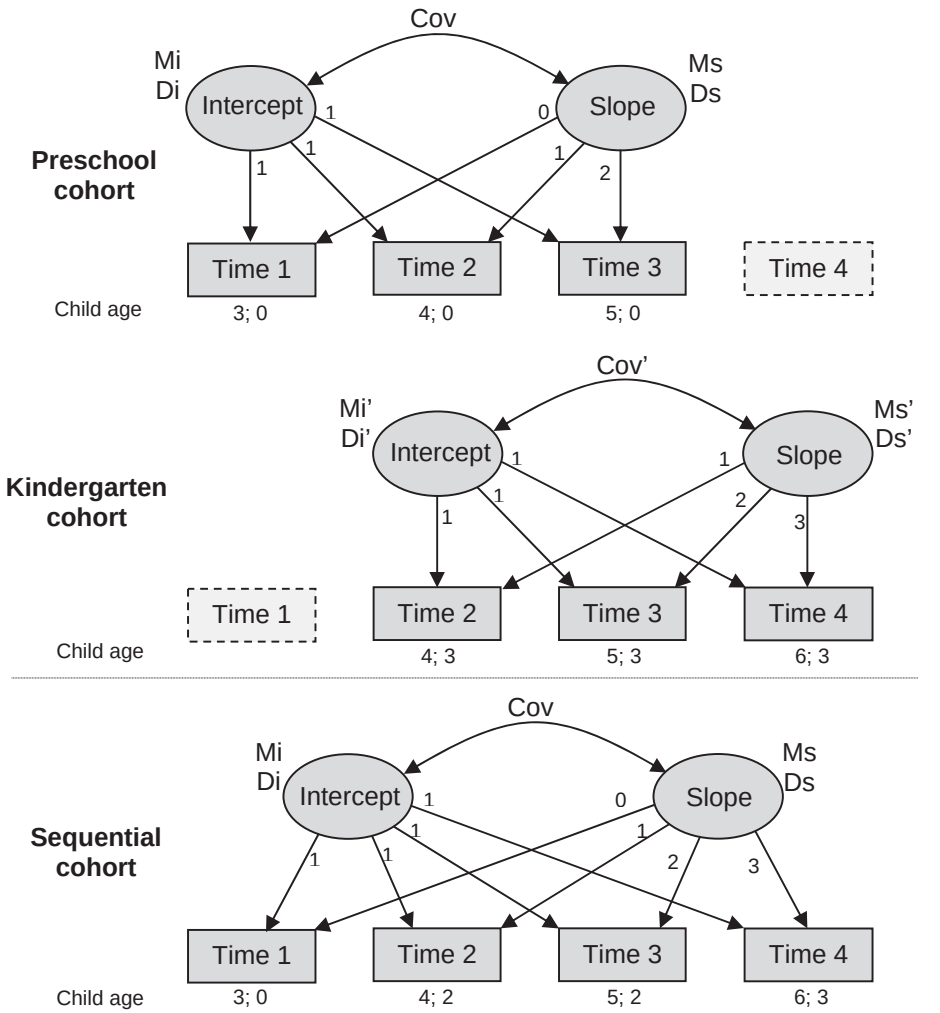

Figure 3.1 Depiction of the merging of the separate preschool and kindergarten cohort into one sequential cohort.
Note. Mi = Mean intercept, Di = Disturbance intercept, Ms = Mean slope, Ds = Disturbance slope, Cov = Covariance.

Normality was checked for each variable. The results of first measurement of literacy and math in the preschool cohort appeared to be skewed due to a floor effect (several children scoring zero). However, skewness was within the acceptable range (< 3).

Table 3.3 shows the means and standard deviations of the observed teacher-managed literacy and math activities by age cohort, program and arrangement. Teacher-managed math activities were quite rare, especially in preschool. Teacher-managed literacy activities

Table 3.2 Correlations, means and standard deviations for literacy and math skills for the preschool (above the diagonal) and the kindergarten cohort (underneath the diagonal)

								Preschool		
		1.	2.	3.	4.	5.	6.	Mean	(*SD*)	*N*
	1. Literacy T1	-	.40**	.46**				7.00	(6.68)	47
	2. Literacy T2	.54**	-	.64**				19.10	(9.01)	47
	3. Literacy T3	.29^{+}	.54**	-				28.10	(10.06)	47
	4. Math T1				-	.72**	.65**	5.32	(4.67)	47
	5. Math T2				.73**	-	.73**	12.99	(7.79)	47
	6. Math T3				.70**	.84**	-	26.74	(8.57)	47
Kindergarten	Mean (*SD*)	18.18 (10.30)	37.68 (11.21)	91.25 (42.48)	16.65 (7.62)	31.22 (8.04)	44.13 (9.06)			
	N	43	43	43	43	43	43			

Note. Above the diagonal the data of the preschool cohort are shown, underneath the diagonal the data of the kindergarten cohort are shown.
$^{+}$ $p < .10$; * $p < .01$.

Table 3.3 Teacher-managed literacy and math activities by cohort, program and mixed arrangement; mean percents and standard deviations

			Teacher-managed literacy activities		Teacher-managed math activities	
			M %	*SD*	*M* %	*SD*
Preschools	Program	No	7.4	5.9	0.3	0.5
		Yes	8.5	6.4	1.4	1.2
	Mixed	No	1.7	1.2	0.7	0.1
		Yes	11.5	5.1	1.4	1.6
Kindergartens	Program	No	17.0	6.6	1.4	1.0
		Yes	13.7	2.7	2.1	0.9
	Mixed	No	15.3	2.0	2.8	0.2
		Yes	13.4	4.5	1.4	0.8

occurred more frequently, but on average not exceeding 17% of the observed classroom time, mostly less. Note, however, that there was overall a considerable range, pointing to differences between teachers. Although use of the education program was meant to enhance the focus on academic content and teacher initiative, observed teacher activities did not significantly differ between classrooms with and without the education program across cohorts ($F_{(1,90)}$ = .045, p = .83 and $F_{(1,90)}$ = .015, p = .90, for literacy and math respectively). In the preschool cohort a remarkably high percent of teacher-managed literacy activities was found for the mixed arrangement, which may have been coincidental. The overall effect of the mixed arrangement across cohorts was significant for literacy ($F_{(1,90)}$ = 6.83, p < .01), not significant for math ($F_{(1,90)}$ = .005, p < .95). The distribution of teacher-managed literacy and math activities was skewed in each cohort and showed some outlier values. For further analyses, therefore, median-split was used to create a dummy variable indicating comparatively high (value = 1) versus low teacher-managed activities (value = 0).

Testing combined growth and structural models

Model selection proceeded in three steps; the results are presented in Table 3.4. First, for each dependent variable, models were tested in which all structural parameters and all covariances between the predictors were constrained to be equal across the cohorts. As can be seen in the Table, the model for literacy did not yield acceptable fit, whereas the model for math fitted the data already reasonably well. Next, models were tested in which all structural parameters (the expected effects of education program, mixed classroom composition and

Table 3.4 Overview of the fit statistics for the fully constrained (Model 0), fully unconstrained (Model 1) and partly unconstrained and trimmed model (Model 2)

Model	Emergent literacy	Emergent math
Model 0 All parameters constrained to be equal across both age cohorts	$\chi^2 = 64.57$, df = 22 $p = .000$ CFI = .593 RMSEA = .147	$\chi^2 = 34.99$, df = 22 $p = .039$ CFI = .943 RMSEA = .081
Model 1 Structural parameter(s) set free to vary between age cohorts	$\chi^2 = 10.52$, df = 13 $p = .651$ CFI = 1.000 RMSEA = .000	$X^2 = 21.13$, df = 13 $p = .070$ CFI = .949 RMSEA = .084
Model 2 – trimmed As the previous model, but with non-significant structural parameters (at $p < .10$) constrained to zero	$\chi^2 = 26.63$, df = 23 $p = .272$ CFI = .965 RMSEA = .042	$X^2 = 32.40$, df = 23 $p = .092$ CFI = .959 RMSEA = .068

teacher activity) were allowed to vary between the preschool and kindergarten cohort. This yielded considerably and statistically significant better fit for the literacy model (a decrease in χ^2 of 54.06, with a loss of 9 degrees of freedom), but a smaller, yet significant improvement of the mathematics model (a decrease of χ^2 of 13.86 against a loss of 9 degrees of freedom). Both models had acceptable fit. Finally, the models were trimmed to obtain the most parsimonious solution. Parameters that were not significant on the $p < .10$ level were fixed to zero. As Table 3.4 shows, this yielded good fit for both models.

Table 3.5 presents the non-standardized estimates of the models for literacy and math based on the final models. Regarding the growth part of the models, the means and variances of the intercepts and slopes were all statistically significant, indicating that the differences between children in the overall level of literacy and math are substantial, that there is substantial overall growth of literacy and math skills over time, and that the individual differences in the rate of growth are substantial as well. Note that the means, variances and covariances of the intercepts and slopes, and the error variances of the measurements were constrained to be equal for the preschool and kindergarten cohort.

Most interesting for the present purpose, are the effects of the predictors on the slopes of literacy and math development, controlling for the intercepts. A remarkable finding, contrary to our expectations, is that there were no significant slope effects of using the education program. However, in line with our expectations, there were positive effects of both mixed arrangement and teacher-managed academic activities on the slopes of math (both in the preschool and kindergarten cohort) and literacy (only in the kindergarten cohort). The standardized regression weights in Table 3.6 reveal small to medium effect sizes.

Table 3.5 Means, variances and covariances for the structural part of the growth only model and parameter estimates of the predictors education program, mixed arrangement and teacher activity in the final model with covariates

	Emergent literacy		Emergent math	
Structural growth model				
Mean (S.E.)				
Intercept	18.39	(1.04)***	14.96	(0.77)***
Slope	9.76	(1.14)***	11.99	(0.88)***
Variance (S.E.)				
Intercept	44.93	(9.36)***	30.71	(5.35)***
Slope	12.61	(6.29)*	8.32	(2.54)***
Covariance				
Intercept-slope	17.21	(5.37)**	9.46	(2.65)***
Predictors				
Preschool cohort				
Effects on intercept (S.E.)				
Program	-		-	
Mixed	-		-	
Teacher activity	-		-	
Effects on slope (S.E.)				
Program	-		-	
Mixed	-		2.36	(0.95)*
Teacher activity	-		2.44	(0.88)**
Kindergarten cohort				
Effects on intercept (S.E.)				
Program	-		-	
Mixed	-		-	
Teacher activity	-		-	
Effects on slope (S.E.)				
Program	-		-	
Mixed	6.96	(2.38)***	2.36	(0.95)*
Teacher activity	5.35	(1.08)***	2.44	(0.88)**

Note. A dash indicates the parameter was constrained to zero.
$^{+}$ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 3.6 Standardized estimates of the significant effects on the slopes of literacy and math development

	Final model literacy		Final model math	
Effects on slope	Preschool cohort	Kindergarten cohort	Preschool cohort	Kindergarten cohort
Program	-	-	-	-
Mixed	-	0.62	0.30	0.38
Teacher activity	-	0.52	0.35	0.36

Discussion

The first goal of this study was to determine whether educationally disadvantaged children showed larger gains in emergent literacy and math skills in mixed than in targeted preschools and kindergartens that are attended in majority by disadvantaged children. The results indicate that children in mixed kindergarten classrooms gained more in literacy and math skills than children in targeted kindergarten classrooms. The positive effect of mixed classrooms was for preschool children only apparent in math skills. One of the explanations is that the share of disadvantaged children in the mixed preschool classrooms was bigger than in mixed kindergarten classrooms. Differently put, the mixed kindergarten classrooms in this study had a socioeconomically more balanced class composition than the mixed preschool classrooms. Overall, the results of the current study are in line with previous research showing that educationally disadvantaged children benefit from a more balanced socioeconomic and ethnic-cultural classroom composition, probably through classroom interactions with peers with better expressive abilities and greater vocabularies (Henry & Rickman, 2007; Mashburn et al., 2009; Schechter & Bye, 2007) and perhaps also because a more balanced classroom composition means less workload for the teacher who can spend more time to instructing and guiding children (Lee et al., 1998).

The second goal of this study was to determine the effect of a special education program on disadvantaged children's literacy and math development. The results indicate that using this program had no detectable effect on children's outcomes. However, before concluding that working with programmed activities is pointless in early childhood education, two findings should be considered. First, based on classroom observations, teachers who were trained in the education program and reported to use the program did not differ from non-trained teachers in providing the language, literacy and math activities in the classroom that were intended by the program, pointing to a lack of implementation fidelity. Second, we found important differences between teachers in the amount of time they spent on initiating and guiding language, literacy and math activities with children. Most importantly, concerning the third main finding of this study, differences between teachers in involvement in language, literacy and math activities were significantly related to child outcomes. The findings are in agreement with findings in other studies, revealing large differences between teachers working within the same program and with the same educational concept (Connor et al., 2006; Early et al., 2005; Klibanoff et al., 2006; Meyer et al., 1993). These studies similarly underscore the importance of what the teacher does (or does not) for child outcomes.

Teacher initiated and guided literacy and math activities occurred rather infrequently according to our observations, yet were significantly related to children's growth in emergent literacy and math in preschool and kindergarten. This finding suggests that even a slight

increase in the amount of time spent on these activities, might lead to gains in emergent school skills. The reasons why teachers on average spent so little time on literacy and math activities are not immediately clear. Further observations revealed that both in preschool and kindergarten much time was lost to transition activities such as gathering material to start an activity, tidying up after an activity and awaiting one's turn when choosing an activity on a plan board (de Haan, Elbers, & Leseman, 2014b). Furthermore, much time was involved in doing daily care routines such as eating snacks. Thus, inadequate classroom time management, from the point of view of promoting emergent academic skills, is a possible explanation. The large range regarding the time teachers spent on academic activities in the preschool and kindergarten classrooms in this study (preschool: 1% to 16%; kindergarten: 11% to 27%), indicates that some teachers were apparently quite able to invest a considerable amount of time in academic content, whereas other teachers failed to do so. Teachers should be urged to reconsider their daily schedules, to incorporate academic content activities in their curriculum and to minimize the loss of time due to mere waiting, transitions between activities, and lack of interesting activities. Improving classroom management skills might propel this process.

Another point to highlight is that, according to further observations, nearly all teacher-managed activities were carried out in whole group settings, despite the fact that in the classrooms working with the education program frequent small group work belonged to the prescribed activities (de Haan et al., 2014b). Presumably the whole group setting limits interaction opportunities and may therefore be less effective in stimulating development (Powell et al., 2008). Indeed, the effect of small group instruction on skill development is reported to be 10 times bigger than instruction in whole group settings (Connor et al., 2006).

In view of the policy issues raised in the introduction of this article, the present results favor a universal approach to early childhood education, allowing for more balanced classroom composition as compared to targeted programs, while making optimal use of the positive contribution of interaction with peers from non-disadvantaged backgrounds to disadvantaged children's development. The present results suggest that merely adopting an education program, however well-designed and theoretically grounded, is perhaps not a powerful enough intervention to change teachers' classroom practices. As the present results support the relevance of teacher-initiated and guided playful academic activities for child outcomes, other strategies to change classroom practice are needed. Strengthening early childhood teachers' professional skills apparently requires more than a manual, a few days of training and occasional feedback, as was the case with the education program studied here. Critical self-reflection, monitoring if planned activities are carried out, improving classroom and time management skills, also by learning from the good practices of colleagues, should be standard procedure in preschools and kindergartens.

Although it is to be recommended to replicate the study with a larger sample size to obtain more robust results, the accelerated LGM-analyses conducted in this study were certainly feasible with the current sample size (see, for example, Nevitt & Hancock, 2004). Note that the children were tested three times and that the preschool and kindergarten cohort were combined into a single sequential cohort with the measurement part of the models being constrained to be equal across both cohorts, which increased the number of data points and decreased the number of free parameters to be estimated, thereby increasing the power of the present analyses (Kline, 2005). In addition, the tested models were not extremely complex. As a matter of fact, the structural part of the models included only three predictors. Future studies should not only include larger samples, but ideally also samples from different geographical areas and ethnic communities to examine the generality of the findings of the current study.

Another limitation of the current study concerns the exclusive focus on one aspect of education quality, namely the amount of teacher-managed language, literacy and mathematics activities. Given the current interest in academic outcomes this limitation was justified, but it should be noted that several studies show that other aspects of both structural and process quality contribute to children's developmental and academic outcomes as well (Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Sylva et al., 2006). Furthermore, more detailed attention should be paid to the quality of the language, literacy and math input provided in the classrooms, instead of only the quantity as in the current study. Future studies should also include family characteristics, such as parenting practices and the availability of educational materials and activities in the home environment. This is important, because family characteristics like these show large socioeconomic and ethnic cultural differences (Bradley, Corwyn, McAdoo, & Coll, 2001) and are clearly related to child outcomes as well (Burchinal, Peisner-Feinberg, Pianta, & Howes, 2002).

Another important point to address in future studies are the effects of being in a mixed classroom for non-disadvantaged children. Although a recent study showed that high-status children's language development was not affected by their peers' language skills (Justice et al., 2011), more studies are needed to address these effects in more detail and to assess whether there are beneficial effects for these children to be in mixed classrooms surrounded by children from different cultural and socioeconomic backgrounds.

Furthermore, future research should include measures of the actual skill level of peers, for example by using standardized assessment of language skills or by observing naturally unfolding peer interaction in the classroom to test the hypothesis that disadvantaged children profit from interaction with higher-skilled peers in more detail.

Chapter 4

Classroom composition and peer verbal interaction in targeted versus mixed preschools and kindergartens as related to children's development in emergent academic skills

de Haan, A. K. E., Elbers, E., & Leseman, P. P. M.
In preparation.

Abstract

The current study examines a) children's interactions in mixed and targeted preschool and kindergarten classrooms and b) the contribution of interactions with native Dutch peers to children's development in emergent academic skills. The results indicate that although most time was spent in large groups, opportunities for children to engage in verbal interaction in this context were limited. The majority of interactions was with peers, especially in kindergarten classrooms. The amount of verbal interaction with native Dutch peers was, compared to targeted classrooms, higher in mixed classrooms. Using cohort-sequential growth modeling, the results show that the amount of verbal interaction with native Dutch peers was positively associated with kindergarten children's gains in emergent literacy skills.

Introduction

Although the importance of early learning is widely agreed upon, one of the remaining questions facing the field of early childhood education today is whether efforts aiming to promote the school readiness skills of young learners should be targeted specifically at children at risk for delays in these basic skills or whether efforts should be aimed at all young children (e.g., Barnett, 2011). One of the main reasons dividing proponents of either a targeted or universal approach are the higher costs associated with universal provisions (Barnett, 2011; Rolnick & Grunewald, 2011). A targeted approach, however, although most cost-effective (for a review, see Rolnick & Grunewald, 2011), has several potential drawbacks. Most importantly, the aggregation of children at risk for developmental delays in classrooms can have important consequences for the learning environment in these targeted classrooms. Given the strong association between exposure to language and children's language skill development (e.g., Hoff & Naigles, 2002; Hoff, 2003; Huttenlocher et al., 2002; Naigles & Hoff-Ginsberg, 1998; Zucker et al., 2013) and since most of the language input in preschools comes from peers (Dickinson & Tabors, 2001), peers are an important resource for children's language learning in preschool. However, because in targeted classrooms children at risk for delays are aggregated in groups, the language learning environment in targeted preschools might be less optimal compared to mixed preschools, serving both disadvantaged and non-disadvantaged children. These so-called direct peer effects are the first of three pathways through which the effects of classroom composition are manifested. The second pathway concerns teacher behavior. The developmental level of children in the classroom may shape teacher practices, both in terms of the quality of interactions and the teacher's ability to provide certain types of learning experiences for children. The third pathway mentioned in the literature (see, for example, Reid & Ready, 2013; Shager, 2012) concerns parent involvement, which is likely to be stronger among parents with

high socioeconomic status. Only the first pathway, direct peer effects, is addressed in the current study.

We will first review studies into the effects of a) socioeconomic/ethnic-cultural classroom composition, and b) mixed age classrooms, and continue with research showing in more detail the potential of peers in enhancing young children's language development.

Classroom composition with regard to socioeconomic and ethnic-cultural factors

Several studies have found long-term associations between the ethnic and socioeconomic composition of the *school* and children's gains in academic skills (see, for example, Aikens & Barbarin, 2008; Kainz & Vernon-Feagans, 2007; Schwartz, 2010). The few studies to date that addressed the effects of the ethnic and socioeconomic composition of the *classroom* in the preschool period yield equivocal results. In a first study, Henry and Rickman (2007) did not find effects of the proportion of minority children in the classroom on children's learning in preschool. In contrast, Lee et al. (1998) showed that children learned less in preschool care and education settings with higher proportions of recent immigrants, minorities, and children whose mothers had had little education. Similarly, previous Dutch research showed that children had smaller gains in vocabulary skills if their kindergarten classroom was composed of a high proportion of second language learners (Mayo & Leseman, 2008). In addition, Schechter and Bye (2007) showed that low-income children in mixed-income preschools, where low-income children made up about 20% of the group, showed larger gains in vocabulary than low-income children in targeted low-income preschools. Furthermore, their results showed that the low-income children in targeted low-income preschools gained at a rate less than their peers from middle- and high-income families. For these children the achievement gap actually increased rather than decreased in preschool. Similarly, a German study showed that immigrant preschool children learned most in classrooms with smaller proportions of children from disadvantaged families (regarding parental educational level, parental employment status, family composition and low proficiency in German language) (Biedinger et al., 2008). Finally, Reid and Ready (2013) found positive associations between the mean socioeconomic status of classmates and individual children's development in language and in math skills, regardless of children's own socioeconomic status and regardless of quality characteristics of their classrooms.

Another line of research on peer effects concerns universal preschool programs in the United States. These programs are deliberately aimed at all children, regardless of their ethnic or socioeconomic background. Although the positive effects of universal preschool on disadvantaged children's cognitive development are well-documented (see, for example,

Fitzpatrick, 2008; Gormley Jr. et al., 2005), studies in which disadvantaged children's development in universal preschool is contrasted with their peers' development in targeted provisions are scarce. A notable exception is a study by Henry et al. (2006) who found that disadvantaged children eligible for Head Start but attending universal preschool, showed better school readiness skills after a year than their peers in targeted Head Start programs. It should be noted, however, that despite thorough analyses to ensure statistically equivalent groups with regard to family characteristics and children's baseline language and cognitive skills, the possibility of a bias favoring children in universal preschool cannot be ruled out entirely. In contrast to these findings, Dotterer et al. (2013) found across a wide range of measures of early academic skills, while controlling for structural and process quality of the preschools, only one significant difference between children in universal and targeted preschool classrooms. Children in universal provisions showed larger gains in expressive language than children in targeted preschool. In sum, the results on universal and targeted programs are equivocal and, more importantly, none of these studies took the actual classroom composition into account. Moreover, possible mechanisms through which the ethnic and socioeconomic composition of the classroom may exert its influence, such as peer interaction, were not taken into account in these studies.

Classroom composition with regard to age

Another line of research comes from mixed-age classrooms, where age can be considered as a proxy of the level of language skills. Moller et al. (2008), for example, found that a wide range in children's ages in a classroom was negatively related to development. This finding could be an argument in favor of targeted classrooms. However, in a replication of the study by Moller et al. (2008), Bell et al. (2013) found that mixed-age classrooms were not significantly related to school readiness skills at all. Guo et al. (2014) found that mixed-age grouping affected younger and older children differently. Younger children showed larger gains in vocabulary during the academic year in classrooms with a wider range in age, whereas older children showed larger gains in classrooms with a more restricted distribution of ages.

Peers' actual ability level

Lee et al. (1998) and Henry and Rickman (2007) also took *the actual ability* of peers in the classroom into account in their studies of classroom composition effects. In the study by Lee et al. the average cognitive ability level of peers was not related to children's learning, whereas in Henry and Rickman's study (2007) evidence for a direct effect of peers was found; children's

cognitive skills, pre-reading skills, and expressive language skills were positively related to the ability level of their peers. Shager (2012) found similar results, indicating positive associations between average peer ability and individual reading and vocabulary skills. Mashburn et al. (2009) also found effects of peer expressive abilities on children's expressive and receptive vocabulary development during the preschool year. However, the authors found an unexpected interaction effect indicating that higher performing peers had a stronger positive influence on receptive language development for children who began preschool with higher receptive language skills. For low achieving children it hardly made a difference whether their peers were low, high or medium achieving. In contrast, Justice and colleagues (2011) showed that peer effects were strongest for low performing children who were in classrooms with relatively many low-skilled peers. These children, who started preschool with poor language abilities, profited from attendance of preschools with a mixture of low- and high-skilled peers. Their findings further showed that the development of high-skilled children was relatively unaffected by the average level of their peers.

In sum, although studies of peer effects in the preschool period yield equivocal results, there are indications that mixing is beneficial for disadvantaged children. None of the above described studies, and to our knowledge no other study to date, examined the actual occurrence of interactions between children and their higher skilled peers nor the contribution of these interactions to children's development. In the current study we will investigate the contribution of interactions with native Dutch peers to children's development, where it is assumed that native Dutch children have well developed language skills.

The Dutch preschool system

In The Netherlands there are separate provisions of early education and care for children under and over four years of age. Next to center-based child day care for children from zero to four years of age coming from families with both parents having a job, children can attend preschools from the age of (mostly) 2;6 onwards. In cities or neighborhoods with a high concentration of socioeconomically disadvantaged families, preschools are often exclusively recruiting disadvantaged children as part of a targeted educational priority policy. The fees are income-dependent and generally very low for disadvantaged families. Generally, children attend preschool for at least ten hours over four (sometimes five) half days a week.

The Dutch kindergarten system is integrated in the primary school system. Kindergarten starts at age four and lasts two to two-and-a-half years. Although attendance is only obligatory from age five onwards, the vast majority of children (about 98%) start attending shortly after their fourth birthday. All children attend kindergarten for five days per week (about 25 hours per week). Primary school is free of charge and since all children are eligible for

participation, the Dutch kindergarten system is an example of a universal provision. Parents enroll their children in a primary school of their choice which will most often be a school in the neighborhood where they live. Despite its universal nature, some kindergartens are attended almost exclusively by disadvantaged children due to the location of the school in neighborhoods with high concentrations of minority and low socioeconomic status families.

Research aims

The first aim of the current study is to examine children's interactions in mixed and targeted preschool and kindergarten classrooms with respect to social context and interaction partners (teacher or peers). The second aim of the current study is to examine the association between the amount of verbal interaction with native Dutch peers and children's development in emergent academic skills.

Method

Participants

The present study involved a sample of 30 preschool and 38 kindergarten children, attending one of 15 participating classrooms (7 preschools and 8 kindergartens), all located in a middle-sized town in the western part of The Netherlands. The majority of the sample are children from immigrant families, mostly with another language than Dutch as first language (see

Table 4.1 Mean age, sex, mean educational attainment level and cultural background of parents for the preschool and kindergarten cohort

		Preschool			Kindergarten		
				N			*N*
N		30			38		
Mean age time 1 (*SD*)		3.0	(0.3)		4.4	(0.2)	
Boys (%)		17	(56.7)		18	(47.4)	
Mean educational level parents		3.1	(1.9)	19	3.7	(2.1)	36
Cultural background parents (%)	Dutch	6	(20.0)		14	(36.8)	
	Moroccan	13	(43.3)		8	(21.1)	
	Turkish	2	(6.7)		3	(7.9)	
	Other/mixed	9	(30.0)		13	(34.2)	

Note. The mean educational attainment level is measured on a seven point scale ranging from 1 (only a few years of primary education at most) to 7 (university degree).

Table 4.1 for descriptive statistics). The majority of children belong to the target group of the national educational priority policy based on family background factors such as parental educational attainment level and parental cultural background. The two cohorts, consisting of preschool and kindergarten children, were included in a cohort-sequential modeling approach, thereby creating a sample spanning the age range from 3 through 6;4 years of age.

Parents of children who started within the last six months in preschool or kindergarten and met our age criteria were thoroughly informed about the longitudinal study of which the current study is a part and asked for their active informed consent. This procedure resulted in participation of the majority of eligible children (84% of preschoolers and 79% of kindergarten children).

Procedures

Child assessment

The Dutch version of the PIPS; Performance Indicators in Primary Schools (Tymms, 2001), a computerized adaptive assessment was used to assess emergent literacy and math skills. Previous research with the PIPS in The Netherlands showed good predictive validity and high reliability (Van der Hoeven-van Doornum, 2005). Children were assessed on three occasions; first shortly after entrance in preschool or kindergarten and, subsequently, one and two years later. On all occasions, assessments were divided over two test sessions of about 30 minutes. Trained research assistants used laptop computers to assess children's skills in a quiet room at preschool or kindergarten.

Classroom observations

In all preschool and kindergarten classrooms six children were randomly selected among the children participating in the longitudinal study of which the current study is a part, for participation in classroom observations. Within a two-week period the first author and trained research assistants spent four half days (about 2.5 to 3.5 hours per half day) observing classroom interactions according to a cyclic interval coding approach in all selected preschool and kindergarten classrooms. These observations were conducted shortly before the child assessments in year two of the longitudinal study. The cyclic interval coding approach that was used to observe the selected children entailed that after ten seconds of individual child observation a number of categories were coded, including the social context of the activity the child was engaged in, the involvement of the teacher and peers in the activity, and the 'interaction mode' displayed during the activity. When the coding of a particular interval was completed (which took on average about 40 seconds), a ten seconds observation of the

next child started. The cycles of observing and coding were repeated from the start until the end of the morning or afternoon. The coding scheme was programmed in E-prime (Schneider et al., 2002) and installed on laptop computers. Teachers were requested to carry out their usual activities and routines. In preschools this observation procedure resulted in a classroom average of 209 (*SD* = 32.2) observed and coded intervals during the morning and a classroom average of 128 (*SD* = 19.6) intervals during the afternoon, which lasted about one hour less than the mornings. In kindergarten, where children spent about one hour longer in the morning and thirty minutes longer in the afternoon than in preschool on average 349 intervals (*SD* = 57.5) and 209 intervals (*SD* = 40.6) were coded per classroom during respectively mornings and afternoons. This procedure resulted in approximately 35 intervals per preschool child during the morning and approximately 21 intervals during the afternoon. Per kindergarten child approximately 58 and 35 intervals were coded during respectively mornings and afternoons.

Inter-rater reliability was determined using detailed transcriptions of realistic classroom situations based on previously made video recordings. These transcriptions and the coding procedure were also programmed in E-prime (Schneider et al., 2002), which made the coding procedure equivalent to the later real time coding procedure in the classroom. On average, research assistants coded 35 different transcriptions which resulted in a satisfactory inter-rater reliability with all coders' Cohen's kappa over .80.

Child measures

Emergent literacy skills

Six subtests of the Dutch version of the PIPS (Tymms, 2001) were combined in a single score for emergent literacy skills: receptive vocabulary, writing, ideas about reading, letter identification, reading words, and reading sentences. The maximum number of items is 180, but due to the adaptive nature of the computerized assessment not all items were administered to all children; after a certain number of errors the concerning subtest ended automatically and the program continued to the next subtest. Missing scores, all missing at random as indicated by Missing Value Analysis in SPSS, were imputed separately for preschool and kindergarten children using the EM method (over the three child assessments the percentage of missing scores ranged between 3% and 30%).

Emergent math skills

Seven subtests of the PIPS (Tymms, 2001) were combined in a single score for emergent math skills: ideas about math, counting, informal sums, digit identification, shapes, adding

and subtracting, and advanced sums. The maximum number of items is 69, and again due to the adaptive nature of the computerized assessment not all items were administered to all children. Missing scores, all missing at random as indicated by Missing Value Analysis in SPSS, were imputed separately for preschool and kindergarten children using the EM method (over the three child assessments the percentage of missing scores ranged between 3% and 30%).

Observation measures

After every ten-second observation interval, the social context the child was observed in, the interaction mode the child was observed in, and the child's (interaction) partner during the observation interval were coded.

Social context

The category social context consisted of five subcategories: solitary, teacher-managed large group, teacher-managed small group, peers and a rest category. The social contexts of teacher-managed large and small group concerned activities during which a teacher is continuously involved. In small groups generally four to five children participated, whereas large groups generally involved the majority of the children in the classroom. If children were involved with one or more peers, without continuous guidance by the teacher, the subcategory peers was coded.

Interaction mode

The category interaction mode consisted of five subcategories: the child 1) shows no interaction behavior, 2) engages in private speech, 3) listens passively (without active involvement in an interaction), 4) interacts non-verbally (e.g., gesturing or waving), and 5) interacts verbally. If two or more categories occurred during one observation interval, priority was given to the 'highest' category regardless of duration within the observation interval. Priority was given because we deemed it important that the data optimally reflected the complexity of children's interactions. For example, if a child was both non-interactive and verbally interactive during one observation interval, 'verbal interactive' was coded, regardless of the duration.

Interaction with native Dutch peers

For every interval in which the child was interacting, the interaction partner of the child was coded. Based on this, we computed for each child a measure of peer interaction with native Dutch peers as the number of intervals coded as peer interaction with one or more

native Dutch peers divided by the total number of intervals coded as containing verbal peer interaction. Since observations were rather evenly divided over the course of mornings and afternoons and the number of intervals per child was quite high, percentages can be considered as accurate approximations of the actual time spending.

Classroom composition

For every classroom the proportion of native Dutch children was computed based on teacher reports.

Mixed classrooms

Based on the proportion of native Dutch children in the classroom mixed and targeted classrooms were distinguished; classroom containing fifty percent or more native Dutch children were labelled mixed, whereas other classrooms were labelled targeted (see Table 4.4 for descriptive statistics). On average, mixed classrooms consisted of 70.8% and 71.6% native Dutch children in respectively preschool and kindergarten, whereas targeted classrooms consisted on average of 18.1% and 23.1% native Dutch children in, respectively, preschool and kindergarten.

Analysis strategy

First, descriptive analyses of interaction opportunities in mixed and targeted classrooms were conducted. Second, a longitudinal cohort-sequential design with three annual child assessments in both the preschool and kindergarten cohort, with two overlapping measurements for both cohorts at age 4 and age 5 was used to examine the association between interaction with native Dutch peers and the growth in emergent academic skills. Through Accelerated Latent Growth Modeling (LGM; Duncan, et al., 2006), with AMOS (version 22; Arbuckle, 2013), the two age cohorts were combined into single models of children's emergent literacy and math development, thereby spanning the age range from 3 to 6 years. The means and the variances of the intercepts and slopes for the dependent variables were constrained to be equal across both cohorts. The estimates of the predictor variables, classroom composition and native Dutch peer interaction, on respectively intercepts and slopes were estimated. More specifically, on a theoretical basis, we assumed an association between classroom composition and children's starting level, or the intercept, and an association between actual peer interaction with native Dutch peers and children's gains in emergent academic skills, or the slope. We presupposed that classroom composition reflects selective placement due to active recruitment as part of a targeted education policy while, as

a distal characteristic, affecting the likelihood that peer learning occurs indirectly, whereas actual peer interaction reflects proximal processes in the classroom and can be regarded a direct measure of possible peer learning effects (see also Hattie, 2002; Wilkinson, 2002). We expected that a relatively high proportion of native Dutch children in the classroom would increase the likelihood that disadvantaged children have interactions with native Dutch peers. In addition, we expected the observed amount of interaction with native Dutch children to contribute to the development of children's emergent academic skills. In short, with regard to the predictor variables the analyses focused specifically on the effect of classroom composition on the intercept and on the effect of peer interaction on the slopes of the growth models of emergent academic skills.

The chi-square (χ^2), the root mean square error of approximation (RMSEA), and the comparative fit index (CFI) were used to evaluate model fit. As a rule of thumb, 1) a non-significant χ^2 or a ratio of χ^2 and the number of degrees of freedom < 2 indicates good model fit, 2) RSMEAs below .05 indicate good fit and below .08 reasonable fit, and 3) CFI greater than .95 can be considered a good fit and values greater than .90 indicate an acceptable fit (Kline, 2005).

The analyses proceeded in two steps. First, a two-group growth model was fitted for emergent literacy and math skills, respectively. Regression weights for the intercepts were fixed to 1. Regression weights for the first measurement in the preschool cohort and the final measurement in the kindergarten cohort were allowed to vary freely in order to examine non-linear growth, whereas weights for overlapping measurement times were constrained to be equal across the two cohorts and to reflect actual measurement times whenever possible (i.e., the weight for T2 in the preschool cohort and T1 in the kindergarten cohort was constrained to 0 and the weight for T3 in the preschool cohort and T2 in the kindergarten cohort was constrained to 1).[1] Similarly, several estimates were constrained to be equal for both cohorts: the measurement errors of the measures taken at the overlapping measurement times (T2 and T3 for the preschool cohort, T1 and T2 for the kindergarten cohort, respectively), the intercept means and variances, the slope means and variances, and the covariances between intercepts and slopes. See Figure 4.1 for a depiction of the accelerated latent growth model.

Second, the effects of two predictor variables were specified, while keeping the growth part of the model constant. Structural paths from the predictors – classroom composition and peer interaction with native Dutch peers – to the intercept and slope, and covariances between the predictor variables were estimated.

1 For the emergent literacy model the regression weight for the second measurement in the kindergarten cohort was constrained to 2 (as opposed to 1 in the preschool cohort). This necessary adaptation is a consequence of the steep increase that kindergarten children on average show between the first and second measurement.

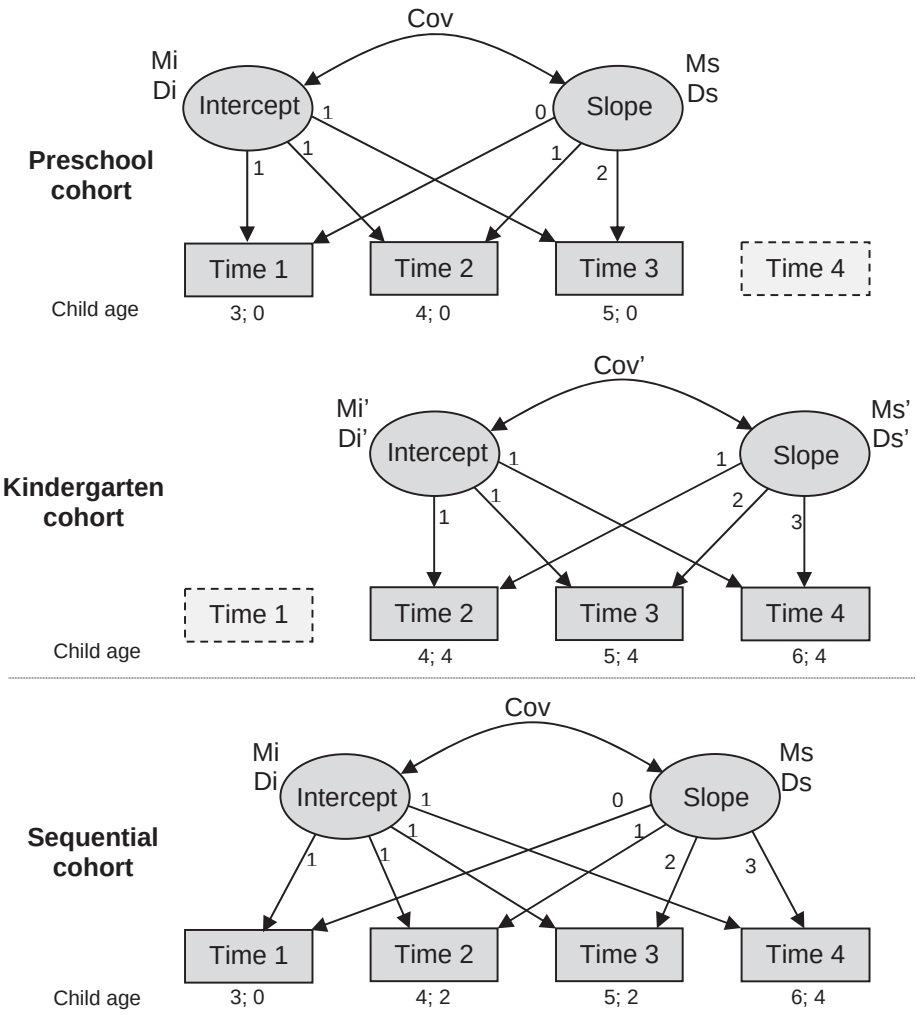

Figure 4.1 Depiction of the merging of the separate preschool and kindergarten cohort into one sequential cohort.
Note. Mi = Mean intercept, Di = Disturbance intercept, Ms = Mean slope, Ds = Disturbance slope, Cov = Covariance.

Results

Interaction opportunities and social context

Observations were conducted in seven preschool and eight kindergarten classrooms. The total number of observed intervals in preschool ranged from 396 to 783 per classroom ($M = 625.88$, $SD = 124.58$) and from 732 to 1299 ($M = 1086.63$, $SD = 171.35$) in kindergarten. In Table 4.2 an

Table 4.2 Social context and interaction mode in percentages for the preschool and kindergarten cohort

	Total		Non-interactive		Private speech		Passive listening		Interactive non-verbal		Interactive verbal	
	P	K	P	K	P	K	P	K	P	K	P	K
Solitary	16.2	9.6	81.4	82.3	5.6	10.4	2.9	2.7	1.4	0.5	8.7	4.2
Large group	41.0	49.5	44.0	31.5	2.5	3.4	23.5	46.4	7.4	2.9	22.6	15.9
Small group	0.6	1.1	0	38.1	0	9.5	83.3	19.0	0	1.2	16.7	32.1
Peers	37.4	36.2	50.9	33.1	4.3	5.0	5.7	6.6	9.1	2.5	30.0	52.7
Other	4.8	3.5	34.7	28.3	1.1	1.3	13.1	9.8	11.4	1.0	39.8	59.6

Note. P = Preschool, K = Kindergarten.

overview of the different social contexts children spent time in is shown in percentages of the total number of observed intervals. As indicated earlier, we regard the percentages as a close approximation of the observed time. Hence, we will refer to these percentages as percentages of the time spent. In both preschool and kindergarten the majority of the time was spent in large groups (preschool: 41.0%; kindergarten: 49.5%) and in unguided small groups with peers (preschool: 37.4%; kindergarten: 36.2%). Teacher-managed activities, initiated and guided by the teacher, in small groups were rarely observed (in preschool 0.6% and in kindergarten 1.1% of the total time). Concerning interaction opportunities it should be noted that large groups, although frequent in nature, provided few possibilities for active participation in verbal interaction (preschool: 22.6%; kindergarten: 15.9%); children were in large group settings mostly non-interactive (preschool: 44.0%; kindergarten: 31.5%) or passively listening (preschool: 23.5%; kindergarten: 46.4%). The second most observed context, unguided small groups with peers (preschool: 37.4%; kindergarten: 36.2%), provided more opportunities for verbal interaction than large groups (preschool: 30.0%; kindergarten: 52.7%).

As shown in Table 4.3, in both targeted and mixed preschool classrooms teacher and peer involvement in children's verbal interactions was about equal (teacher involvement was respectively 37.4% and 26.4% of the time spent on verbal interaction (direct interaction with the teacher was, respectively, 34.2% and 23.2%; interaction with the teacher through teacher involvement in peer interaction was, respectively, 3.2% and 3.2%), whereas peer involvement in interactions was respectively 38.4% and 35.5% of the time spent on verbal interaction. In mixed preschool classrooms children's interactions with peers involved far more often native Dutch peers compared to targeted preschool classrooms (85.5% vs. 30.0% in targeted classrooms; $Z = 10.10$, $p < .001$).

In both targeted and mixed kindergarten classrooms, however, the teacher was engaged in a minority of the observed interactions (teacher involvement was, respectively, 13.6% and 11.1% of the time children were engaged in verbal interaction); the majority of verbal interaction was with peers (targeted: 76.9%; mixed: 80.7%). Children in mixed kindergarten classrooms,

Table 4.3 Children's interaction partners – percentages

	Preschool		Kindergarten	
	Targeted	Mixed	Targeted	Mixed
Direct teacher interaction	34.2	23.2	13.0	10.7
Teacher involved in peer interaction	3.2	3.2	0.6	0.4
Peer(s)	38.4	35.5	76.9	80.7
Native Dutch peers	*30.0*	*85.5*	*26.9*	*70.8*
Missing	24.1	38.1	9.4	8.1

Table 4.4 Mean percent of interaction with native Dutch children opposed to the mean proportion of native Dutch children in the classroom

		Preschool		Kindergarten	
		Targeted	Mixed	Targeted	Mixed
Peer interaction Dutch		30.0	85.5	26.9	70.8
Classroom composition		18.1	70.8	23.1	71.6
	Range	0.0–42.9	58.3–83.3	4.0–38.9	50.0–90.0

however, had significantly more interactions with native Dutch peers than children in targeted kindergarten classrooms (70.8% vs. 26.9%; $Z = 18.95$, $p < .001$).

The higher proportion of interactions with native Dutch children in mixed classrooms was clearly associated with the classroom composition. In targeted preschool and kindergarten classrooms the mean percentage of native Dutch peers was respectively 18.1% and 23.1% (respective range of 0.0% to 42.9% and 4.0% to 38.9%). In contrast, in mixed preschool and kindergarten classrooms the mean percent of native Dutch peers in the classroom was respectively 70.8% and 71.6% (respective range of 58.3% to 83.3% and 50.0% to 90.0%; see Table 4.4). As expected, the correlation between peer interaction with native Dutch peers and the percentage of native Dutch peers in the classroom is strong (preschool: $r = .79$, $p < .001$; kindergarten: $r = .81$, $p < .001$).

Table 4.5 shows the correlation matrix and the means and standard deviations of the dependent variables for each age cohort (preschool cohort: above the diagonal; kindergarten cohort: below the diagonal). For each variable normality was checked. No outliers with a Mahalanobis distance > 16.27 (df = 3, $p = .001$) were detected. Skewness was for all variables and time points in the acceptable range (< | 3 |). Multivariate kurtosis was also in the acceptable range with critical ratios ranging between .09 and 1.21.

Testing growth models

Model selection proceeded in two steps; the results are presented in Table 4.6. First, a cohort-sequential growth model in which all parameters but one were constrained to be equal across cohorts[2], was fitted for emergent literacy and math skills, respectively (Model 1). As can be seen in the Table 4.6, these models yielded acceptable fit. Second, the estimates of two predictor variables were specified, while keeping the growth part of the model constant

2 For the emergent literacy model the regression weight for the second measurement in the kindergarten cohort was constrained to 2 (as opposed to 1 in the preschool cohort). This necessary adaptation finds its origin in the steep increase that kindergarten children on average show between the first and second measurement.

Table 4.5 Correlations, means and standard deviations for literacy and math skills for the preschool (above the diagonal) and the kindergarten cohort (underneath the diagonal)

								Preschool		
		1.	2.	3.	4.	5.	6.	Mean	(*SD*)	*N*
	1. Literacy T1	-	.64**	.63**	.64**	.68**	.57**	6.60	(6.16)	30
	2. Literacy T2	.71**	-	.83**	.60**	.60**	.51**	19.13	(11.42)	30
	3. Literacy T3	.69**	.67**	-	.63**	.65**	.57**	25.78	(11.83)	30
	4. Math T1	.78**	.64**	.68**	-	.73**	.59**	4.94	(4.66)	30
	5. Math T2	.64**	.66**	.66**	.80**	-	.66**	11.83	(6.76)	30
	6. Math T3	.60**	.55**	.79**	.69**	.73**	-	25.18	(7.31)	30
Kindergarten	Mean (*SD*)	23.76 (12.58)	44.93 (16.88)	107.73 (40.57)	20.79 (8.94)	34.51 (6.90)	46.42 (7.38)			
	N	38	38	38	38	38	38			

Note. Above the diagonal the data of the preschool cohort are shown, underneath the diagonal the data of the kindergarten cohort are shown.
$^{+}$ $p < .10$; * $p < .01$.

Table 4.6 Overview of the fit statistics

		χ^2	df	*p*	CFI	RMSEA
Emergent literacy	Model 1	18.10	13	.15	.97	.08
	Model 2	22.79	17	.16	.97	.07
Emergent math	Model 1	17.27	12	.14	.97	.08
	Model 2	20.18	16	.21	.98	.06

(Model 2). The estimates of classroom composition on the intercept and interaction with native Dutch peers on the slope were taken into account, the other estimates were constrained to zero. This yielded an even better fit for both models (see Table 4.6).

In Table 4.7 the non-standardized estimates of the final models for emergent literacy and math skills are shown. The means and variances of the intercept and slope were significant for both outcome measures, emergent literacy and math skills, which indicates that differences between children in the overall level of these skills are substantial, there is substantial growth in these skills over time, and that there are substantial individual differences in the rate of growth.

Most interesting for the present purpose, are the effects of the predictor variables. As mentioned before, theoretically we assumed an association between classroom composition and children's starting level, represented by the intercept, and an association between the observed amount of verbal interaction with native Dutch peers and children's gains in emergent academic skills, represented by the slope. Furthermore, based on a previous study with this sample, the estimates of predictor variables were unconstrained between the two cohorts, because the previous study revealed differential effects of mixed classroom composition for the preschool and kindergarten cohort (de Haan, Elbers, Hoofs, & Leseman, 2013). This approach led to a good model fit. Again, different associations were found for the two cohorts. In the preschool cohort only one borderline significant association was found; the percentage of native Dutch children in the classroom was positively associated with children's initial emergent literacy skills. For the kindergarten cohort, however, several significant associations were found. The percentage of native Dutch children in the classroom was positively associated with children's initial emergent literacy and math skills. Furthermore, the percentage of actual interaction with native Dutch peers in the classroom was positively associated with the gains children showed over time in emergent literacy skills, indicating that children who had more interactions with native Dutch peers showed larger gains in emergent literacy skills over time. For gains in emergent math skills no such association was found. The standardized regression weights in Table 4.8 reveal medium to large effect sizes (according to Kline, 2005).

Table 4.7 Means, variances and covariances for the structural part of the growth model and parameter estimates of the predictors peer verbal interaction and verbal interaction with native Dutch peers in the final model

	Emergent literacy		Emergent math	
Structural growth model				
Mean (S.E.)				
Intercept	15.78	(1.82)***	12.04	(1.18)***
Slope	7.05	(1.00)***	13.83	(0.90)***
Variance (S.E.)				
Intercept	62.96	(13.66)***	22.66	(4.98)***
Slope	7.97	(2.78)**	4.70	(2.56) †
Covariance				
Intercept-slope	15.43	(4.48)***	1.83	(2.43)
Predictors				
Preschool cohort				
Effects on intercept (S.E.)				
Classroom composition Dutch	.09	(0.05) †	-.01	(0.04)
Native Dutch peer interaction	-		-	
Effects on slope (S.E.)				
Classroom composition Dutch	-		-	
Native Dutch peer interaction	.02	(0.01)	.02	(0.02)
Kindergarten cohort				
Effects on intercept (S.E.)				
Classroom composition Dutch	.18	(0.04)***	.18	(0.02)***
Native Dutch peer interaction	-		-	
Effects on slope (S.E.)				
Classroom composition Dutch	-			
Native Dutch peer interaction	.06	(0.02)**	-.02	(0.02)

Note. A dash indicates the parameter was constrained to zero.
† $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 4.8 Standardized estimates of the significant effects on the slopes of literacy and math development

		Final model literacy		Final model math	
Effects of predictor variables		Preschool cohort	Kindergarten cohort	Preschool cohort	Kindergarten cohort
Intercept	Classroom composition	.27†	.56***	-.07	.73***
	Native Dutch peer interaction	-	-	-	-
Slope	Classroom composition	-	-	-	-
	Native Dutch peer interaction	.23	.53**	.26	-.27

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

Discussion

The first research aim concerned the interaction opportunities in relation to classroom characteristics. The results showed that children spent the majority of their time in large groups and in unguided small groups with peers. That children spent much time in large groups is worrisome given the limited interaction opportunities; children were mostly non-interactive or passively listening in large groups (see also de Haan, Elbers, & Leseman, 2014c). Our results showed that the second most observed context, unguided small groups with peers, provided ample opportunities for active participation in verbal interaction, especially for kindergarten children. In preschool classrooms the proportion of interactions with peers and teachers was about equal. The finding that preschool children had proportionally more interactions with their teachers than kindergarten children had, is partly related to the number of children and teachers; in preschool classrooms there are generally 15 children and two teachers, whereas in kindergarten there are generally about 20 children and one teacher. Children in mixed preschool and kindergarten classrooms had significantly more interactions with native Dutch children, which is associated with the composition of the classroom with respect to native Dutch children.

With regard to the second research aim, regarding the association between verbal interaction with native Dutch peers and children's gains in emergent literacy and math skills, we found that interaction with native Dutch children was related to kindergarten children's gains in emergent literacy skills. However, according to our results, in preschool interaction with native Dutch peers was not associated with children's development. A likely explanation is that preschool children, compared to kindergarten children, have fewer interactions with peers, that is, a smaller share of the linguistic input comes from peers (see also Winsler et al., 2002). Another factor that might be associated with a lack of effects is the lower intensity of preschool compared to kindergarten; preschool children spent only about half the time in preschool compared to the time children spent in kindergarten.

The lack of effects of interaction with native Dutch peers on both preschool and kindergarten children's gains in emergent math development is not surprising. We think it is unlikely that 'math talk' occurs more frequently in naturally unfolding peer interaction than it does in teacher language input, where math talk occurs rarely (de Haan et al., 2013; de Haan et al., 2014b; Graham et al., 1997; Howes et al., 2008; Phillips et al., 2007; Phillips et al., 2009). If teachers, however, deliberately compose groups with high and relatively low skilled children during activities, children perhaps may benefit more from interaction with their peers also with regard to math skills. This requires excellent classroom management skills, skills that have been identified earlier as important for the success of mixed classrooms (Guo et al., 2014). A study by Winsler et al. (2002) also suggests that it takes a deliberate

effort by the teacher to enable children to benefit from mixed classrooms. These authors studied a transition from two same age classrooms (three- to four-year-olds) to mixed age classrooms and found that after the transition children in the mixed age classrooms had fewer interactions with peers of a different age. These results suggest that children preferably interact with same-aged peers. Although it is unclear whether actual language skills were related to these preferences, perhaps children with well developed language skills prefer each other as playmates, thereby restricting opportunities for lower skilled peers to interact with higher skilled children. These processes should be taken into account in future studies, because they might neutralize potential positive effects of mixing.

Although the accelerated LGM-analyses were feasible with the current sample size (see, for example, de Haan et al., 2013; Kline, 2005; Nevitt & Hancock, 2004), we recommend to include a larger sample size in future studies. A limitation of the current study is that no measures of process quality were taken into account. Not only did other studies show that aspects of both structural and process quality contribute to children's developmental and academic outcomes (Burchinal et al., 2010; Sylva et al., 2006), some studies also suggest that the effects of mixed classrooms are different for classrooms with good quality process management (especially classroom management behavior) compared to classrooms with low quality process management. More specifically, Guo et al. (2014) found that children in mixed age classrooms showed larger gains in vocabulary in classrooms where teachers employ good behavior management (i.e., in classrooms in which rules and expectations for behaviors are clearly stated, few or no problematic behaviors occur, and teachers effectively redirect misbehavior as opposed to classrooms in which rules and expectations are absent, unclear or inconsistently enforced). As mentioned before, a precondition for children to benefit from higher skilled peers might be that their teacher is capable to plan and adapt activities to children's developmental level, that is, has good classroom management skills.

To our knowledge this study is the first to take measures of actual peer interaction into account in examining direct peer effects. The current study thereby adds to a growing body of studies into the effects of mixed preschool and kindergarten classrooms. Previous studies showed mixed, though mainly positive, effects of mixed classrooms on disadvantaged children's development. Since some previous studies showed mixed and different peer effects for high and low-skilled peers (Gormley Jr. et al., 2005; Guo et al., 2014; Justice et al., 2011), future studies should also take potential differences between disadvantaged and non-disadvantaged children into account. In addition, to study peer effects in more detail, we recommend to include measures of peers' academic skills. In the current study we presumed that native Dutch children have relatively well developed language skills, but a direct measure of their skills is recommendable. Furthermore, in addition to measures of peers' actual skill level it would be interesting to include detailed analyses of naturally

unfolding peer interaction in the classroom to further unravel peer effects in the preschool classroom. More precise measures may enable researchers to confirm that the language learning environment in targeted classrooms, consisting mainly of children at risk for developmental delays, is indeed less optimal than in mixed classrooms.

To conclude, for both the preschool and kindergarten cohort the classroom composition was related to children's initial skills. Interaction with native Dutch peers was only related to kindergarten children's development in emergent literacy skills. The current study adds to a growing body of research on mixed preschool and kindergarten classrooms by including actual measures of peer interaction, but much remains to be explored.

Chapter 5

Teacher language input across contexts in preschool and kindergarten: Effects on children's immediate language output

de Haan, A. K. E., Elbers, E., & Leseman, P. P. M.
In preparation.

Abstract

Video-observations of teacher-managed activities were used to explore teacher language practices in Dutch preschool and kindergarten classrooms. First, differences between a) mixed and targeted classrooms and b) small and large group activities were examined. Few differences were found between mixed and targeted classrooms. During small and large group activities, preschool and kindergarten teachers showed opposite patterns in question types and utterances' cognitive demands. Furthermore, during large groups teachers accounted for the majority of utterances and were often in prolonged interaction with only a few children. Second, the associations between teachers' language input and preschool and kindergarten children's responses were examined. Although different associations between teachers' language input and preschool and kindergarten children's responses were found, for both groups the cognitive demands of teachers' utterances were strongly related to the cognitive demands of child responses.

Introduction

Children's language skills are related to the quantity and quality of the language they are exposed to; much language input containing a wide variety of sentence structures and a rich vocabulary serves as an accelerator of language development (e.g., Hoff & Naigles, 2002; Hoff, 2003; Huttenlocher et al., 2002; Naigles & Hoff-Ginsberg, 1998; Zucker et al., 2013). In general, children from low socioeconomic status (SES) homes are less likely than peers from middle and high SES homes to be exposed to a large quantity of language input or a wide variety of sentence structures and a rich vocabulary (Hart & Risley, 1995; Hoff, 2003; Zucker et al., 2013). Preschool education is particularly important for this group of disadvantaged children. Generally, it is acknowledged that preschool education needs to be of high quality in order to have any effect on children's development (see, for example, Bryant et al., 1994; Burchinal et al., 2000; Howes et al., 2008; Mashburn et al., 2008; Peisner-Feinberg et al., 2001). Usually, when teacher-child interactions are taken into account in research on the quality of early childhood education, global rating measures are used to determine the quality of these interactions. Often, these studies find that teacher language strategies deemed important for children's language development (such as frequent conversation, open-ended questions, repetition and extension, advanced language), are used infrequently (Cabell, DeCoster, LoCasale-Crouch, Hamre, & Pianta, 2013; Justice et al., 2008b). Despite the studies of teacher-child interaction that are carried out to date, much remains to be explored on teacher language practices in the preschool and kindergarten classroom.

In an earlier study (de Haan et al., 2014b) we shed light on the activities that teachers and children are engaged in during the preschool or kindergarten day. In order to gain a

better understanding of what children are exposed to in terms of language practices and to gain solid directions for professional development interventions, in this study, we will explore teachers' language practices in depth and subsequently relate teachers' language practices to children's immediate language output. Video-observations of teacher-managed activities in large and small groups are used to first provide general descriptive statistics of the types of utterances preschool and kindergarten teachers use and the cognitive demands (or distancing level) of questions and statements. Subsequently, we will compare two contexts and explore whether there are differences in the types of questions posed and the cognitive demands of the most frequently observed utterances (statements, open-ended and closed questions). The first context concerns the composition of the classroom with regard to the proportion of disadvantaged children and the second context concerns the number of children involved in an activity. In addition, we will examine whether small group activities provide more opportunities for children to make a verbal contribution to the activity than large group activities. The last part of this study concerns effects of characteristics of teacher language input on children's immediate language output.

Teacher language input and its effect on children's language output and skills

To the best of our knowledge, few studies have extensively reported on preschool and kindergarten teachers' types of utterances. Several studies, focusing on teacher questioning, have shown that generally, the majority of teacher utterances were statements, whereas about one third of teacher utterances were questions (de Rivera, Girolametto, Greenberg, & Weitzman, 2005; Durden & Rainer Dangel, 2008; Gest, Holland-Coviello, Welsh, Eicher-Catt, & Gill, 2006; Massey, Pence, Justice, & Bowles, 2008; Tompkins, Zucker, Justice, & Binici, 2013; Zucker, Justice, Piasta, & Kaderavek, 2010). Due to their potential in fostering children's language development, teachers' questioning is often addressed in more detail. In these studies the focus is often on the question form (e.g., open-ended or closed questions) (see de Rivera et al., 2005; Girolametto, Weitzman, van Lieshout, & Duff, 2000).

Several studies have shown that teacher questions concerned mainly closed questions or requests for simple yes/no responses; few open-ended or choice questions were asked (de Rivera et al., 2005; Lee & Kinzie, 2012; Powell et al., 2008; Siraj-Blatchford & Manni, 2008). Few studies of teacher questioning took child responses into account. A first exception concerns a study by de Rivera et al. (2005) in which the effects of closed, open-ended, topic-continuing and topic-initiating questions on children's answers were examined. They found that preschool children (over 30 months of age) used more multi-word utterances following open-ended and topic-continuing questions than following closed questions. It

is important to note, however, that in this study - as seen in other studies - about 73% of all questions were closed, thereby presumably limiting opportunities for children to use multi-word sentences. These effects were not found for younger children (18 to 30 months of age); the question type did not affect the length of younger children's utterances. In addition, a study by Girolametto et al. (2000) in toddler and preschooler groups in day care showed that the use of open-ended 'wh-questions' was positively associated with children's immediate language output in terms of number of utterances, number of different words and number of multi-word combinations. Also, Lee and Kinzie (2012) found indications that open-ended questions elicited more complex utterances with regard to sentence structure than closed questions in preschool children.

Intervention studies aiming to improve teachers' language stimulation strategies, such as asking open-ended questions, also suggest that open-ended questions are positively associated with preschool children's language output (number of utterances and multi-word combinations) (e.g., Girolametto, Weitzman, & Greenberg, 2003) and long-term language learning (Justice et al., 2008a; Peterson, Jesso, & McCabe, 1999; Wasik, Bond, & Hindman, 2006). The positive impact of open-ended questions on children's immediate language output and language skills development on the long-term might be caused by the higher cognitive demands these questions place on children's thinking. For example, a question like 'What is the cat doing?' requires a different type of response from children than a question like 'Is that a cat?' and places higher cognitive demands on children. Although closed questions do not necessarily restrict a child to a one word answer, open-ended questions may encourage the child to choose from several options and in that way stimulate children's thinking and learning processes, expand their thinking and help develop abstract language skills such as predicting, analyzing, and inferring (Blank, Rose, & Berlin, 1978; Wasik et al., 2006). Several studies have shown that, in general, teachers use relatively little cognitively demanding talk (including open-ended questions), also referred to as decontextualized, non-immediate, representational or inferential language (Durden & Rainer Dangel, 2008; Gest et al., 2006; Zucker et al., 2010). Instead, teachers' attempts to promote joint interaction often include rhetorical or test questions that do not invite extended conversation, and place few demands on the children to use language to explain, predict or provide information. Zucker et al. (2010) found that children's responses to questions with a high cognitive demand were on average longer than utterances in response to cognitively less challenging questions. Moreover, teacher questions were more likely than by chance to be followed by a child response at the same level of cognitive demands (see also Danis, Bernard, & Leproux, 2000). Similar results were found in a study of small groups play sessions (Tompkins et al., 2013). However, in the latter study no relation between the cognitive demand of teachers' questions and the length of children's responses was found. In addition to immediate effects of cognitively

demanding questions on child responses, several studies found long-term positive effects of these questions on children's emergent literacy and vocabulary skills (Hindman, Connor, Jewkes, & Morrison, 2008; Hindman, Wasik, & Erhart, 2012). To conclude, generally, teachers ask mostly closed questions, whereas open-ended and cognitively demanding questions are related to more complex immediate child language output and long-term language development.

Teacher-managed activities in large versus small groups

Although some studies focused on the variation of teachers' language use across classroom contexts such as book reading and free play, few studies have made a comparison between teacher-managed activities in large and small groups, although there are indications that the quality and quantity of teacher language input may vary according to group size (Morrow & Smith, 1990). Smith and Dickinson (1994), for example, found that teachers showed more didactic talk, characteristic of language routines such as counting, reciting, and discussing classroom rules, during large group activities than during small group activities. In addition, they found that cognitively demanding talk was also observed less often during large group activities than during small group activities. Similarly, Pence Turnbull, Anthony, Justice, and Bowles (2009) found that the frequency of language stimulation techniques such as using open-ended questions was lower when the group size was large as opposed to small. Pellegrino and Scopesi (1990) found that teachers showed fewer utterances aiming to initiate or continue verbal interaction with children in large versus small groups.

Other indications that teacher language input varies according to group size comes from studies comparing small group activities to one-to-one interaction. Kontos and Keyes (1999), for example, observed more complex language interactions between teachers and children in one-to-one interactions than in situations when a teacher is working with a group of children. These authors showed that teachers were more likely to extend or elaborate on children's social bids and engaged more in prolonged conversations when they worked individually with a child, a finding also reported by Schaffer and Liddell (1984). The latter authors also found that children received more language input and less directive and prohibitive language during activities in which they were alone with a teacher than during small group activities (Schaffer & Liddell, 1984). In contrast to these earlier findings, however, Marinac, Ozanne, and Woodyatt (2004) found that teachers' language was more complex (higher mean length of utterance (MLU), elaborate lexicon, more complex syntax) in groups with two or more children as compared to one-to-one interactions.

In a previous study (de Haan et al., 2014b) we found that nearly all teacher-managed activities were carried out in large groups. Other studies, however, suggest that large groups

limit interaction opportunities; children in large groups are more likely to be passively listening or merely watching than talking and acting (de Haan et al., 2014a; Powell et al., 2008). This is not unexpected, since the teacher has to divide his or her attention between many children and usually one child at a time is able to contribute to the ongoing activity. Opportunities for prolonged joint interactions and teacher language adjusted to individual children's skill level therefore, are limited during large group activities (Girolametto et al., 2000; Powell et al., 2008; Schaffer & Liddell, 1984). However, several studies have shown that during small group activities children are also in the majority of instances listening to their teacher (Aukrust, 2008; Durden & Rainer Dangel, 2008; Girolametto et al., 2003; Lawton & Fowell, 1989). Moreover, Girolametto et al. (2000) showed correlational evidence that teacher dominance was not only negatively associated with the number of children's utterances, but also with the complexity of children's utterances (number of different words and number of multi-word combinations). Morrow and Smith (1990) also found that children expressed more comments and questions during teacher-managed story book reading in small groups and one-to-one interactions than in large groups (see also Phillips & Twardosz, 2003). Moreover, presumably due to the abovementioned reasons story book reading in large groups was, compared to one-to-one and small group reading, least effective in enhancing children's story book comprehension. A finding that might be true for large group activities in general (Connor et al., 2006). In short, research allows the general conclusion that large group activities provide less opportunities to use (complex) language than small group activities.

Targeted versus mixed classrooms

Generally, two main types of interventions are distinguished in early childhood education. First, interventions that are targeted at children most in need. In the Dutch context targeted interventions are generally aimed at children from families with non-western backgrounds and/or low educational attainment. In contrast, universal, or mixed interventions are deliberately aimed at all children, regardless of background. Targeted interventions seem most cost-effective since resources are aimed at a smaller number of children (Heckman, 2006).

Since targeted provisions serve almost exclusively children at risk for delays in basic school readiness skills, there may be negative consequences for the classroom learning environment. The accumulation of disadvantaged children in groups may exert its potential negative consequences through at least two pathways. First, a high concentration of children at risk for delays may negatively affect the language input of peers in such classrooms, because input from higher skilled peers with larger vocabularies and better expressive language abilities is lacking. For this first pathway is only supporting evidence through

studies into the composition of the classroom with regard to 1) socioeconomic and ethnic-cultural background (Biedinger et al., 2008; Dotterer et al., 2013; Henry et al., 2006; Henry & Rickman, 2007; Lee et al., 1998; Mayo & Leseman, 2008; Reid & Ready, 2013; Schechter & Bye, 2007), 2) age (Bell et al., 2013; Guo et al., 2014; Moller et al., 2008), and 3) actual peer ability levels (Henry & Rickman, 2007; Justice et al., 2011; Lee et al., 1998; Mashburn et al., 2009; Shager, 2012). Overall, these studies showed mixed, though mainly positive effects of mixed classrooms on children's development. The second pathway through which peer effects may exert their influence concerns teacher behavior and is addressed in the current study. Since teachers may adapt the level of their activities and language input to children in their classroom, as a result the cognitive and linguistic demands that are placed on children may be lower in targeted classrooms as compared to mixed provisions (Girolametto & Weitzman, 2002; Justice et al., 2011). Indications for the second pathway, teacher behavior, stem predominantly from studies in which children's age or language skills were taken into account in exploring teacher language practices. Girolametto and Weitzman (2002), for example, found that child care providers 'matched' their utterances to children's language skills; children with advanced language skills received another type of language input from their teacher than children who produced single words or simple two word sentences. Similar results were found in a recent study by Justice, McGinty, Zucker, Cabell, and Piasta (2013). They found that the syntactic complexity of teacher and child utterances were mutually dependent; a complex teacher utterance was more likely to be followed by a complex child utterance and vice versa. This finding implies that teachers in classrooms with a large share of children with advanced language skills use relatively many syntactically complex utterances. Furthermore, their results indicated that the likelihood that a child immediately mirrors the teacher's syntactically complex utterance is greatest in classrooms with low average language levels. This finding underlines that especially for second language learners complex language input is an important resource for children's language development.

Similarly, Girolametto et al. (2000) found that child care providers showed a lower mean length of utterance (MLU) in groups which included a child with language delay, as compared to groups consisting of normally developing children. In addition, De Rivera et al. (2005) showed that teachers placed less cognitive and linguistic demands on young children; teachers asked proportionally slightly less open-ended questions with toddlers (18 to 30 months of age) than preschoolers (over 30 months of age).

No studies of which we are aware examined differences in the characteristics of teachers' language input between targeted and mixed classrooms. Studies in which children's age or language skills were taken into account, however, do show indications that teachers adapt their language practices to the ability level of children in their classroom.

The current study

For the current study, video-observations were made during small and large group teacher-managed activities in both mixed and targeted preschools and kindergartens in The Netherlands. Dutch pre-primary schooling consist of two systems; preschool for children under four years of age and kindergarten, which is integrated in primary school, from four years onwards, thereby spanning the age range of (mostly) 2;6 to about six years of age (kindergarten lasts two to two-and-a-half years). In the municipality where the current study was conducted, children start preschool from age 2;6 and parents pay a small income-dependent fee. At the time this research was carried out about 60% of eligible disadvantaged children participated four half days per week (about 2.5 hours per day). In The Netherlands, eligibility for preschool is defined by municipalities and in the municipality under study indicated by family background factors such as parental educational attainment level and parental cultural background. The vast majority of children – over 98% – start in kindergarten at or shortly after their fourth birthday (kindergarten is compulsory from age five onwards). Children attend kindergarten about 25 hours per week divided over five days.

At the beginning of the 21st century the renewed attention for preschool as a means of promoting disadvantaged children's skills has led to the transformation of preschool (or play groups) into 'educational preschool' targeted exclusively at children at risk for developmental delays, generally indicated by their parents' educational and cultural background. In the current study also mixed educational preschools were included, in which groups consist of both disadvantaged and non-disadvantaged children. Although the Dutch kindergarten system is universal in nature, since all children are equally eligible for participation, some kindergartens are attended almost exclusively by children at risk for delays, mostly due to the location of the school in neighborhoods with a large share of disadvantaged families. Especially in these neighborhoods most schools have implemented an education program aiming to accelerate disadvantaged children's development.

In addition to providing general descriptive statistics of the types of utterances preschool and kindergarten teachers use during teacher-managed activities and the cognitive demands, or distancing level, of questions and statements, we will attempt to answer two main research questions. First, do mixed versus targeted preschools and kindergartens and small versus large group activities differ in a) the type of questions teachers pose; and b) the cognitive demands, or distancing level, of teachers' statements and questions? In addition, we will determine whether the opportunities for children to make a verbal contribution to the activity are greater during small group activities than during large group activities. Second, is teacher language input related to children's immediate language output: a) is the type of question related to the length of children's responses; b) are cognitively demanding

utterances associated with longer responses in preschool and kindergarten children; c) is the cognitive demand or distancing level of teacher utterances associated with the distancing level of child responses?

Method

Participants

The participants in this study were 6 preschool and 6 kindergarten teachers and the children in their classroom in the area of Utrecht, The Netherlands, working in respectively four preschools and four kindergartens. All preschool teachers completed at least an education at the intermediate vocational training level (one teacher completed higher vocational training). All kindergarten teachers completed higher vocational training (bachelor). All teachers were female, had on average 6.5 years (SD = 4.2) of experience (preschool: M = 7.0, SD = 3.08, range 3–10 years; kindergarten: M = 6.0, SD = 5.5, range 0–15 years) and were responsible for curriculum planning in their respective classrooms.

All preschool groups consisted of 15 children in the age range of 2;6 to 4 years and were led by two qualified teachers. Kindergarten classrooms consisted of 13 to 23 children (M = 18.75, SD = 4.65) and were led by one teacher. Based on the proportion of native Dutch and disadvantaged children in the classroom mixed and targeted classrooms were distinguished. In targeted classrooms on average 17% (SD = 17.3) (range 0% to 36%; preschool: 0–36%; kindergarten: 4–26%) of the children were native Dutch. The other children were mainly Moroccan-Dutch, Turkish-Dutch or from other cultural origin. In contrast, mixed classrooms had a comparatively high representation of native Dutch children (M = 28%, SD = 23.1), the range was 14% to 63% (preschool: 14–15%; kindergarten: 20–63%). In each of the eight classrooms (four preschool and four kindergarten) four children were selected among the children participating in the longitudinal study of which the current study is a part. We aimed to randomly select two boys and two girls per classroom and in mixed classrooms in addition, we aimed to randomly select two disadvantaged and two non-disadvantaged children.[1] Whether children were considered disadvantaged is indicated, as determined by the municipality, by parental educational attainment level and parental cultural background. At the time of the video-observations, the selected preschool children were on average 3.66 years old (SD = 0.24, range 3.17 to 4.00 years), the selected kindergarten children were

1 The final selection differed from our aims in the two mixed kindergarten classrooms due to the composition of the group of children participating in the longitudinal study; in one mixed kindergarten three disadvantaged and one non-disadvantaged child was selected and in the other mixed classroom three girls and one boy were selected.

on average 5.18 years old (*SD* = 0.21, range 4.58 to 5.33 years). The preschool children in targeted preschools, all disadvantaged, attended preschool four half days per week. The non-disadvantaged children in mixed preschools attended preschool for two half days per week; disadvantaged children in mixed preschools attended four half days (as in targeted preschools). All kindergarten children attended five days, about 25 hours per week. All children attended preschool or kindergarten for about ten months prior to the study.

Procedure

At the beginning of the longitudinal study, parents of children who started recently in preschool or kindergarten – within the past six months – were informed by letter, brochure and occasionally a group meeting. The majority of parents gave active informed consent for their child's participation in the longitudinal study (84% of preschoolers and 79% of kindergarten children; see below for the procedure regarding the video-recordings).

Video-recordings

The video-recordings took place in the fall of 2008. Parents of all children in the classroom were informed by a letter about the video-recordings and gave passive or active informed consent depending on the (pre)school policy (in some (pre)schools parents were asked permission for their child's participation in video-recordings at enrollment). The first author worked with research assistants to make the video-recordings of large and small group teacher-managed activities and during free play sessions (not reported in this study). We aimed to record two large group activities and two small group activities per classroom. Teachers were requested to carry out their usual program and activities, but were informed that we were aiming to record small and large group teacher-managed activities during our visits. Mostly, the video-recordings were made on two separate days within a two week period. Teachers were informed that the purpose of the study was to observe adult and child communication during teacher-managed small and large group activities. We captured, regardless of length, the complete activity. All classrooms were first visited on a morning or afternoon to make pilot recordings in order to familiarize teachers and children with the procedure and equipment. To gather qualitatively good and audible recordings, two wireless microphones were used. One of the microphones was worn by the teacher, the other microphone was worn by one of the children. The microphone for the children was hidden in a jacket made of jeans material (picture available from the first author). In order to make the jacket appealing to children, the child that wore the jacket was allowed to choose one of three Velcro patches (Cars, Bob the Builder or a 'girly' pink one with an unfamiliar comic character depicted on it) that could be attached to the front of the jacket. Two of the selected

children refused to wear the jacket. Instead, for both children a classmate also participating in the longitudinal study was included in the current study.

Measures and reliability

Transcription and coding

All video-recordings were transcribed using an adaptation of the CHAT conventions. Transcripts included all teachers' and children's utterances and descriptions of actions that were necessary to understand the course of the interaction. If a teacher, for example, said 'Stop that!' it is necessary to know what sort of behavior the child had to stop in order to understand the teacher's utterance. In addition, all non-verbal actions (like nodding, putting up a finger, etcetera) relevant to the interaction were transcribed into the file. All transcripts were made in SPSS 17.0. Transcription in SPSS allowed us to code a number of variables for each utterance or non-verbal action: the agent, the type of action (utterance or non-verbal action), the recipient, the activity, the social context, cognitive demands or distancing level and, if the action was an utterance, the type of utterance. Type of utterance and distancing level were not coded simultaneously with the transcription and coding of other variables, but afterwards when the transcript was completed and checked by the assistant.

Four research assistants made respectively 25, 2, 2 and 1 transcripts. The assistants who made 2 or less transcripts participated in the transcription of the free play situations that were also recorded (not included in the current study) and therefore, had gained as much experience as the assistant who made the majority of the transcripts for the current study. All assistants participated in a thorough training which started with a group instruction by the first author and was followed up by extensive proof transcriptions. Finally, once the research assistant completed the transcripts, the first author verified for each assistant at least 20% of the transcripts to ensure accurate representations of the recorded interactions (in case an assistant transcribed two or less teacher-managed activities all transcripts were verified). Every line in the transcript was checked against the video-recording and disagreements were noted. Agreement was found for over 90% of all lines in the checked transcripts.

Utterance

In the transcription process two rules of thumb concerning utterances were that a) the speaker's intonation often marks the end of an utterance and b) utterances often have a meaning in itself (except when an utterance is aborted because the speaker stops talking or is interrupted, then the (incomplete) utterance may be 'meaningless'). The latter rule of thumb is mainly applied if a speaker showed more than one utterance in one turn. For example, a

child telling about a birthday party shows consecutive utterances that are distinguished by sentence structure and meaning. Conjunctions are important in distinguishing utterances. The conjunction 'and', for example, often signals a new utterance, except when the verb is the same. The utterance 'My presents were a car and a book' was transcribed as one utterance. However, the sequence 'First we ate cake and candy, and then we went to a museum and I liked the party a lot!', was transcribed in three separate utterances (1. first we ate cake and candy; 2. and then we went to a museum; 3. and I liked the party a lot!). The sequence 'I liked the museum a lot, because I saw a dinosaur skeleton' was transcribed in one utterance, because the second part of the utterance is clearly related to the first; the second part of the utterance is 'meaningless' without the first part of the utterance. A 'Yes' or 'No' in response to a question is also considered an utterance.

Social context

We coded the social context of each line in the transcript. In small group activities a selection of children participated, generally about 4 to 5. Large group activities involved a larger group of children (between 10 and 23). Transcripts were coded as either a small group or large group activity. There was 100 percent agreement in the coding of social context.

Type of utterance

For all utterances the type of utterance was coded (for an overview of the coding scheme see Appendix 5.1). For utterances reflecting attention gaining (e.g., 'Hello?', 'Look!', mentioning a name, etcetera), expression (e.g., 'Wow', 'Whoops', 'So what', etcetera), affirmation, and negation a hierarchical coding rule was used in the coding process; if these types of utterances occurred simultaneously with another type of utterance in one utterance, the latter type of utterance overruled the former. For example, the utterance 'Yes this giraffe has a very long neck', where an affirmation occurs with a statement within one utterance, was coded as a statement.

Type of utterance was coded by the same assistant who made the transcript, so four research assistants coded respectively 25, 2, 2, and 1 transcripts. At least 16 percent of transcripts were double coded by the first author (for the assistants who coded two or less transcripts all transcripts were double coded). Reliability as assessed with Cohen's kappa was overall .93 (ranging from .91 to .96 for individual assistants) and therefore, considered very good. Percent of exact agreement was overall 94.8% (ranging from 93.4% to 97.4%), whereas the agreement by chance was 31.1% (ranging from 23.8% to 33.3% for individual assistants). The intra class correlation was overall .83 (ranging from .82 to .90).

Cognitive demands or distancing level

The four level framework of Blank et al. (1978) was used to code the cognitive demands of utterances. Due to reliability difficulties, the four level framework was, as in other studies (e.g., Massey et al., 2008), adapted to a two level framework. The underlying continuum of this scale concerns the distance between perception and language and therefore, the levels are also referred to as 'distancing levels'. Utterances coded as low in distancing are utterances focusing on perceptual characteristics or on a specific characteristic of what the child perceives or recently perceived and involve labeling and locating objects or characters (typical examples are 'What is this?' while pointing at an image or 'What color is this umbrella?'). In sum, language coded as 'low distancing' is given in perception, based on direct of recent perception, and an immediate shared context is a prerequisite. If a child, for example, says 'My bedroom door is red' there is no shared context and, therefore, this utterance should be coded as high in distancing, or cognitively demanding. If a child, however, says 'The classroom door is red', then there is a shared context and the utterance should be coded as low in distancing. In short, the second, or highest level of cognitive demands, requires children to distance themselves from their direct perceptual experiences and manipulate their experiences. Examples are summarizing, defining, comparing, contrasting, referring to a plan or a strategy, reasoning, problem solving, and predicting. Typical questions are why-questions.

Three research assistants and the first author coded respectively 14, 6, 4, and 5 transcripts. The procedure, in which distancing was coded after the transcription was completed, allowed us to have other assistants than the one who made the transcript to code the distancing variable. This was the case for nineteen transcripts. At least 21 percent of transcripts were double coded by the first author. Reliability as assessed with Cohen's kappa was overall .59 (ranging from .45 to .73 for individual assistants) and therefore, considered moderate. Percent of exact agreement was overall 79.2% (ranging from 71.0% to 87.0%), whereas the agreement by chance was 50.8% (ranging from 47.5% to 55.0% for individual assistants).

Length of utterance in response

As in other studies (e.g., de Rivera et al., 2005; Girolametto et al., 2003; Zucker et al., 2010) the length of utterance was used as a measure of complexity of utterances. For every utterance in the transcript the number of words was computed and inserted at the respective line in the SPSS-file. Expressions like 'uh' and 'oh' within an utterance were excluded from the number count.

Results

We aimed to record two large group activities and two small group activities in each classroom. In one preschool and in one kindergarten classroom we were unable to schedule a second video recording of a small group activity (see Table 5.1 for an overview).

The mean length of large group activities was 16.76 minutes (SD = 5.53) (preschool: M = 14.28, SD = 4.22; kindergarten: M = 19.24, SD = 5.80). The mean length of small group activities was 15.61 minutes (SD = 5.96) (preschool: M = 12.91, SD = 5.21; kindergarten: M = 18.32, SD = 5.72). An ANOVA with context and cohort (preschool or kindergarten) as factors revealed a large main effect of cohort on the length of activities (F (1, 29) = 7.26, η^2 = .22, p = .01); activities in both large and small groups lasted shorter in preschool classrooms. This difference is not unexpected given the younger age and shorter concentration span of preschool children. The recorded activities concerned mainly language activities (see Table 5.2 for an overview).

Teachers' types of utterances and cognitive demands

All teachers' utterances addressed to one or more children in the classroom were coded for type of utterance. As shown in Table 5.3 most teacher utterances concerned statements (overall: 44.9%; preschool: 43.8%; kindergarten: 45.9%). The second most observed type of utterance were closed questions (overall: 14.6%; preschool: 16.9%; kindergarten: 12.5%). Open-ended questions yielded similar percentages as closed questions (overall: 13.2%; preschool: 15.2%; kindergarten: 11.4%) and together these three categories (statements, closed and open-ended questions) made up over two third of all teacher utterances. Preschool and kindergarten teachers' proportions of utterances differed significantly for several types of utterances (statement (Z = -1.99, p < .05); request (Z = -2.74, p < .01); imitation (Z = 2.63, p < .01); imitation with expansion (Z = 4.59, p < .01); reading (Z = -12.83, p < .01)). However, the differences between preschool and kindergarten teachers were most apparent for closed questions and open-ended questions; preschool teachers used relatively more closed and

Table 5.1 Overview of the number of small and large group teacher-managed activities in preschool and kindergarten

		Preschool			Kindergarten		
		Targeted	Mixed	Total	Targeted	Mixed	Total
Social context	Large group	4	4	8	4	4	8
	Small group	3	4	7	4	3	7
	Total	7	8	15	8	7	15

Table 5.2 Overview of the observed activities in large and small groups in preschool and kindergarten classrooms

		Preschool			Kindergarten		
		Large group	Small group	Total	Large group	Small group	Total
Activity	Fantasy play					1	1
	Book reading				3		3
	Academic language/literacy activities	3	5	8	3	4	7
	Academic math activities				1	1	2
	Other academic activities/routines	5	2	7	1	1	2
	Total	8	7	15	8	7	15

Table 5.3 Type of teacher utterance addressed to one or more children

	Overall		Preschool		Kindergarten	
	N	%	*N*	%	*N*	%
Attention/expression	349	3.8	158	3.6	191	3.9
Statement	4178	44.9	1913	43.8	2265	45.9
Request (imperative)	184	2.0	68	1.6	116	2.3
Closed question	1354	14.6	737	16.9	617	12.5
Multiple choice question	37	0.4	19	0.4	18	0.4
Completion question	143	1.5	77	1.8	66	1.3
Open-ended question	1226	13.2	662	15.2	564	11.4
Affirmation w/o expansion	393	4.2	166	3.8	227	4.6
Negation w/o expansion	69	0.7	32	0.7	37	0.7
Imitation	538	5.8	282	6.5	256	5.2
Imitation with expansion	146	1.6	96	2.2	50	1.0
Reading	445	4.8	77	1.8	368	7.5
Rest/incomplete/unintelligible	242	2.6	79	1.8	163	3.3
	9304		4366		4938	

open-ended questions than kindergarten teachers (closed questions: 16.9% vs. 12.5%, $Z = 5.99$, $p < .01$; open-ended questions: 15.2% vs. 11.4%, $Z = 5.32$, $p < .01$).

Because the majority of teachers' utterances concerned statements, closed and open-ended questions, we examined the cognitive demands of these types of utterances. As Table 5.4 shows, there was a weak association between these types of utterances and cognitive demands (preschool: χ^2 (2, $N = 2457$) = 25.91, $p = .00$, Cramer's $V = .10$; kindergarten: χ^2 (2, $N = 2280$) = 20.19, $p = .00$, Cramer's $V = .09$). It is apparent that, generally, kindergarten teachers' utterances were more often cognitively demanding than preschool teachers' utterances (see Table 5.4); kindergarten teachers' statements, closed and open-ended questions were in

Table 5.4 Distancing level of teachers' statements, closed and open-ended questions

		Distancing			
		Low		High	
		%	*N*	%	*N*
Preschool	Overall	50.0	1229	50.0	1228
	Statement	51.8	647	48.2	602
	Closed question*	41.1	240	58.9	344
	Open-ended question*	54.8	342	45.2	282
Kindergarten	Overall	38.8	884	61.2	1396
	Statement*	42.7	564	57.3	758
	Closed question*	32.8	149	67.2	305
	Open-ended question*	33.9	171	66.1	333

Note. * = the column proportions differ significantly from each other at the .05 level.

majority cognitively demanding, whereas the pattern was less clear for preschool teachers. Preschool teachers' closed questions were significantly more often cognitively demanding, whereas open-ended questions were significantly more often low in cognitive demand.

Subsequently, a direct logistic regression analysis was performed separately for the preschool and kindergarten cohort with cognitive demands as dependent variable and utterance (statement, open-ended and closed question) as predictors. A test of the full model with all predictors against a constant-only model was statistically significant for both cohorts (preschool: χ^2 (2, N = 2457) = 26.02, p < .001; kindergarten: χ^2 (2, N = 2280) = 20.33, p < .001), indicating that type of utterance was related to cognitive demands. The variance in cognitive demands accounted for was small, however, with Nagelkerke R^2 = .01 for the preschool and kindergarten cohort. Classification was unimpressive (preschool: 50.0%; kindergarten: 54.3%). Table 5.5 shows regression coefficients, Wald statistics, odds ratios for each of the dummy coded variables of the predictor variable with open-ended questions as reference category. Unexpectedly, in preschool classrooms both statements and closed questions were more likely to be cognitively demanding than open-ended questions, although the effect was not significant for statements. In kindergarten classrooms open-ended questions were more likely to place high cognitive demands than closed questions, but significantly more likely to place high cognitive demands on children than statements.

Table 5.5 Logistic regression analysis with distancing level as dependent variable and utterance (statement, closed and open-ended question) as predictor

		β	*SE* β	Wald's χ^2	df	*p*	e^{β} (odds ratio)
Preschool	Constant	-.19	.08	5.75	1	.02	.83
	Statement	.12	.10	1.51	1	.22	1.13
	Closed question	.55	.12	22.57	1	.00	1.74
	Open-ended question*			25.69	2	.00	
Kindergarten	Constant	.67	.09	50.19	1	.00	1.95
	Statement	-.37	.11	11.52	1	.00	.69
	Closed question	.05	.14	.13	1	.72	1.05
	Open-ended question*			20.10	2	.00	

Note. * = reference category.

Differences between two contexts in the types of questions and cognitive demands of teacher utterances

Mixed versus targeted preschools and kindergartens

To test whether there are differences between teachers in mixed and targeted provisions in the types of questions and cognitive demands of their utterances a chi squared test was conducted. Although a cross table reveals that the proportion of open-ended questions was slightly higher in mixed than in targeted preschools, there was no significant relationship (see Table 5.6, χ^2 (1, N = 1399) = 0.78, p = .38, Cramer's V = .02). Although also not significant, in kindergarten the pattern was opposite; the proportion of open-ended questions was slightly higher in targeted than in mixed kindergartens (see Table 5.6, χ^2 (1, N = 1181) = 0.47, p = .50, Cramer's V = -.02). With regard to distancing, we found that in mixed preschools statements and questions were generally more often of a high distancing level than in targeted preschools (χ^2 (1, N = 2550) = 37.81, p = .000, Cramer's V = .12). The association was weak, however. In kindergartens, no such relationship was found (χ^2 (1, N = 2356) = 1.46, p = .23, Cramer's V = .03).

Large versus small group teacher-managed activities

To test whether there are differences between teacher-managed activities in large and small groups again a chi squared test was conducted. In preschools, both during large and small group activities, closed questions were asked more often than open-ended questions, but closed questions were significantly more often shown in small group activities, whereas open-ended questions were significantly more often observed during large group than small group activities (see Table 5.6, χ^2 (1, N = 1399) = 4.26, p = .04, Cramer's V = .06). In kindergartens the pattern was opposite; during large group activities the majority of questions

Table 5.6 Overview of closed and open-ended questions and distancing level of statements and questions in targeted and mixed classrooms

		Type of question				Distancing level			
		Closed		Open		Low		High	
		%	*N*	%	*N*	%	*N*	%	*N*
Preschool	Overall	52.7	737	47.3	662	50.3	1283	49.7	1267
	Targeted	54.0	338	46.0	288	56.6	703[a]	43.4	540[a]
	Mixed	51.6	399	48.4	374	44.4	580[b]	55.6	727[b]
Kindergarten	Overall	52.2	617	47.8	564	38.9	916	61.1	1440
	Targeted	51.1	257	48.9	246	40.3	408	59.7	605
	Mixed	53.1	360	46.9	318	37.8	508	62.2	835
Preschool	Large group	50.1	378[a]	49.9	376[a]	44.1	596[a]	55.9	756[a]
	Small group	55.7	359[b]	44.3	286[b]	57.3	687[b]	42.7	511[b]
Kindergarten	Large group	57.4	311[a]	42.6	231[a]	31.0	329[a]	69.0	731[a]
	Small group	47.9	306[b]	52.1	333[b]	45.3	587[b]	54.7	709[b]

Note. Subscript letters denote a significant difference at the .05 level between context (targeted vs. mixed) or group size (large group vs. small group).

concerned closed questions, whereas during small group activities the majority of questions concerned open-ended questions (χ^2 (1, N = 1181) = 10.59, p = .001, Cramer's V = .10). In both preschool and kindergarten the association was weak.

With regard to cognitive demands we found that in preschools, statements and questions of the highest distancing level were generally posed more often during large group activities, whereas during small group activities statements and questions of a low distancing level were more apparent (see Table 5.6, χ^2 (1, N = 2550) = 44.69, p = .000, Cramer's V = .13). In kindergartens, the highest distancing level was more apparent during both large group and small group activities. Nevertheless, statements and questions of a high distancing level were proportionally more prevalent during large group activities as compared to small group activities (χ^2 (1, N = 2356) = 49.87, p = .000, Cramer's V = .15). Again, in both preschool and kindergarten the association was weak.

With regard to teacher dominance we found that, generally, teachers were more dominant than children. In large groups, teachers accounted for 71.9% of all utterances (see Table 5.7, preschool: 73.6%, kindergarten: 70.2%). Teachers were generally slightly less dominant during small group situations, where about four children took part in the activity, than during large group activities, where more children were involved (overall: 65.3% vs. 71.9%; preschool: 70.0% vs. 73.6%; kindergarten: 61.7% vs. 70.2%). We also determined how many different children were verbally involved during large group and small group

Table 5.7 An overview of the agent of utterances in large and small groups

	Overall				Preschool				Kindergarten			
	Teacher		Children		Teacher		Children		Teacher		Children	
	N	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
Large group	4537	71.9	1774	28.1	2302	73.6	825	26.4	2235	70.2	949	29.8
Small group	4382	65.3	2329	34.7	2040	70.0	876	30.0	2342	61.7	1453	38.3

Table 5.8 Percentage of different children making a contribution to the interaction

		Overall				Preschool				Kindergarten			
		N	Mean %	(*SD*)	Range	*N*	Mean %	(*SD*)	Range	*N*	Mean %	(*SD*)	Range
At least one utterance													
Social context	Large group	16	78.1	(14.70)	58.3–100	8	71.9	(10.43)	58.3–92.3	8	84.3	(16.36)	61.5–100
	Small group	14	100	(0)		7	100	(0)		7	100	(0)	
More than three utterances													
Social context	Large group	16	44.0	(3.85)	7.7–72.7	8	45.9	(3.14)	30.8–60.0	8	42.1	(7.25)	7.7–72.7
	Small group	14	100	(0)		7	100	(0)		7	100	(0)	

activities. Despite the large number of children involved during large group activities, in preschool about 15 and in kindergarten 10 to 23, the majority of children were verbally involved at some point during the activity (see Table 5.8, preschool: 71.9%; kindergarten: 84.3%). The majority of children's utterances stemmed from only a few children, however. As shown in Table 5.8, generally 44.0% of children in a classroom made more than three utterances during large group activities (preschool: 45.9%; kindergarten: 42.1%). In short, a minority of the children accounted for the majority of the utterances during large group activities. As shown in Table 5.8, during small group activities all participating children made substantial contributions to the activity in terms of utterances.

Effects of teacher input on children's immediate language output

Types of questions

The relationship between different types of questions and the length of children's direct responses was examined. As shown in Table 5.9 kindergarten children's responses were generally longer than preschool children's responses (preschool: $M = 1.81$, $SD = 1.27$; kindergarten: $M = 2.11$, $SD = 1.59$, $d = -0.21$). An ANOVA showed that the type of question was not associated with differences in the length of preschool children's direct responses (F (2, 417) = 2.46, $p = .09$, $\eta^2 = .01$). In kindergarten classrooms, however, the length of children's direct responses was related to the type of question (F (2, 348) = 9.93, $p = .000$, $\eta^2 = .05$). Post hoc analyses showed that children's responses to open-ended questions were significantly longer than responses to closed questions and multiple choice/completion questions (closed: $M = 1.75$, $SD = 1.46$; multiple choice/completion question: $M = 1.53$, $SD = 0.98$; open-ended $M = 2.44$, $SD = 1.69$; Mean difference: 0.69, $p = .000$; respectively Mean difference: 0.91, $p = .002$). So, for kindergarten children, open-ended questions seemed associated with longer child responses. The effect size, however, was small. For preschool children no differences were found.

Table 5.9 Effects of different types of questions on children's length of utterance in response

	Overall			Preschool			Kindergarten		
Sentence length	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Closed	283	1.78	1.39	154	1.81	1.33	129	1.75	1.46
Multiple choice/completion	81	1.48	1.01	49	1.45	1.04	32	1.53	0.98
Open-ended	407	2.15	1.51	217	1.89	1.27	190	2.44	1.69
Total	771	1.95	1.44	420	1.81	1.27	351	2.11	1.59

Cognitive demands

We also tested whether the cognitive demands, or distancing level of teachers' utterances was related to the length of children's direct responses. An ANOVA with teachers' distancing level as factor showed that preschool children's direct responses to cognitively demanding utterances (statements, open-ended and closed questions were taken into account) were significantly longer than responses to low distancing utterances (see Table 5.10, preschool: F (1, 547) = 5.72, η^2 = .01, p = .02). The effect size, however, was small. For kindergarten children no significant differences in the length of children's responses to low and high distancing utterances were found (F (1, 428) = 0.26, η^2 = .00, p = .61).

Finally, we tested whether the distancing level of teacher utterances was related to the distancing level of children's responses. As shown in Table 5.11, a low distancing utterance by a teacher is more likely to be followed by a low distancing child response. This strong significant relationship was found for both cohorts (preschool: χ^2 (1, N = 559) = 385.66, p = .000, Cramer's V = .83; kindergarten: χ^2 (1, N = 465) = 288.65, p = .000, Cramer's V = .79).

Table 5.10 Effect of distancing on children's length of utterance in response

	Overall			Preschool			Kindergarten		
Distancing teacher	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Low	418	2.13	1.56	263	1.92	1.40	155	2.48	1.76
High	561	2.39	1.67	286	2.21	1.47	275	2.57	1.84
Overall	979	2.28	1.63	549	2.07	1.44	430	2.54	1.81

Table 5.11 Distancing level teacher's utterances and children's distancing level of immediately following utterances

		Distancing level children			
		Low		High	
Distancing level teacher		%	*N*	%	*N*
Preschool	Low	88.2	254	5.2	14
	High	11.8	34	94.8	257
Kindergarten	Low	84.6	148	6.6	19
	High	15.4	27	93.4	271

Discussion

With regard to the first aim of this study, differences in teacher language input across contexts in preschool and kindergarten, we found mixed results. Few differences were found between mixed and targeted preschool and kindergarten; the only significant difference was that in mixed preschools teachers' statements and questions were generally more often of a high distancing level than in targeted preschools. The absence of differences between teachers in mixed and targeted classrooms may be related to composition of the groups; the proportion of native Dutch children in mixed classrooms approached – especially in preschools and during small group activities – the proportion in targeted classrooms. In addition, differences in language skills between children in mixed and targeted groups were perhaps not large enough to exert an influence on teacher behaviour. Possibly, differences between children's language skills need to be substantial, as they presumably were in the studies that did find that teachers adjusted their language practices (de Rivera et al., 2005; Girolametto et al., 2000; Girolametto & Weitzman, 2002; Justice et al., 2013), for teachers to deliberately adjust their language practices to the (developmental level of) children in their classroom.

With regard to teacher-managed activities in small and large groups, we found that although preschool teachers asked in both small and large groups more closed than open-ended questions, open-ended questions were proportionally more often asked during large group activities than during small group activities. This pattern was opposite for kindergarten teachers; they asked more closed questions during large group activities and more open-ended questions during small group activities. Presumably, preschool teachers often chose for closed questions, because these questions place less linguistic and cognitive demands on the young children in their classroom, who often hardly speak Dutch. At the other side of that continuum, with very advanced Dutch language speakers, teachers might be hesitant to ask open-ended questions during large groups because this may lead to off topic prolonged conversation with one of the children instead of whole group involvement in the planned activity (see also Wasik et al., 2006).

Kindergarten teachers did ask relatively many questions that are cognitively demanding; in both small and large groups their utterances were in majority of a high distancing level. Unexpectedly, preschool teachers' utterances in large groups were in majority of a high distancing level, whereas their utterances in small groups were in majority of a low distancing level. That many teachers employ a direct teaching mode in small groups, in which they, for example, check whether children are familiar with certain words, might explain this finding (see also Girolametto & Weitzman, 2002). In sum, we found more differences in teacher language practices between large and small group teacher-managed activities than between mixed and targeted classrooms. However, the associations were weak.

With regard to opportunities for children to make a verbal contribution to an activity we found – in line with earlier studies – that the majority of utterances in large groups were made by the teacher (Aukrust, 2008; Durden & Rainer Dangel, 2008; Girolametto et al., 2003; Lawton & Fowell, 1989; Rydland, Grøver, & Lawrence, 2014). Although teachers nevertheless did succeed to verbally involve a majority of children at some point during the activity a minority of the children accounted for more than three utterances during large group activities, which suggests that teachers are often engaged in prolonged conversation with only a small number of children during large group activities. In contrast, in small groups all participating children were able to make significant contributions to the activity in terms of number of utterances. It is important to note however, that as in large groups, teachers also accounted for the majority of utterances in small group activities. In sum, the current study showed that small groups offer more opportunities for active child engagement, nevertheless earlier studies showed that teacher-managed small group activities occur rarely (Chien et al., 2010; de Haan et al., 2014a; de Haan et al., 2014b; Elicker & Mathur, 1997).

With regard to the second research aim – the effect of different types of questions and utterances with varying cognitive demands on children's immediate language output – we also found mixed results. In kindergarten classrooms open-ended questions elicited the longest child responses, whereas in preschool no relationship between different types of questions and the length of child responses was found. In preschool, however, cognitively demanding utterances were associated with longer child responses than low distancing utterances, whereas in kindergartens no such effect was found. As in other studies however, the distancing level of teacher's utterances was strongly related to the distancing level of child responses in both preschool and kindergarten (Tompkins et al., 2013; Zucker et al., 2010).

Limitations and directions for future research

Although our study explored a topic that is not often addressed in previous research, there are some limitations associated with the current study. First, although our analyses were on utterance level and the number of data points was therefore large, the number of classrooms and teachers involved in this study was small and the multilevel character of the data, meaning that utterances were nested in classrooms, was not taken into account.

In the current study the type of activity was coded, but due to the small sample size teacher language practices were not related to activity type. In order to further unravel teacher language practices it is recommended though, to relate these practices to types of activities in future studies, since earlier research showed that different types of activities might elicit different types of teacher language/behavior (e.g., Gest et al., 2006; Girolametto et al., 2000; Lee & Kinzie, 2012).

For future studies it would be very interesting to also take sequences of utterances into account. Especially in preschools we saw some interesting patterns that were not revealed by our quantitative analyses. One recurring pattern for example, is that the teacher starts with an open-ended question (see Excerpt 5.1: 'What do you need to do C.?') and while the child remains silent the teacher narrows the child's options by asking a closed question ('Can you go outside with your nice jacket?'). After the child's answer (she nods no), the teacher returns to the open-ended question she previously asked ('What do you need to do then?'). Lee and Kinzie (2012) and Siraj-Blatchford and Manni (2008) found a similar pattern; if children were unable to answer an open-ended question, teachers often changed the type of question from open-ended to closed. Another typical pattern concerns the repetition of a question after a 'wrong' answer was given by one of the children (de Rivera et al., 2005).

Probably, the conversation as shown in the excerpt, was rather confusing for the child, because she was convinced that she could wear her regular coat when it is raining – which might be quite true for this child in real life. The excerpt also shows that it can be hard for a teacher to achieve goals she set for an activity. The teacher wants to focus on a couple of words whereas these words may be meaningless to some children or do not align with their current interests or way of thinking. In light of the current tendency to use information about individual children's development (observations, test data, etcetera) for deliberate planning of purposeful (educational) activities adjusted to children's knowledge and skills (an approach that is coined 'opbrengstgericht werken' in Dutch) in Dutch preschools and kindergartens, it is important to note that teachers should never let their purposes and goals prevail over the child's interests and ideas; activities and teacher language should be adapted to the child's (verbal) skills and spontaneous behavior. The excerpt forms an illustration of the phenomenon that teacher-child interaction tends to be characterized by short exchanges directed by the teacher, in which the teacher is not planning to have a genuine conversation with the child (Hayes & Matusov, 2005).

Related to question patterns, it seems that teachers often give children little time to answer a question. Children, especially young second language learners, need time to think about a question and formulate an answer. Given the documented positive effects of a prolonged wait time (at least three seconds whereas teachers typically wait less than a second) on children's language output and active participation, future studies should also take this wait time into account (de Rivera et al., 2005; Rowe, 1986; Siraj-Blatchford & Manni, 2008).

Practical implications

This study showed that a minority of the children in a classroom accounts for the majority of children's utterances during large group activities, which indicates that teachers are often

in prolonged conversation with a few children. In an earlier study (de Haan et al., 2014a) we saw nevertheless that children did spent quite a lot of their time in large groups. Especially in preschool classrooms, where teachers work in couples, time spent in large groups should be reduced to benefit optimally from the availability of two teachers. Professional development efforts should not only focus on teaching teachers why working in small groups is such a powerful means in stimulating children's development, but should also show them what they need to do, since working in small groups requires classroom management skills to manage a small group activity while other children work independently (Powell et al., 2008; Wasik et al., 2006; Zucker et al., 2013). In addition, teachers need to know how to stimulate active child participation in small groups in order to fully achieve the potential of activities in small groups in accelerating child development, since a recent study showed that if teachers engaged in small group activities they tended to adhere to their dominant pattern of giving directions and instructions, a pattern that inhibits active child engagement.

In addition to a shift from large group to small group teacher-managed activities, it might be important to create more situations in which children are able to interact one-to-one in meaningful activities with their teacher. Furthermore, interactions with peers might be very powerful in fostering children's language development; these interactions are repeatedly associated with more complex language output of children (Deunk, 2009; Kontos & Keyes, 1999). Therefore, it is important for teachers to guide children's interactions with peers.

To conclude, this study is a next step in unraveling child experiences in early childhood education. The current study showed that there are meaningful differences between teacher-managed small and large group activities, especially in terms of opportunities for active child engagement. With regard to mixed and targeted classrooms, few differences were found in teacher language practices. The study therefore, does not lend support for the hypothesis that peer effects exert their influence through teacher behaviour; teachers do not seem to adapt their language practices to the children in their classroom. Given the importance of responsive teacher-child interactions though, professional development efforts should be targeted at these skills the coming years in order to elevate quality and effectiveness of early childhood education programs in The Netherlands.

Appendix 5.1 Coding scheme of type of utterance

Type of utterance	Description
Attention/expression	Utterances to gain someone's attention (e.g., 'hello?', 'look!', mentioning a name, etcetera) or utterances reflecting expression (e.g., 'wow', 'whoops', 'so what', etcetera).
Statement	Statements, including restating and repetitions of an utterance. Also words or sentences that finish another speaker's utterance are coded in this category.
Request (imperative)	A request in the form of an imperative (e.g., 'come', 'quiet').
Closed question	Requires a 'yes' or 'no' as an answer. Including restating and repetitions of a closed question.
Multiple choice question	Provides a limited number of choices ('was your present a car, plane or something else?'). Including restating and repetitions.
Completion question	Requires the recipient to fill in the answer ('this is a …?'). Including restating and repetitions.
Open-ended question	The answer to this type of question is not 'yes' or 'no', and the answer is not given in the question as in multiple choice questions.
Affirmation without expansion	Affirmation of an utterance without elaboration (e.g., 'yes', 'all right', etcetera).
Negation without expansion	Negation of an utterance without elaboration (e.g., 'no', 'not', etcetera).
Imitation	Imitation of an utterance or any type of question of another speaker without elaboration.
Imitation with expansion	Imitation of an utterance or any type of question of another speaker with elaboration.
Reading	The utterance concerns reading aloud of printed text.
Rest/incomplete/unintelligible	The utterance could not be coded in one of the above categories or was incomplete or unintelligible.

Excerpt 5.1 An example of a preschool teacher's conversation with children during a teacher-managed academic language activity centering around the theme autumn

At the start of the activity the teacher describes the weather. It is raining, storming and very windy; the weather is horrible.	
teacher	what do you need to do C.? *wat moet je doen C.?*
teacher	can you go outside with your nice jacket? *kun je zo naar buiten met je mooie vestje?*
child	nods no *schudt nee*
teacher	what do you need to do then? *wat moet je doen dan?*
other child	no with your coat *nee met je jas*
child	with a coat *met je jas*
teacher	yes, you'll wear your regular coat? *ja moet je gewone jas aan?*
child	nods yes *knikt ja*
teacher	yes? *ja?*
teacher	and what else over your coat? *en wat nog meer dan er overheen?*
teacher	what do you have to do when it rains C.? *wat moet je doen als het regent C.?*
teacher	wear a regular coat? *een gewone jas aan?*
child	nods yes *knikt ja*
child	yes? *ja?*
teacher	and what else do you wear over your coat? *en wat doe je er overheen dan?*
child	uh umbrella *uh paraplu*
teacher	an umbrella goes over your… *een paraplu doe je boven…*
teacher	points to her head *wijst naar haar hoofd*
At this point the teacher asks the child to get something that she can use when it rains. The child picks a rain coat and the teacher asks the child to name the coat. The child answers that it is a coat, whereas one of the other children remarks that it is a pink coat. Subsequently, the teacher asks the group whether it is a regular coat, upon which the child answers that it is a 'a coat, just a coat'. Finally one of the other children shouts that it is a rain coat.	

Chapter 6

Summary and general discussion

Aim of this dissertation

This study is one of the first in The Netherlands to take actual child experiences in the preschool and kindergarten classroom into account in evaluating the effects of early childhood education programs on disadvantaged children's development in emergent academic skills. Earlier studies of education programs that were carried out in the Dutch context, yielded equivocal results and focused on the mere fact of using an education program. Since internationally, many studies have shown that implementation fidelity is an important factor in the success of a program it is surprising that no study in The Netherlands to date took into account what actually happens in the classroom; what activities are done and how they are done. Insight into child experiences may be more important than the mere reported use of a program (see also Chien et al., 2010; Early et al., 2010). In addition to the effect of an education program, measures of child experiences were used to answer one of the remaining questions in the field of early childhood education: Should education programs aiming to promote the school readiness skills of young learners be targeted specifically at children at risk for delays in these basic skills or should these efforts be aimed at all young children (e.g., Barnett, 2011)? In the city of Utrecht, The Netherlands, an experiment was conducted in which preschool education was deliberately aimed at all children, regardless of background.

In short, in the present dissertation measures of actual child experiences in the preschool and kindergarten classroom were taken into account in evaluating the effects of an education program and the classroom composition on participating children's development in emergent academic skills. The measures of child experiences were obtained through two types of observation methods. First, intensive classroom observations were conducted to gain an in-depth insight into children's everyday activities and interactions in preschools and kindergartens. Second, video-observations were conducted, to unravel interaction processes in more detail. In this chapter the results of the four empirical chapters are summarized and discussed.

Summary of main findings

The first main finding is that we did not find an effect of an education program on disadvantaged children's development. Another important finding, and at the same time an explanation for the null effect, is that the education program was not carried out as intended. This lack of implementation fidelity is indicated by the low average level of teacher-managed activities in small groups and the overall low occurrence of teacher-managed activities. Furthermore, the use of an education program was not decisive in the time spent on teacher-

managed activities; teachers working without a program succeeded as well, or even better, in spending a considerable amount of classroom time on these activities. Importantly, these teacher-managed academic activities were related to child development. In short, the use of an education program was not crucial for the time spent on teacher-managed activities, but the time spent on these activities was important for child development.

The second main finding, with regard to classroom composition, is that especially children in mixed kindergarten classrooms showed accelerated development compared to children in targeted classrooms. The positive effects of mixed classrooms seem to be moderated through interaction with native Dutch peers and not through teacher behavior, two of the supposed pathways through which peer effects exert their influence. In the remainder we will summarize and discuss findings in more detail.

Children's experiences in preschool and kindergarten

Activities

The study in Chapter 2 showed that classrooms differed in the amount of time spent on **teacher-managed academic activities**. The time teachers spent on math activities was remarkably low. In preschool classrooms these activities were rarely observed. In addition, informal observations revealed that the rarely occurring teacher-managed math activities hardly required higher order mathematical thinking, but rather involved counting or shape naming. Compared to math activities, teacher-managed language and literacy activities occurred more frequently, but on average little time was spent on these activities. Despite the overall low occurrence of teacher-managed language and literacy activities, there was considerable variation between teachers. The study in Chapter 3 revealed that classrooms where the education program was implemented, did not differ from classrooms without such a program in the amount of teacher-managed activities.

In both preschool and kindergarten much time was spent on **transition activities** such as gathering material to start an activity, tidying up after an activity, and awaiting one's turn when choosing an activity on a plan board. Most of the transition time concerned mere waiting time. Much time was also spent on care **routines** such as eating snacks and drinking. A remarkable predominant activity in terms of time spending in both preschools and kindergartens in our sample was **outdoor play**. A large share of time was spent on **'other' activities**, which included computer work (only in kindergarten), gymnastics and dancing, music and singing activities, other fine motor activities, and activities that did not fit in any of the categories of the coding scheme.

Interaction opportunities and social context

In both preschool and kindergarten, children spent the majority of time in large groups or in unguided small groups with one or more peers. Engagement in teacher-managed activities in small groups was rare; small group activities initiated and guided by the teacher were rarely observed. Although large group activities were frequent in nature, possibilities for active participation in terms of verbal interaction were limited; during large group activities children were mostly non-interactive or listening passively. Activities during which children were in unguided small groups with peers, provided more opportunities for verbal interaction.

In both targeted and mixed preschool classrooms about two third of interactions were with the teacher. In kindergartens, however, teachers were involved in only a minority of children's interactions in both targeted and mixed classrooms; the vast majority of children's interactions in kindergarten classrooms were with peers. Children in mixed classrooms, both preschool and kindergarten, had more interactions with native Dutch children than children in targeted classrooms. This was associated with the classroom composition; there was a high correlation between interaction with native Dutch peers and the percentage of native Dutch peers in the classroom, which suggests that children interacted with native Dutch peers as expected based on the classroom composition.

Teachers' language input

The study in Chapter 5 showed that the majority of teacher utterances concerned statements, closed and open-ended questions. Generally, kindergarten teachers' utterances were proportionally more often cognitively demanding than preschool teachers' utterances; kindergarten teachers' statements, closed and open-ended questions were in majority cognitively demanding, whereas the pattern was less clear for preschool teachers. Preschool teachers' closed questions were more often cognitively demanding, whereas their open-ended questions were more often not cognitively demanding.

Large versus small group teacher-managed activities

We made a comparison between two types of social contexts: teacher-managed activities in small and large groups. Note that teachers were asked to plan and execute a small group activity when the video-observations were scheduled, and that our classroom observations revealed that naturally occurring teacher-managed small group activities were scarce. In preschools, during both small and large group activities, the majority of questions concerned closed questions. Closed questions were more often observed during small group activities, whereas open-ended questions were significantly more often observed during large group than small group activities. In kindergartens the pattern was opposite; during large group activities the majority of questions were closed, whereas during small group activities the

majority of questions concerned open-ended questions. The associations between social context and type of question, however, were weak in both preschool and kindergarten.

The associations between social context and cognitive demands of utterances were also mixed and weak. In preschools, statements and questions of the highest cognitive distancing level were generally posed more often during large group activities, whereas during small group activities statements and questions of a low distancing level were observed more frequently. In kindergartens, the highest distancing level was observed more often during both large group and small group activities.

During large group activities, teachers accounted for the majority of utterances. During small group activities teachers were slightly less dominant than during large group activities, but still accounted for the majority of utterances. Although during large group activities the majority of children were verbally involved at some point during the activity, the majority of children's utterances stemmed from a minority of the children in the classroom. In preschool and kindergarten about half of the children made more than three utterances during large group activities. In short, a minority of the children accounted for the majority of the utterances during large group activities. During small group activities, on the other hand, all participating children made substantial contributions to the activity in terms of utterances.

Effects of teacher language input on children's immediate language output

Types of questions

The study reported in Chapter 5 showed that the type of question did not elicit differences in the length of preschool children's direct responses. In contrast, although the effect size was small, the length of kindergarten children's direct responses was related to the type of question; children's responses to open-ended questions were longer than responses to closed questions and multiple choice/completion questions.

Cognitive demands

Preschool children's direct responses to cognitively demanding utterances were longer than to low distancing utterances. Note that the effect size was small. For kindergarten children no significant differences in the length of children's responses to cognitively demanding and low distancing utterances were found. For both preschool and kindergarten children, however, the level of the teacher's utterance was likely to be followed by a corresponding child response; low distancing teacher utterances were most likely to be followed by a low distancing child response.

Effects of education program, academic activities and mixed classrooms

Education program

In the study reported in Chapter 3 Accelerated Latent Growth Modeling (LGM; Duncan et al., 2006), with AMOS (version 17; Arbuckle, 2006), was used to combine the preschool and kindergarten cohort in single models of children's emergent literacy and math development, spanning the age range from 3 to 6 years. The analyses focused specifically on the effects of three predictor variables – *the use of an education program*, *the amount of teacher-managed academic activities*, and *the classroom composition with respect to disadvantaged children (targeted vs. mixed classrooms)* – on children's development in emergent academic skills. The results of the study reported in Chapter 3 indicated that the use of an education program to promote disadvantaged children's emergent literacy and math was not effective in accelerating participating children's development.

Teacher-managed academic activities

The effects of teacher-managed academic activities, typical of education programs, on children's development were examined in two studies. First, in the study reported in Chapter 2 all children, regardless of background, were included and effects were estimated over a one year period using repeated measures ANOVAs. In the study reported in Chapter 3 only disadvantaged children, the target population of education programs, were included in Accelerated Latent Growth Models (LGM; Duncan et al., 2006). The unique effects of the amount of teacher-managed academic activities on disadvantaged children's development over a two year period were examined over and above two other predictor variables, the use of an education program and the classroom composition with respect to disadvantaged children (targeted vs. mixed classrooms).

In the study in Chapter 2 we found that children, regardless of background, showed larger gains in emergent academic skills over a one year period if their teacher engaged in relatively many teacher-managed activities. Particularly preschool children showed larger gains in emergent academic skills if their teacher devoted a relatively large proportion of classroom time to language and literacy activities. For kindergarten children, only gains in math skills were associated with teacher's engagement in academic activities; these children showed larger gains in math skills if their teacher initiated relatively many language and literacy activities.

Whereas in the study in Chapter 2 all children in the sample, regardless of background, were included in the analyses, in the study reported in Chapter 3 only disadvantaged children were included in the analyses (this concerned about two third of the original sample). Because the study in Chapter 3 focused specifically on the effect of the education program,

only children at whom these programs are targeted were included in the analyses. In the study reported in Chapter 3 Accelerated Latent Growth Modeling (LGM; Duncan et al., 2006) showed that the extent to which teachers engaged in teacher-managed activities, the kind of activities that were intended by the education program, was significantly related to disadvantaged children's gains in emergent academic skills over a two year period. In the preschool cohort a positive effect was found for emergent math skills, whereas in the kindergarten cohort positive effects were found for both emergent literacy and math skills. Again, the effects were small to medium sized. Note that although on average teacher-managed academic activities occurred infrequently, there was considerable variation between teachers and the occurrence of these activities seemed independent of the use of an education program.

Child-managed academic activities

In the study reported in Chapter 2 we also examined the effect of child-managed language and literacy and math activities on preschool and kindergarten children's development over a one year period. The results showed that preschool and kindergarten children's emergent literacy development benefited from child-managed language and literacy activities. In contrast, engagement in child-managed math activities was not related to preschool and kindergarten children's gains in math skills.

Mixed classrooms

In the study reported in Chapter 3 the unique effects of mixed classrooms on disadvantaged children's development over a two year period were examined over and above two other predictor variables, the use of an education program and the amount of teacher-managed academic activities. The results showed that disadvantaged children in mixed preschool and kindergarten classrooms showed accelerated development in emergent academic skills compared to disadvantaged children in targeted classrooms. In the preschool cohort a positive effect was found for math skills, whereas in the kindergarten cohort positive effects were found for both literacy and math skills. The effects were small to medium sized.

In the study reported in Chapter 4 we further disentangled the effects of mixed classrooms by examining whether the proportion of interaction with native Dutch children was related to children's development in emergent literacy and math skills. This study further explored direct peer effects, the first pathway through which peer effects may exert their influence. We presupposed that classroom composition, as a distal characteristic, affected the likelihood that peer learning occurs indirectly, whereas actual peer interaction reflects proximal processes in the classroom and can be regarded a direct measure of possible peer learning effects (see also Hattie, 2002; Wilkinson, 2002). Therefore, with regard to the

predictor variables the analyses focused specifically on the effect of classroom composition on the intercept and on the effect of peer interaction on the slopes of the growth models of emergent academic skills.

The results of the study in Chapter 4 correspond to the effects of being in a mixed classroom as found in the study reported in Chapter 3. In the preschool cohort one borderline significant effect was found; the percentage of native Dutch children in the classroom was positively associated with children's initial emergent literacy skills. Several effects were found for the kindergarten cohort; the percentage of native Dutch children in the classroom was positively associated with children's initial emergent literacy and math skills. Furthermore and most important, the percentage of actual interaction with native Dutch peers in the kindergarten classroom was positively associated with children's development in emergent literacy skills, which indicates that children who had more interactions with native Dutch peers showed larger gains in emergent literacy skills over time. The effects were medium to large sized.

In the study reported in Chapter 5 the second pathway through which peer effects may exert their influence, teacher behavior, was examined. No differences in teacher language practices were found between mixed and targeted classrooms; the current study, therefore, does not lend support for the hypothesis that peer effects exert their influence through teacher behaviour.

Discussion of findings

The lack of an effect of the implementation of the education program

As in previous Dutch studies (see, for example, Karssen et al., 2013; Veen et al., 2002; Veen et al., 2006; Veen et al., 2008), we did not find an effect of the implementation of an education program. One explanation is that the program was not carried out as intended as indicated by the low average level of teacher-managed activities in small groups and the overall low occurrence of teacher-managed activities. The lack of teacher-managed small group activities is problematic since these activities were considered as the most important means of stimulating children's development in the program Taalrijk. The finding that teachers did not succeed in providing sufficient activities in small groups raises the question whether training and coaching was intensive enough. Working in small groups requires excellent classroom management skills, which may need more attention than a manual, a few days of training and occasional feedback, as was the case with the education program studied here. In addition, the amount of time spent on teacher-managed academic activities did not

differ between classrooms where an education program was implemented and classrooms working without such a program; teachers working without a special education program were equally able to spent time on teacher-managed academic activities (see also Veen et al., 2006). In short, an education program is not decisive for children's experiences of teacher-managed activities. However, engagement in these teacher-managed activities is important for children's development. This dissertation lends further support for the claim that it is not about the mere fact of using an education program, rather it is about how effectively the teacher uses the program in the context of children's needs and the overall climate in the classroom (Phillips et al., 2009).

This dissertation is one of the first studies of early childhood education in The Netherlands which took actual child experiences in the classroom into account in evaluating the effects of an education program. Although our classroom observations yielded insights in children's experiences in the classroom, they do not paint the full picture; to fully understand classroom experiences it is important to also take a qualitative approach since teachers vary considerably in their behavior. For example, with regard to the finding that teacher-managed activities were associated with children's development, it is important to note, however, that relatively much time spent on teacher-managed academic activities does not necessarily equals a high quality educational climate (Phillips et al., 2009). The reverse may be true; the amount of teacher-managed activities might be related to teachers' focus on emergent academic skills, which in turn results in didactic teaching or teacher-directed drill and practice of school skills. Therefore, it is very important to take an in-depth perspective on how activities are employed in future studies.

The fact that studies of early childhood education programs in other countries generally yield more favorable results than studies in The Netherlands, might be associated with the less frequent attendance of children in Dutch provisions, especially in preschool classrooms where children generally attend about ten hours per week. In addition, although programs in The Netherlands generally include a parent component, efforts aiming to include parents may be less intensive in The Netherlands compared to other countries. Since children spend the majority of their time at home (and not in preschool or kindergarten), the home environment is a very powerful means of influencing children's development, and presumably therefore, a combination of center- and home-based programs has proven to be most effective in stimulating children's development (Blok et al., 2005). For that reason, it is highly recommended to include parents in programs.

Generally, evaluation studies focus on measurable outcomes such as assessments of children's skills. However, maybe not all effects of participation in early childhood education are easily translated into measurable and tangible outcomes. Anecdotal evidence suggests that kindergarten teachers and also parents whose older children did not attend preschool report

that preschool attendees do show better school readiness skills than peers or older siblings who did not attend preschool. In preschool, children, for example, become familiarized with school routines and develop their (passive) vocabulary, although some children do not speak a single word during their stay in preschool. In addition, especially in neighborhoods with a high concentration of families that face the accumulated effects of risk factors such as poverty, unemployment and insufficient parenting skills, preschool has an important role in the early detection of developmental delays. This early detection often leads to a referral to specialized care. If these children would not attend preschool, referral is likely to be postponed for at least a year and half when children start in kindergarten.

Positive effects of mixed classrooms

Kindergarten children seem to benefit more from attendance of mixed classrooms than preschool children. Overall, the results of the current study for kindergarten children are in line with previous research showing that disadvantaged children benefited from a more balanced socioeconomic and ethnic-cultural classroom composition, probably through classroom interactions with peers with better expressive abilities and greater vocabularies (Henry & Rickman, 2007; Mashburn et al., 2009; Schechter & Bye, 2007). One explanation for the more favorable results for kindergarten children is that preschool children interact less with peers than kindergarten children. In kindergarten classrooms more language input comes from peers, which increases the likelihood that children benefit from higher skilled peers. Another explanation for the different effects for preschool and kindergarten children may concern the larger share of disadvantaged children in mixed preschool classrooms compared to mixed kindergarten classrooms. Differently put, mixed kindergarten classrooms in this study had a socioeconomically more balanced classroom composition than mixed preschool classrooms which were still attended by a relative small share of native Dutch children. In addition, the fact that preschool children spent about half the time in preschool compared to children in kindergarten, might be associated with the lack of effects. Moreover, disadvantaged children attended preschool four half days, whereas their non-disadvantaged peers attended preschool two half days. For disadvantaged children this meant that they encountered their disadvantaged peers more often and this may have led to preferences for these children. Such preferences may reduce the potential positive effects of mixed classrooms.

The current study expands the growing body of studies of mixed preschool and kindergarten classrooms by taking actual peer interaction into account in examining peer effects. In addition, the second pathway through which peer effects may exert their influence, teacher behavior, was examined. Since no differences in teacher language practices were

found between mixed and targeted classrooms, the current study suggests that the effects of mixed classrooms are moderated through peer interaction and not teacher behaviour.

Despite the fact that targeted provisions seem most cost-effective, the results of the current study favor a mixed, or universal, approach over a targeted approach, especially for kindergarten children. If the preschool system was truly universal in nature, meaning that all children are eligible for full and equal attendance, this might lead to positive effects of mixing for preschool children as well. An important means in doing so may concern an integration of the scattered landscape of provisions for preschool children in one system. Moreover, such an integration of center-based day care, which is more focused on care, and preschool, which is more focused on education, may be beneficial for all participating children, regardless of background (Zigler, Gilliam, & Jones, 2006a).

Limitations and directions for future research

Although the current study expands the research literature by including children's actual experiences in the classroom, there are some limitations. First, although the small scale of the current study allowed for an in-depth examination of child experiences in early childhood classrooms, it is recommended that future studies include a larger sample size – from different geographical areas and ethnic communities – in order to examine the generalizability of the findings. In addition, a large and diverse sample enables an examination of effects for subgroups of children with respect to, for example, cultural background or home environment. It is important to take children's home environment into account in greater detail than in the current study, since parenting practices and the educational climate at home are important predictors of children's skills (Bradley et al., 2001). In the current study broad indicators, such as parents' educational attainment level and cultural background, were considered as a proxy of children's home environment. Second, although a study (Justice et al., 2011) found that the development of non-disadvantaged or high-skilled peers was relatively unaffected by their less-skilled peers' language skills, effects for these children should also be explored further. Third, future studies should also include measures of the actual skill level of peers, for example by using standardized assessment of language skills or by observing naturally unfolding peer interaction in the classroom to test the hypothesis that disadvantaged children profit from interaction with higher-skilled peers in more detail. Perhaps high-skilled children prefer each other as playmates, thereby restricting opportunities for lower skilled peers to interact with high-skilled children. These processes should be taken into account as well, because they might reduce potential positive effects of mixed classrooms. More precise measures enable researchers to examine whether

the language learning environment in targeted classrooms, consisting mainly of children at risk for developmental delays, is actually less optimal than in mixed classrooms.

Practical implications

Practical implications for teachers

Given the limited opportunities for children to actively engage in verbal interaction during large group activities, it is recommended to limit large group activities to a minimum (for instance to a welcome and greeting of children) and engage in small group activities instead. Generally, for preschool teachers in The Netherlands this is easily feasible since they often work in couples. With adequate classroom management skills, however, this is also feasible in kindergarten; while children are playing independently in the classroom, teachers can engage in activities with a small group of children. Teacher-managed small group activities have two important advantages compared to large group activities. First, small group activities provide more opportunities for both verbal and non-verbal engagement of children. Second, because there are fewer children involved, it is easier for a teacher to fine-tune activities to participating children's level of development and needs. Presumably, due to this fine-tuning activities in small groups are – if well designed – for none of the participating children too hard or too easy. This is an important difference with large group activities in which it is nearly impossible to adapt an activity to all participating children's level of development.

It is recommended to closely monitor children's development and use this information for designing activities and composing groups for purposeful small group activities. Although it is important to plan these teacher-managed activities in classroom routines, spontaneously occurring learning activities may be even more important, since children learn most when they are interested and motivated. To save time for meaningful (teacher-managed) activities it is recommended to spend classroom time as effectively as possible. Our results showed that much classroom time was spent on waiting, tidying up or doing nothing at all. In addition, also routine activities such as lunch time or waiting in line before going outside can be used effectively by playing a game, for example by letting children arrange themselves from small to tall.

It is important that the classroom is a rich learning environment, also with respect to teacher language input. Language input should not be simplified; children are only able to learn new words and expressions when they are exposed to them (Henrichs, 2010). Therefore, teachers should ask open-ended and challenging questions whenever possible and bear in

mind that children need time to formulate an answer. Furthermore, it is important to guide children's interactions with peers, especially when there are large differences in language skills.

General practical implications: coaching and exchange of best practices

In order to optimize care and education for young children our general recommendations concern ongoing coaching and an exchange of best practices between teachers. Several studies, including the present dissertation, have shown that the mere use of a program is no guarantee that teachers carry out the program as intended (Connor et al., 2006; Early et al., 2005; Klibanoff et al., 2006; Meyer et al., 1993). Moreover, hypothetically, a teacher who follows a manual exactly as prescribed, might be less effective since she does not adjust activities to children's needs (Phillips et al., 2009). It is important to realize that a program is a means of stimulating children's development, not a goal in itself. Ongoing professionalization, for which coaching may be helpful, is highly needed because the care and education of young children is highly complex and dynamic, and few professions faced so many changes in a relative short period. The key to ongoing professionalization is reflection on one's and colleagues' behavior, monitoring of the schedule and results, and a willingness to experiment with new approaches. In short, it is important that there is a learning climate in (pre)school and that teachers develop into reflective practitioners (see, for example, Bowman, Donovan, & Burns, 2001, p. 151; Leseman & Slot, 2011). Since most of teachers' learning takes place in the workplace, coaching on the job seems most valuable in enhancing quality. Several studies have shown positive effects of coaching on the job on teacher practices (Dickinson & Caswell, 2007; Landry, Swank, Smith, Assel, & Gunnewig, 2006; Neuman & Cunningham, 2009). Coaching should be employed by professionals in child development, which are also experts in adult work place learning. Ideally, these coaches equip teachers with knowledge and tools to make decisions independently while relying on their professional experience and knowledge. After an initial training for the education program, especially preschool teachers have been overwhelmed over the years with training initiatives aimed at other or more specific domains. Early childhood education programs are sometimes for example complemented with programs specifically aiming at improving children's vocabulary (such as 'Met Woorden in de Weer'). Trainers of different programs sometimes give conflicting advice and ideally, teachers are equipped with knowledge and skills that enable them to confidently rate advice at its true value and decide what is best for the children in their classroom. In addition, the use of additional programs may have shifted attention from the main comprehensive program away, with potentially harmful consequences for the implementation of the main comprehensive program.

General conclusion

There are two main conclusions. First, we did not find an effect of an education program on disadvantaged children's development. An explanation for the null effect is that the education program was not carried out as intended, indicated by the low average level of teacher-managed activities in small groups and the overall low occurrence of teacher-managed activities. Although no effect was found of the implementation of an education program, there are indications that teacher-managed activities, typical for such programs, have potential in accelerating participating children's development. Second, the present dissertation provides evidence that mixed classrooms have beneficial effects for young children. The effects of mixed classrooms seem to be moderated through interaction with native Dutch peers and not through teacher behavior, two of the supposed pathways through which peer effects exert their influence. This dissertation expands research in the field of early childhood education by showing the importance of including actual child experiences in the classroom in examining the effect of an education program and classroom composition on children's development. Ongoing coaching, the exchange of best practices between teachers and organizations, in short embedding teachers in learning communities, are considered as critical in ensuring high quality education and care.

References

Aikens, N., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology, 100*, 235-251.

Aikens, N., Kopack Klein, A., Tarullo, L., & West, J. (2013). *Getting ready for kindergarten: Children's progress during Head Start. FACES 2009 report. OPRE report 2013-21a.* Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

Aikens, N., Tarullo, L., Hulsey, L., Ross, C., West, J., & Xue, Y. (2010). *ACF-OPRE report: A year in Head Start: Children, families and programs.* Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

Anderson, L. M., Shinn, C., Fullilove, M. T., Scrimshaw, S. C., Fielding, J. E., Normand, J., & Carande-Kulis, V. G. (2003). The effectiveness of early childhood development programs: A systematic review. *American Journal of Preventive Medicine, 24*(3), 32-46.

Arbuckle, J. L. (2006). *Amos 7.0 user's guide (version 7.0).* Spring House, PA: Amos Development Corporation.

Arbuckle, J. L. (2013). *IBM SPSS Amos 22 user's guide.* Amos Development Corporation.

Aukrust, V. G. (2008). Turkish-speaking first graders in Norway acquiring second language vocabulary, listening comprehension and literacy skills. *Scandinavian Journal of Educational Research, 52*, 293-314.

Barnett, W. S. (1993). Benefit-cost analysis of preschool education: Findings from a 25-year follow-up. *American Journal of Orthopsychiatry, 63*, 500-508.

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children, 5*(3), 25-50.

Barnett, W. S. (1998). Long-term cognitive and academic effects of early childhood education on children in poverty. *Preventive Medicine, 27*, 204-207.

Barnett, W. S. (2011). Four reasons the United States should offer every child a preschool education (chapter 5). In E. Zigler, W. S. Gilliam & W. S. Barnett (Eds.), *The pre-K debates. current controversies and issues* (pp. 34-39). Baltimore, MA: Paul H. Brookes Publishing Co., Inc.

Bell, E. R., Greenfield, D. B., & Bulotsky-Shearer, R. J. (2013). Classroom age composition and rates of change in school readiness for children enrolled in Head Start. *Early Childhood Research Quarterly, 28*, 1-10.

Biedinger, N., Becker, B., & Rohling, I. (2008). Early ethnic educational inequality: The influence of duration of preschool attendance and social composition. *European Sociological Review, 24*, 243-256.

Blank, M., Rose, S. A., & Berlin, L. J. (1978). *The language of learning: The preschool years*. New York: Grune & Stratton.

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*, 35-47.

Bodrova, E. (2008). Make-believe play versus academic skills: A Vygotskian approach to today's dilemma of early childhood education. *European Early Childhood Education Research Journal, 16*, 357-369.

Bowman, B., Donovan, M. S., & Burns, S. (Eds.) (2000). *Eager to learn: Educating our preschoolers*. Washington, DC: National Research Council.

Bradley, R. H., Corwyn, R. F., McAdoo, H. P., & Coll, C. G. (2001). The home environments of children in the United States part I: Variations by age, ethnicity, and poverty status. *Child Development, 72*, 1844-1867.

Bruggers, I., Driessen, G., & Gesthuizen, M. (2014). Voor- en vroegschoolse voorzieningen, effectief of niet? De samenhang tussen deelname aan voor- en vroegschoolse voorzieningen en de taal- en rekenprestaties van leerlingen op de korte en langere termijn [Does early childhood education have effect? The association between participation in early childhood education programs and children's short- and long-term language and math skills]. *Mens en Maatschappij, 89*, 117-150.

Bryant, D. M., Burchinal, M., Lau, L. B., & Sparling, J. J. (1994). Family and classroom correlates of Head Start children's developmental outcomes. *Early Childhood Research Quarterly, 9*, 289-309.

Buitenhek Management & Consult (2014). Peuterspeelzaalwerk NL: Facts & figures 2014. Feitenonderzoek ter voorbereiding op "Een betere basis voor peuters". Retrieved February 2, 2015, from http://www.rijksoverheid.nl/bestanden/documenten-en-publicaties/rapporten/2014/06/17/rapport-peuterspeelzaalwerk-nl-facts-figures-2014/rapport-peuterspeelzaalwerk-nl-facts-figures-2014.pdf on 27-02-2015.

Burchinal, M. R., & Cryer, D. (2003). Diversity, child care quality, and developmental outcomes. *Early Childhood Research Quarterly, 18*, 401-426.

Burchinal, M. R., Peisner-Feinberg, E., Bryant, D. M., & Clifford, R. (2000). Children's social and cognitive development and child-care quality: Testing for differential associations related to poverty, gender, or ethnicity. *Applied Developmental Science, 4*, 149-165.

Burchinal, M. R., Peisner-Feinberg, E., Pianta, R., & Howes, C. (2002). Development of academic skills from preschool through second grade: Family and classroom predictors of developmental trajectories. *Journal of School Psychology, 40*, 415-436.

Burchinal, M. R., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly, 25*, 166-176.

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

Bus, A. G., Leseman, P. P. M., & Neuman, S. B. (2012). Methods for preventing early academic difficulties. In K. R. Harris, S. Graham, T. Urdan, A. G. Buss, S. Mayor, & H. L. Swanson (Eds.), *APA educational psychology handbook, Vol 3: Application to teaching and learning* (pp. 527-550). Washington, DC: American Psychological Association.

Cabell, S. Q., DeCoster, J., LoCasale-Crouch, J., Hamre, B. K., & Pianta, R. C. (2013). Variation in the effectiveness of instructional interactions across preschool classroom settings and learning activities. *Early Childhood Research Quarterly, 28*, 820-830.

Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *The Teachers College Record, 112*, 579-620.

Campbell, F. A., Pungello, E. P., Miller-Johnson, S., Burchinal, M., & Ramey, C. T. (2001). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology, 37*, 231-242.

Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development, 65*, 684-698.

Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian project. *Applied Developmental Science, 6*, 42-57.

Centraal Bureau voor de Statistiek. (2012). *Jaarrapport integratie 2012 [Annual report on integration 2012]*. Den Haag/Heerlen, The Netherlands: Centraal Bureau voor de Statistiek.

Centraal Bureau voor de Statistiek, Visser, H., & Huynen, B. (2014). Minder kinderen naar kinderdagverblijven. Retrieved February 2, 2015, from http://www.cbs.nl/nl-NL/menu/themas/arbeid-sociale-zekerheid/publicaties/artikelen/archief/2014/2014-4024-wm.htm.

Chien, N. C., Howes, C., Burchinal, M., Pianta, R. C., Ritchie, S., Bryant, D. M., . . . Barbarin, O. A. (2010). Children's classroom engagement and school readiness gains in prekindergarten. *Child Development, 81*, 1534-1549.

Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology, 98*, 665-689.

Danis, A., Bernard, J., & Leproux, C. (2000). Shared picture-book reading: A sequential analysis of adult-child verbal interactions. *British Journal of Developmental Psychology, 18*, 369-388.

de Goede, D., & Reezigt, G. J. (2001). *Implementatie en effecten van de voorschool in Amsterdam [Implementation and effects of preschool in Amsterdam]*. Groningen, The Netherlands: GION.

de Haan, A. K. E., Elbers, E., Hoofs, H., & Leseman, P. P. M. (2013). Targeted versus mixed preschools and kindergartens: Effects of class composition and teacher-managed activities on disadvantaged children's emergent academic skills. *School Effectiveness and School Improvement, 24*, 177-194.

de Haan, A. K. E., Elbers, E., & Leseman, P. P. M. (2014a). Classroom composition and peer verbal interaction in targeted versus mixed preschools and kindergartens as related to children's development in emergent academic skills. *Manuscript in preparation.*

de Haan, A. K. E., Elbers, E., & Leseman, P. P. M. (2014b). Teacher- and child-managed academic activities in preschool and kindergarten and their influence on children's gains in emergent academic skills. *Journal of Research in Childhood Education, 28*, 43-58.

de Haan, A. K. E., Elbers, E., & Leseman, P. P. M. (2014c). Teacher language input across contexts in preschool and kindergarten: Effects on children's immediate language output. *Manuscript in preparation.*

de Rivera, C., Girolametto, L., Greenberg, J., & Weitzman, E. (2005). Children's responses to educators' questions in day care play groups. *American Journal of Speech-Language Pathology, 14*, 14-26.

Deunk, M. I. (2009). Use and development of speech acts in different contexts in preschool classrooms. *Discourse practices in preschool: Young children's participation in everyday classroom activities* (pp. 101-131). Doctoral dissertation. University of Groningen, The Netherlands.

Dickinson, D. K., & Caswell, L. (2007). Building support for the language and early literacy in preschool classrooms through in-service professional development: Effects of the literacy environment enrichment program (LEEP). *Early Childhood Research Quarterly, 22*, 243-260.

Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E. S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology, 95*, 465-481.

Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school.* Baltimore, MD: Paul H Brookes Publishing.

Doolaard, S., & Leseman, P. P. M. (2008). *Versterking van het fundament* [Strengthening the fundament]. Groningen, The Netherlands: GION.

Dotterer, A. M., Burchinal, M., Bryant, D., Early, D., & Pianta, R. C. (2013). Universal and targeted pre-kindergarten programmes: A comparison of classroom characteristics and child outcomes. *Early Child Development and Care, 183*, 931-950.

Driessen, G. (2000). The limits of educational policy and practice? The case of ethnic minorities in The Netherlands. *Comparative Education, 36*, 55-72.

Driessen, G. (2012). *Variatie in voor- en vroegschoolse educatie. Een onderzoek naar de uiteenlopende wijzen waarop in gemeenten vorm wordt gegeven aan VVE [Variation in early childhood education. A study of differences between municipalities in the implementation of early childhood education]*. Nijmegen, The Netherlands: ITS, Radboud Universiteit Nijmegen.

Driessen, G., & Merry, M. S. (2014). Trends in educational disadvantage in Dutch primary school. *Educational Review, 66*, 276-292.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues and applications* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Durden, T., & Rainer Dangel, J. (2008). Teacher-involved conversations with young children during small group activity. *Early Years, 28*, 251-266.

Early, D. M., Barbarin, O., Bryant, D., Burchinal, M., Chang, F., Clifford, R., . . . Ritchie, S. (2005). *Prekindergarten in eleven states: NCEDL's multi-state study of pre-kindergarten and study of state-wide early education programs (SWEEP).* Chapel Hill, NC: National Center for Early Development & Learning.

Early, D. M., Iruka, I. U., Ritchie, S., Barbarin, O. A., Winn, D. M. C., Crawford, G. M., . . . Pianta, R. C. (2010). How do pre-kindergarteners spend their time? Gender, ethnicity, and income as predictors of experiences in pre-kindergarten classrooms. *Early Childhood Research Quarterly, 25*, 177-193.

Eldering, L., & Vedder, P. (1999). The Dutch experience with the home intervention program for preschool youngsters (HIPPY). In L. Eldering, & P. P. M. Leseman (Eds.), *Effective early education: Cross-cultural perspectives* (pp. 259-285). New York: Falmer.

Elicker, J., & Mathur, S. (1997). What do they do all day? Comprehensive evaluation of a full-day kindergarten. *Early Childhood Research Quarterly, 12*, 459-480.

Fitzpatrick, M. D. (2008). Starting school at four: The effect of universal pre-kindergarten on children's academic achievement. *The B.E. Journal of Economic Analysis & Policy, 8*(1), Article 46.

Gest, S. D., Holland-Coviello, R., Welsh, J. A., Eicher-Catt, D. L., & Gill, S. (2006). Language development subcontexts in Head Start classrooms: Distinctive patterns of teacher talk during free play, mealtime, and book reading. *Early Education and Development, 17*, 293-315.

Ginsburg, H. P., Lee, J. S., & Boyd, J. S. (2008). Mathematics education for young children: What it is and how to promote it. *Social Policy Report, 22*(1), 3-22.

Girolametto, L., & Weitzman, E. (2002). Responsiveness of child care providers in interactions with toddlers and preschoolers. *Language, Speech, and Hearing Services in Schools, 33*, 268-281.

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.

Girolametto, L., Weitzman, E., van Lieshout, R., & Duff, D. (2000). Directiveness in teachers' language input to toddlers and preschoolers in day care. *Journal of Speech, Language, and Hearing Research, 43*, 1101-1114.

Gorey, K. M. (2001). Early childhood education: A meta-analytic affirmation of the short- and long-term benefits of educational opportunity. *School Psychology Quarterly, 16*, 9-30.

Gormley Jr., W. T., Gayer, T., Phillips, D., & Dawson, B. (2005). The effects of universal pre-K on cognitive development. *Developmental Psychology, 41*, 872-884.

Gormley Jr., W. T., & Phillips, D. (2005). The effects of universal pre-k in Oklahoma: Research highlights and policy implications. *Policy Studies Journal, 33*, 65-82.

Graham, T. A., Nash, C., & Paul, K. (1997). Young children's exposure to mathematics: The child care context. *Early Childhood Education Journal, 25*, 31-38.

Graue, E. (2008). Teaching and learning in a post-DAP world. *Early Education and Development, 19*, 441-447.

Guarino, C. M., Hamilton, L. S., Lockwood, J. R., Rathbun, A. H., & Hausken, E. G. (2006). *Teacher qualifications, instructional practices, and reading and mathematics gains of kindergartners*. Washington, D.C.: National Center for Education Statistics.

Guimarães, S., & McSherry, K. (2002). The curriculum experiences of pre-school children in Northern Ireland: Classroom practices in terms of child-initiated play and adult-directed activities. *International Journal of Early Years Education, 10*, 85-94.

Gullo, D. F., & Hughes, K. (2011a). Reclaiming kindergarten: Part I: Questions about theory and practice. *Early Childhood Education Journal, 38*, 323-328.

Gullo, D. F., & Hughes, K. (2011b). Reclaiming kindergarten: Part II: Questions about policy. *Early Childhood Education Journal, 38*, 393-397.

Guo, Y., Tompkins, V., Justice, L., & Petscher, Y. (2014). Classroom age composition and vocabulary development among at-risk preschoolers. *Early Education and Development, 25*, 1016-1034.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Paul H Brookes Publishing.

Hattie, J. A. C. (2002). Classroom composition and peer effects. *International Journal of Educational Research, 37*, 449-481.

Hayes, R., & Matusov, E. (2005). Designing for dialogue in place of teacher talk and student silence. *Culture & Psychology, 11*, 339-357.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*, 1900-1902.

Henrichs, L. (2010). *Academic language in early childhood interactions. A longitudinal study of 3- to 6-year-old Dutch monolingual children*. Doctoral dissertation. University of Amsterdam.

Henry, G. T., Gordon, C. S., & Rickman, D. K. (2006). Early education policy alternatives: Comparing quality and outcomes of Head Start and state prekindergarten. *Educational Evaluation and Policy Analysis, 28*, 77-99.

Henry, G. T., & Rickman, D. K. (2007). Do peers influence children's skill development in preschool? *Economics of Education Review, 26*, 100-112.

Hindman, A. H., Connor, C. M., Jewkes, A. M., & Morrison, F. J. (2008). Untangling the effects of shared book reading: Multiple factors and their associations with preschool literacy outcomes. *Early Childhood Research Quarterly, 23*, 330-350.

Hindman, A. H., Wasik, B. A., & Erhart, A. C. (2012). Shared book reading and Head Start preschoolers' vocabulary learning: The role of book-related discussion and curricular connections. *Early Education and Development, 23*, 451-474.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*, 1368-1378.

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development, 73*, 418-433.

Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Childhood Research Quarterly, 23*, 27-50.

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*, 381-404.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45*, 337-374.

Hyson, M. C. (1991). The characteristics and origins of the academic preschool. *New Directions for Child and Adolescent Development, 1991*(53), 21-29.

Inspectie van het Onderwijs. (2013). *Extra aandacht nodig voor achterstanden bij het jonge kind. Eindrapport bestandsopname voor- en vroegschoolse educatie in Nederland [Developmental delays in young children warrant extra attention. Final report on the inventory of early childhood education in The Netherlands]*. Ministerie van Onderwijs, Cultuur en Wetenschap.

Jepma, I., Rutten, S., & Beekhoven, S. (2009). *Quick scan VVE-programma's [Quick scan early childhood education programs]*. Utrecht, The Netherlands: Sardes.

Jordan, N. C., & Levine, S. C. (2009). Socioeconomic variation, number competence, and mathematics learning difficulties in young children. *Developmental Disabilities Research Reviews, 15*, 60-68.

Justice, L. M., Mashburn, A., Pence, K. L., & Wiggins, A. (2008a). Experimental evaluation of a preschool language curriculum: Influence on children's expressive language skills. *Journal of Speech, Language and Hearing Research, 51*, 983-1001.

Justice, L. M., Mashburn, A. J., Hamre, B. K., & Pianta, R. C. (2008b). Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly, 23*, 51-68.

Justice, L. M., McGinty, A. S., Zucker, T. A., Cabell, S. Q., & Piasta, S. B. (2013). Bi-directional dynamics underlie the complexity of talk in teacher-child play-based conversations in classrooms serving at-risk pupils. *Early Childhood Research Quarterly, 28*, 496-508.

Justice, L. M., Petscher, Y., Schatschneider, C., & Mashburn, A. (2011). Peer effects in preschool classrooms: Is children's language growth associated with their classmates' skills? *Child Development, 82*, 1768-1777.

Kainz, K., & Vernon-Feagans, L. (2007). The ecology of early reading development for children in poverty. *The Elementary School Journal, 107*, 407-427.

Karssen, M., van der Veen, I., Veen, A., van Daalen, M., & Roeleveld, J. (2013). *Effecten van deelname aan en kwaliteit van voor-en vroegschoolse educatie op de ontwikkeling van kinderen [Effects of participation in and quality of early childhood education on children's development].* Amsterdam: Kohnstamm Instituut.

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk". *Developmental Psychology, 42*, 59-69.

Kline, R. B. (2005). *Principles and practices of structural equation modeling* (2nd ed.). New York, NY: The Guilford Press.

Kontos, S., & Keyes, L. (1999). An ecobehavioral analysis of early childhood classrooms. *Early Childhood Research Quarterly, 14*, 35-50.

Korat, O. (2005). Contextual and non-contextual knowledge in emergent literacy development: A comparison between children from low SES and middle SES communities. *Early Childhood Research Quarterly, 20*, 220-238.

La Paro, K. M., & Pianta, R. C. (2000). Predicting children's competence in the early school years: A meta-analytic review. *Review of Educational Research, 70*, 443-484.

Landry, S. H., Swank, P. R., Smith, K. E., Assel, M. A., & Gunnewig, S. B. (2006). Enhancing early literacy skills for preschool children: Bringing a professional development model to scale. *Journal of Learning Disabilities, 39*, 306-324.

Lawton, J. T., & Fowell, N. (1989). A description of teacher and child language in two preschool programs. *Early Childhood Research Quarterly, 4*, 407-432.

Lee, V. E., Loeb, S., & Lubeck, S. (1998). Contextual effects of prekindergarten classrooms for disadvantaged children on cognitive development: The case of chapter 1. *Child Development, 69*, 479-494.

Lee, Y., & Kinzie, M. B. (2012). Teacher question and student response with regard to cognition and language use. *Instructional Science, 40*, 857-874.

Leseman, P. P. M. (2009). The impact of high quality education and care on the development of young children: Review of the literature. *Tackling social and cultural inequalities through early childhood education and care in Europe* (pp. 17-49). Brussels, Belgium: European Commission.

Leseman, P. P. M. & Slot, P. L. (2011). Increasing team competence in examining practice critically. *Children in Europe*, 21, 10-11.

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 1*, 86-92.

Magnuson, K. A., & Waldfogel, J. (2005). Early childhood care and education: Effects on ethnic and racial gaps in school readiness. *The Future of Children, 15*(1), 169-196.

Malone, L., Hulsey, L., Aikens, N., West, J., & Tarullo, L. (2010). *ACF-OPRE report: Data tables for FACES 2006 Head Start children go to kindergarten report.* Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

Marinac, J., Ozanne, A., & Woodyatt, G. (2004). Talking to children: The within-classroom nature of everyday adult language input in the early childhood educational environment. *International Journal of Speech-Language Pathology, 6*, 237-246.

Mashburn, A. J., Justice, L. M., Downer, J. T., & Pianta, R. C. (2009). Peer effects on children's language achievement during pre-kindergarten. *Child Development, 80*, 686-702.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., . . . Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development, 79*, 732-749.

Massey, S. L., Pence, K. L., Justice, L. M., & Bowles, R. P. (2008). Educators' use of cognitively challenging questions in economically disadvantaged preschool classroom contexts. *Early Education and Development, 19*, 340-360.

Mayo, A. Y., & Leseman, P. P. M. (2008). Off to a good start? Vocabulary development and differences in early family and classroom experiences of children from native-Dutch and ethnic minority families in The Netherlands. *Educational and Child Psychology, 25*(1), 70-82.

McKey, R. H., Condelli, L., Ganson, H., Barret, B. J., McConkey, C., & Plantz, N. C. (1985). *The impact of Head Start on children, families, and communities: Executive summary*. Washington, DC: CSR, Inc.

Melhuish, E. C. (2004*). A literature review of the impact of early years provision on young children, with emphasis given to children from disadvantaged backgrounds.* London: Institute for the Study of Children, Families & Social Issues.

Meyer, L. A., Wardrop, J. L., Hastings, C. N., & Linn, R. L. (1993). Effects of ability and settings on kindergartners' reading performance. *The Journal of Educational Research, 86*, 142-160.

Moller, A. C., Forbes-Jones, E., & Hightower, A. D. (2008). Classroom age composition and developmental change in 70 urban preschool classrooms. *Journal of Educational Psychology, 100*, 741-753.

Morrow, L. M., & Smith, J. K. (1990). The effects of group size on interactive storybook reading. *Reading Research Quarterly, 25*, 213-231.

Naigles, L. R., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language, 25*, 95-120.

Nap-Kolhoff, E., Van Schilt-Mol, T., Simons, M., Sontag, L., Van Steensel, R., & Vallen, T. (2008). *VVE onder de loep. Een studie naar de uitvoering en effectiviteit van voor- en vroegschoolse educatieve programma's [Focus on early childhood education. A study of the execution and effects of early childhood education programs]*. Tilburg, The Netherlands: IVA.

Nelson, G., Westhues, A., & MacLeod, J. (2003). A meta-analysis of longitudinal research on preschool prevention programs for children. *Prevention & Treatment, 6*, Article 31.

Neuman, S. B., Copple, C., & Bredekamp, S. (2000). *Learning to read and write: developmentally appropriate practice*. Washington, DC: NAEYC.

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*, 532-566.

Nevitt, J., & Hancock, G. R. (2004). Evaluating small sample approaches for model test statistics in structural equation modeling. *Multivariate Behavioral Research, 39*, 439-478.

OECD. (2006). *Starting strong II: Early childhood education and care*. Paris: Organisation for Economic Cooperation and Development.

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 child-care quality to children's cognitive and social developmental trajectories through second grade. *Child Development, 72*, 1534-1553.

Pellegrino, M. L. M., & Scopesi, A. (1990). Structure and function of baby talk in a day-care centre. *Journal of Child Language, 17*, 101-114.

Pence Turnbull, K., Beckman Anthony, A., Justice, L., & Bowles, R. (2009). Preschoolers' exposure to language stimulation in classrooms serving at-risk children: The contribution of group size and activity context. *Early Education and Development, 20*, 53-79.

Peterson, C., Jesso, B., & McCabe, A. (1999). Encouraging narratives in preschoolers: An intervention study. *Journal of Child Language, 26*, 49-67.

Phillips, D., Gormley Jr., W. T., & Lowenstein, A. (2007, March). *Classroom quality and time allocation in Tulsa's early childhood programs*. Paper presented at the Meeting of the Society for Research in Child Development, Boston, MA.

Phillips, D., Gormley Jr., W. T., & Lowenstein, A. E. (2009). Inside the pre-kindergarten door: Classroom climate and instructional time allocation in Tulsa's pre-K programs. *Early Childhood Research Quarterly, 24*, 213-228.

Phillips, L. B., & Twardosz, S. (2003). Group size and storybook reading: Two-year-old children's verbal and nonverbal participation with books. *Early Education and Development, 14*, 453-478.

Powell, D. R., Burchinal, M. R., File, N., & Kontos, S. (2008). An eco-behavioral analysis of children's engagement in urban public school preschool classrooms. *Early Childhood Research Quarterly, 23*, 108-123.

Ramey, C. T., & Campbell, F. A. (1984). Preventive education for high-risk children: Cognitive consequences of the Carolina Abecedarian project. *American Journal of Mental Deficiency, 88*, 515-523.

Ramey, C. T., & Ramey, S. L. (2004). Early learning and school readiness: Can early intervention make a difference? *Merrill-Palmer Quarterly, 50*, 471-491.

Reid, J. L., & Ready, D. D. (2013). High-quality preschool: The socioeconomic composition of preschool classrooms and children's learning. *Early Education and Development, 24*, 1082-1111.

Reynolds, A. J., & Temple, J. A. (2008). Cost-effective early childhood development programs from preschool to third grade. *Annual Review of Clinical Psychology, 4*, 109-139.

Reynolds, A. J., Temple, J. A., Ou, S. R., Arteaga, I. A., & White, B. A. (2011). School-based early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. *Science, 333*, 360-364.

Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2001). Long-term effects of an early childhood intervention on educational achievement and juvenile arrest: A 15-year follow-up of low-income children in public schools. *JAMA, 285*, 2339-2346.

Rolnick, A., & Grunewald, R. (2011). The economic case for targeted preschool programs (chapter 3). In E. Zigler, W. S. Gilliam & W. S. Barnett (Eds.), *The pre-k debates. current controversies and issues* (pp. 22-26). Baltimore, MA: Paul H. Brookes Publishing Co., Inc.

Rowe, M. B. (1986). Wait time: Slowing down may be a way of speeding up! *Journal of Teacher Education, 37*(1), 43-50.

Royce, J. M., Darlington, R. B., & Murray, H. W. (1983). As the twig is bent: Lasting effects of preschool programs. In Consortium for Longitudinal Studies (Ed.), *Pooled analyses: Findings across studies* (pp. 411-459). Hillsdale, NJ: Erlbaum.

Rydland, V., Grøver, V., & Lawrence, J. (2014). The second-language vocabulary trajectories of Turkish immigrant children in Norway from ages five to ten: The role of preschool talk exposure, maternal education, and co-ethnic concentration in the neighborhood. *Journal of Child Language, 41*, 352-381.

Schaffer, H., & Liddell, C. (1984). Adult-child interaction under dyadic and polyadic conditions. *British Journal of Developmental Psychology, 2*, 33-42.

Schechter, C., & Bye, B. (2007). Preliminary evidence for the impact of mixed-income preschools on low-income children's language growth. *Early Childhood Research Quarterly, 22*, 137-146.

Schmit, S. (November 6, 2014). A new lens: Early childhood data from the office of civil rights. Retrieved from http://www.clasp.org/issues/child-care-and-early-education/in-focus/a-new-lens-early-childhood-data-from-the-office-of-civil-rights

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh, PA: Psychology Software Tools Inc.

van Schooten, E., & Sleegers, P. (2009). *Onderzoek naar de effectiviteit van VVE- en peuterspeelzalen in Oosterhout en Den Bosch [A study of the effects of educational preschool and play groups in Oosterhout and Den Bosch].* Amsterdam: SCO-Kohnstamm Instituut.

Schwartz, H. (2010). Housing policy is school policy: Economically integrative housing promotes academic success in Montgomery county, Maryland. *Education Digest: Essential Readings Condensed for Quick Review, 76*(6), 42-48.

Schweinhart, L. J., & Weikart, D. P. (1988). Education for young children living in poverty: Child-initiated learning or teacher-directed instruction? *The Elementary School Journal, 89*, 212-224.

Schweinhart, L. J., & Weikart, D. P. (1997). The High/Scope preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly, 12*, 117-143.

Shager, H. M. (2012). *What role do peer effects play in early childhood education? Evidence from the 2003 Head Start Family and Child Experiences Survey (FACES).* Unpublished doctoral dissertation. University of Wisconsin-Madison.

Siraj-Blatchford, I., & Manni, L. (2008). 'Would you like to tidy up now?' An analysis of adult questioning in the English foundation stage. *Early Years, 28*, 5-22.

Slot, P. L. (2014). Summary and general discussion. In P. L. Slot, *Early childhood education and care in The Netherlands. Quality, curriculum, and relations with child development*. pp. 131-147.

Slot, P. L., Leseman, P. P. M., Verhagen, J., & Mulder, H. (2014). *Associations between structural quality aspects and process quality in Dutch early childhood education and care settings.* Unpublished manuscript.

Slot, P. L., Mulder, H., Verhagen, J., & Leseman, P. P. M. (2014). *Domain-general and domain-specific quality characteristics of early childhood education and care predict growth of two-year olds' vocabulary and attention skills over one year.* Unpublished manuscript.

Smith, M. W., & Dickinson, D. K. (1994). Describing oral language opportunities and environments in Head Start and other preschool classrooms. *Early Childhood Research Quarterly, 9*, 345-366.

Smith, S. S., & Dixon, R. G. (1995). Literacy concepts of low-and middle-class four-year-olds entering preschool. *The Journal of Educational Research, 88*, 243-253.

Stipek, D., Feiler, R., Byler, P., Ryan, R., Milburn, S., & Salmon, J. M. (1998). Good beginnings: What difference does the program make in preparing young children for school? *Journal of Applied Developmental Psychology, 19*, 41-66.

Stipek, D., Feiler, R., Daniels, D., & Milburn, S. (1995). Effects of different instructional approaches on young children's achievement and motivation. *Child Development, 66*, 209-223.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2004). *The effective provision of pre-school education (EPPE) project: Final report: A longitudinal study funded by the DfES 1997-2004*. London: Institute of Education.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2010). *Early childhood matters: Evidence from the effective pre-school and primary education project*. New York, NY: Routledge.

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*, 109-124.

Sylva, K., Siraj-Blatchford, I., Taggart, B., Sammons, P., Melhuish, E., Elliot, K., & Totsika, V. (2006). Capturing quality in early childhood through environmental rating scales. *Early Childhood Research Quarterly, 21*, 76-92.

Tompkins, V., Zucker, T. A., Justice, L. M., & Binici, S. (2013). Inferential talk during teacher-child interactions in small-group play. *Early Childhood Research Quarterly, 28*, 424-436.

Tonyan, H. A., & Howes, C. (2003). Exploring patterns in time children spend in a variety of child care activities: Associations with environmental quality, ethnicity, and gender. *Early Childhood Research Quarterly, 18*, 121-142.

Tudge, J. R. H., & Doucet, F. (2004). Early mathematical experiences: Observing young black and white children's everyday activities. *Early Childhood Research Quarterly, 19*, 21-39.

Tymms, P. (2001). The development of a computer-adaptive assessment in the early years. *Educational and Child Psychology, 18*(3), 20-30.

U.S. Department of Health and Human Services, Administration for Children and Families. (January 2010). *Head Start impact study. Final report.* Washington, DC.

Van den Heuvel-Panhuizen, M., & van den Boogaard, S. (2008). Picture books as an impetus for kindergartners' mathematical thinking. *Mathematical Thinking and Learning, 10*, 341-373.

Van den Heuvel-Panhuizen, M., van den Boogaard, S., & Doig, B. (2009). Picture books stimulate the learning of mathematics. *Australasian Journal of Early Childhood, 34*(2), 30-39.

Van der Hoeven-van Doornum, A. (2005). *Development on scale, instruction at measure. Obis, a system of value added indicators in primary education*. Nijmegen, The Netherlands: ITS.

Van Horn, M. L., Karlin, E. O., Ramey, S. L., Aldridge, J., & Snyder, S. W. (2005). Effects of developmentally appropriate practices on children's development: A review of research and discussion of methodological and analytical issues. *The Elementary School Journal, 105*, 325-351.

Van Tuijl, C., & Leseman, P. P. M. (2007). Increases in the verbal and fluid cognitive abilities of disadvantaged children attending preschool in The Netherlands. *Early Childhood Research Quarterly, 22*, 188-203.

Van Tuijl, C., Leseman, P. P. M., & Rispens, J. (2001). Efficacy of an intensive home-based educational intervention programme for 4- to 6-year-old ethnic minority children in The Netherlands. *International Journal of Behavioral Development, 25*, 148-159.

Veen, A., Derriks, M., & Roeleveld, J. (2002). *Een jaar later. Vervolgonderzoek evaluatie van Kaleidoscoop en Piramide [A year later. Follow up evaluation study Kaleidoscoop and Piramide]*. Amsterdam: SCO-Kohnstamm Instituut.

Veen, A., Fukkink, R., & Roeleveld, J. (2006). *Evaluatie van Startblokken en Basisontwikkeling. Implementatie en effecten van het programma Startblokken en Basisontwikkeling in het kader van voor- en vroegschoolse educatie [Evaluation of Startblokken and Basisontwikkeling. Implementation and effects of the education program Startblokken en Basisontwikkeling in early childhood education]*. Amsterdam: SCO-Kohnstamm Instituut.

Veen, A., Roeleveld, J., & Leseman, P. (2000). *Evaluatie van Kaleidoscoop en Piramide. Eindrapportage [Evaluation of Kaleidoscoop and Piramide. Final report]*. Amsterdam: SCO-Kohnstamm Instituut.

Veen, A., Roeleveld, J., & van Daalen, M. (2008). *Implementatie en effecten van voor- en vroegschoolse educatie in Rotterdam. Samenvatting [Implementation and effects of early childhood education in Rotterdam. Summary]*. Amsterdam: SCO-Kohnstamm Instituut.

Walsh, D. J. (1989). Changes in kindergarten: Why here? Why now? *Early Childhood Research Quarterly, 4*, 377-391.

Wasik, B. A., Bond, M. A., & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology, 98*, 63-74.

Wilkinson, I. A. G. (2002). Introduction: Peer influences on learning: Where are they? *International Journal of Educational Research, 37*, 395-401.

Winsler, A., & Carlton, M. P. (2003). Observations of children's task activities and social interactions in relation to teacher perceptions in a child-centered preschool: Are we leaving too much to chance? *Early Education and Development, 14*, 155-178.

Winsler, A., Caverly, S. L., Willson-Quayle, A., Carlton, M. P., Howell, C., & Long, G. N. (2002). The social and behavioral ecology of mixed-age and same-age preschool classrooms: A natural experiment. *Journal of Applied Developmental Psychology, 23*, 305-330.

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley Jr., W. T., . . . Zaslow, M. J. (2013, October). *Investing in our future: The evidence base on preschool education*. Society for Research in Child Development and Foundation for Child Development.

Zigler, E., Gilliam, W. S., & Jones, S. M. (2006a). The universal preschool movement. In E. Zigler, W. S. Gilliam, & S. M. Jones (Eds.), *A vision for universal preschool education* (pp. 4). New York: Cambridge University Press.

Zigler, E., Gilliam, W. S., & Jones, S. M. (2006b). The universal preschool movement. In E. Zigler, W. S. Gilliam, & S. M. Jones (Eds.), *A vision for universal preschool education* (pp. 1-18). New York: Cambridge University Press.

Zill, N., Resnick, G., Kim, K., O'Donnell, K., Sorongon, A., McKey, R. H., . . . D'Elio, M. A. (2003). *Head Start FACES 2000: A whole-child perspective on program performance. Fourth progress report.* Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

Zucker, T. A., Justice, L. M., Piasta, S. B., & Kaderavek, J. N. (2010). Preschool teachers' literal and inferential questions and children's responses during whole-class shared reading. *Early Childhood Research Quarterly, 25*, 65-83.

Zucker, T. A., Solari, E. J., Landry, S. H., & Swank, P. R. (2013). Effects of a brief tiered language intervention for prekindergartners at risk. *Early Education and Development, 24*, 366-392.

Samenvatting
(Summary in Dutch)

Introductie

Onderwijsachterstanden en voor- en vroegschoolse educatie (VVE)

Kinderen uit gezinnen met een lage sociaal-economische status (SES) of niet-Nederlandse culturele achtergrond hebben, vergeleken met andere kinderen bij de aanvang van het basisonderwijs regelmatig een achterstand in taal- en beginnende lees- en rekenvaardigheden. Omdat deze vaardigheden belangrijke voorspellers zijn van latere schoolprestaties, zijn in Nederland net als in veel andere geïndustrialiseerde landen, voor- en vroegschoolse educatie programma's (VVE-programma's) geïmplementeerd om deze groep kinderen zo goed mogelijk voor te bereiden op het onderwijs vanaf groep 3. Uit internationaal onderzoek blijkt dat dergelijke programma's op de korte en lange termijn effectief zijn in het stimuleren van de ontwikkeling van deelnemende kinderen. De kwaliteit van het aanbod blijkt een belangrijke factor in het bewerkstelligen van effecten; kinderen die een aanbod van hoge kwaliteit kregen, lieten betere resultaten zien dan kinderen die een aanbod van lagere kwaliteit kregen. Ook in Nederland is onderzoek gedaan naar VVE-programma's; de implementatie van drie programma's, Piramide, Kaleidoscoop en Startblokken/Basisontwikkeling, is middels quasi-experimenteel onderzoek gevolgd. Ondanks dat de implementatie vaak niet optimaal was (beroepskrachten[1] waren bijvoorbeeld soms niet gecertificeerd om met het programma te werken), waren de resultaten van de eerste onderzoeken veelbelovend. Deze evaluaties hebben de grootschalige inzet van VVE-programma's aan het begin van de 21ste eeuw ingeluid. De positieve effecten van VVE-programma's zijn in Nederland tot op heden echter niet onomstotelijk vastgesteld. Daarbij moet opgemerkt worden dat veel studies waarin geen of zelfs negatieve effecten van VVE-programma's gerapporteerd werden, belangrijke methodologische beperkingen hadden. Bovendien is doorgaans onderzocht of de implementatie van een programma, onafhankelijk van hoe dit programma daadwerkelijk werd uitgevoerd, effect had op de ontwikkeling van deelnemende kinderen.

Hoewel verschillende cohortstudies hebben laten zien dat de prestaties van 'doelgroepkinderen' over de loop van de jaren zijn verbeterd, is er nog een lange weg te gaan, want de groep doelgroepkinderen met een niet-Nederlandse achtergrond blijft nog achter bij Nederlandse doelgroepkinderen.

1 Met beroepskrachten worden in deze samenvatting zowel pedagogisch medewerkers als leerkrachten bedoeld.

Gemengde groepen

Het belang van goede ontwikkelingsmogelijkheden voor jonge kinderen, of dit nu met of zonder VVE-programma is, staat over het algemeen niet ter discussie. Een belangrijke vraag, zowel internationaal als nationaal, is of inspanningen om de ontwikkeling van jonge kinderen te stimuleren zich vooral op doelgroepkinderen zouden moeten concentreren of dat deze zich op alle kinderen zouden moeten richten. In de Nederlandse situatie heeft deze vraag vooral betrekking op de voorschoolse periode, aangezien in Nederland alle kinderen vanaf vier jaar naar de basisschool kunnen. Het voordeel van concentreren op doelgroepkinderen is een financieel voordeel; de groep is simpelweg kleiner.

Een consequentie van een dergelijke geconcentreerde aanpak is dat groepen uitsluitend bestaan uit kinderen met een risico op ontwikkelingsachterstanden, een risico dat bepaald is op basis van criteria zoals opleidingsniveau en culturele achtergrond van ouders. In dergelijke groepen kunnen doelgroepkinderen niet 'profiteren' van de aanwezigheid van niet-doelgroepkinderen. Niet-doelgroepkinderen kunnen op minimaal twee manieren invloed uitoefenen op de taalleeromgeving in de groep en daarmee invloed op de ontwikkeling van groepsgenootjes. Op de eerste plaats is er mogelijk een direct effect; de taalinput van groepsgenootjes is in gemengde groepen mogelijk van een hoger niveau (gevarieerd woordgebruik, complexe zinnen). Omdat taalinput een belangrijke bron is voor de taalontwikkeling van kinderen, zou door het hogere niveau van de taalinput van groepsgenootjes, de taalleeromgeving in gemengde groepen gunstiger kunnen zijn dan die in niet-gemengde groepen. Op de tweede plaats beïnvloeden groepsgenootjes de taalleeromgeving mogelijk ook indirect, namelijk via de beroepskracht. Een gemengde of heterogene groep zou er toe kunnen leiden dat de beroepskracht een ander aanbod, zowel wat betreft activiteiten als taalgebruik, biedt dan in niet-gemengde groepen.

Voor deze jonge leeftijdsgroep is weinig onderzoek gedaan naar het effect van gemengde groepen op de ontwikkeling. Uit onderzoek naar publieke voorschool in de Verenigde Staten en naar de samenstelling van de groep wat betreft sociaal-economische en culturele achtergrond, leeftijd en de vaardigheden van groepsgenootjes, is af te leiden welk effect gemengde groepen mogelijk kunnen hebben. De studies die op deze gebieden zijn gedaan laten wisselende, maar overwegend positieve effecten zien van gemengde groepen op de ontwikkeling van kinderen.

Voor- en vroegschoolse educatie in Nederland

Er zijn over de loop van de jaren veel initiatieven geweest om de uitgangspositie van kinderen bij de aanvang van het onderwijs te verbeteren. In de loop van de jaren '90 werd duidelijk

dat deze initiatieven zich vooral moesten richten op de voorschoolse periode en dan met name op centrumgerichte programma's. Verschillende programma's zijn toen ontwikkeld en de implementatie daarvan is nauwgezet gevolgd middels quasi-experimenteel onderzoek. Positieve evaluaties leidden tot de 'Regeling Voor- en Vroegschoolse Educatie' (Ministerie van OCW). Deze regeling luidde de grootschalige implementatie van VVE-programma's in.

In de voorschoolse periode (tot vier jaar) worden VVE-programma's aangeboden op voorscholen die doorgaans zijn voortgekomen uit peuterspeelzalen die in de jaren zestig voornamelijk op initiatief van ouders ontstonden. In de loop van de jaren is het accent van het leren samenspelen met andere kinderen verschoven naar een combinatie met leerzame activiteiten en sinds de komst van VVE-programma's zijn veel peuterspeelzalen, vooral in buurten en wijken met een hoog percentage doelgroepkinderen, omgevormd naar voorschool. De belangrijkste verschillen tussen voorscholen en peuterspeelzalen zijn dat op de eerste een VVE-programma wordt gebruikt, kinderen vaker komen (minimaal tien uur, meestal verspreid over vier dagdelen), er twee in plaats van één pedagogisch medewerker werkzaam zijn en er een intensieve samenwerking met een nabijgelegen basisschool is. Gemeenten kunnen zelf bepalen welke kinderen tot de doelgroep van de voorschool behoren. Vaak zijn voorscholen vooral of uitsluitend gericht op doelgroepkinderen, zoals bepaald aan de hand van criteria zoals opleidingsniveau en culturele achtergrond van de ouders. Basisscholen zijn in principe gericht op alle kinderen, maar vooral in buurten met een hoog percentage doelgroepkinderen zijn de groepen vaak als vanzelf relatief niet-gemengd.

Deze dissertatie

Een belangrijke vraag, zowel internationaal als nationaal, is of inspanningen om de ontwikkeling van jonge kinderen te stimuleren zich vooral op doelgroepkinderen zouden moeten concentreren of dat ze op alle kinderen gericht zouden moeten worden. In dit proefschrift is het effect van twee factoren van voor- en vroegschoolse educatie op de ontwikkeling van kinderen onderzocht. Ten eerste de implementatie van een VVE-programma, wat tot nu toe wisselende resultaten liet zien in eerder Nederlands onderzoek. Ten tweede is de samenstelling van de groep, wat betreft gemengde en niet-gemengde groepen, op de ontwikkeling van kinderen onderzocht, waarmee een experiment dat in Utrecht werd gedaan met doelbewuste menging van groepen in de voorschoolse groepen is geëvalueerd.

In tegenstelling tot eerder Nederlands onderzoek naar VVE-programma's, dat zich vaak richtte op het feit of er een VVE-programma werd gebruikt, zijn in dit proefschrift gegevens over de daadwerkelijke activiteiten van kinderen (verkregen door verschillende typen observaties) meegenomen in het bepalen van effecten op de ontwikkeling. Rekening

houden met de daadwerkelijke activiteiten van kinderen is belangrijk, omdat uitvoering conform het programma een belangrijke factor is in het bepalen van het succes van een programma. Met andere woorden, dat er 'op papier' een VVE-programma wordt gebruikt, wil niet automatisch zeggen dat het handelen van de beroepskrachten ook conform dat programma is.

Kortom, in dit proefschrift wordt de ontwikkeling van doelgroepkinderen vergeleken tussen: a) groepen waar een VVE-programma werd uitgevoerd en groepen waar geen VVE-programma werd gebruikt en b) gemengde groepen en niet-gemengde groepen.

Voor- en vroegschoolse educatie in Utrecht

In Utrecht is in de voorschoolse periode een experiment gedaan, waarbij ook niet-doelgroepkinderen in aanmerking kwamen voor deelname aan de voorschool. De aanvankelijke aanpak, waarbij alleen doelgroepkinderen in aanmerking kwamen, heeft zoals eerder genoemd belangrijke nadelen die de positieve effecten van deelname aan voorschool mogelijk negatief beïnvloeden. Het experiment bestond eruit dat in wijken met een gemengde populatie gemengde voorscholen werden opgezet. Op gemengde voorscholen bestonden groepen uit tien doelgroepkinderen die net als op reguliere, niet-gemengde, voorscholen vier dagdelen kwamen. Daarnaast waren er twee groepjes van vijf niet-doelgroepkinderen die ieder twee dagdelen, zoals op reguliere peuterspeelzalen, naar de gemengde voorschool kwamen. De basisscholen in buurten waar gemengde voorscholen waren opgezet, waren vanwege de samenstelling van de buurt als vanzelf ook relatief gemengd van samenstelling.

Op alle voorscholen en kleutergroepen met een VVE-programma, zowel de gemengde als niet-gemengde, werd het VVE-programma Taalrijk, een aangepaste versie van het programma Ko Totaal, gebruikt. Dit programma is door de afwisseling van beroepskracht-gestuurde activiteiten in grote en kleine groepen heel geschikt voor gebruik in gemengde en niet-gemengde groepen.

Resultaten

Wat gebeurt er in de groep?

Activiteiten en interactie

In hoofdstuk 2 zagen we dat 'beroepskracht-gestuurde' activiteiten, activiteiten die worden gedaan onder continue begeleiding van de beroepskracht, over het algemeen weinig voorkwamen. In kleine groepjes kwamen deze activiteiten amper voor; als beroepskracht-

gestuurde activiteiten werden gedaan was dit bijna altijd in de grote groep. Wel was er behoorlijke variatie tussen groepen in het voorkomen van deze activiteiten. In hoofdstuk 3 zagen we dat beroepskrachten die geen VVE-programma gebruikten er net zo goed of soms zelfs nog beter in slaagden om een behoorlijk gedeelte van de tijd te besteden aan dit type activiteiten als beroepskrachten die wel een VVE-programma gebruikten. Ook zagen we dat veel tijd 'verloren' gaat aan overgangsactiviteiten zoals wachten en opruimen.

In hoofdstuk 4 zagen we dat kinderen tijdens grote groepsactiviteiten meestal aan het luisteren of helemaal niet bij een interactie betrokken waren. De mogelijkheden om zelf actief een bijdrage te leveren tijdens grote groepsactiviteiten zijn beperkt. In hoofdstuk 5 zagen we dat beroepskrachten zowel tijdens grote als kleine groepsactiviteiten de meerderheid van alle verbale uitingen voor hun rekening nemen. Het merendeel van de uitingen van kinderen wordt tijdens grote groepsactiviteiten bovendien door een minderheid van de kinderen gedaan; het lijkt erop dat beroepskrachten tijdens grote groepsactiviteiten vaak langere tijd in gesprek zijn met een beperkt aantal kinderen. In tegenstelling tot grote groepsactiviteiten, zijn tijdens kleine groepsactiviteiten alle kinderen in de gelegenheid om een substantiële bijdrage te leveren aan de activiteit.

In hoofdstuk 4 zagen we ook dat als peuters een interactie hebben, dit in de meerderheid van de gevallen blijkt te zijn met een beroepskracht. Kleuters daarentegen hebben de meerderheid van hun interacties met groepsgenootjes, wat gedeeltelijk te verklaren is door het grotere aantal kinderen in kleutergroepen. Peuters en kleuters blijken in gemengde groepen meer interacties te hebben met kinderen van Nederlandse afkomst vergeleken met kinderen in niet-gemengde groepen. Deze laatste bevinding hangt sterk samen met de samenstelling van de groep.

Taalgebruik van beroepskrachten: Type vragen en 'abstractieniveau'

In hoofdstuk 5 zagen we dat het taalgebruik van beroepskrachten wat betreft type vragen en het abstractieniveau, het beroep dat op het denkvermogen van kinderen wordt gedaan (vergelijk een vraag als 'Waar is de kat?' met 'Waarom verstopt de kat zich?'), weinig bleek te verschillen tussen verschillende contexten (kleine versus grote groep; gemengde versus niet-gemengde groepen). Open vragen lijken een langer antwoord uit te lokken bij kleuters dan andere soorten vragen. Bij peuters maakte het type vraag geen verschil voor de lengte van het antwoord, maar het abstractieniveau wel; antwoorden waren langer als het abstractieniveau van de uiting van de beroepskracht hoog was. Bij kleuters maakte het abstractieniveau van de uiting van de beroepskracht geen verschil voor de lengte van het antwoord. Voor zowel peuters als kleuters bleek het abstractieniveau van de uiting van de beroepskracht samen te hangen met het abstractieniveau van hun antwoord; als het abstractieniveau van de uiting

van de beroepskracht hoog was, was het abstractieniveau van het antwoord van het kind vaak ook hoog.

Effecten van het VVE-programma

In hoofdstuk 3 zagen we dat de implementatie van een VVE-programma geen effect heeft op de ontwikkeling van deelnemende doelgroepkinderen. Een eerste verklaring is dat het programma niet werd uitgevoerd zoals dat bedoeld was. Dit is af te leiden uit de lage frequentie van beroepskracht-gestuurde academische activiteiten in kleine groepen, maar ook aan de over het algemeen lage frequentie van dit type activiteiten. De lage frequentie van beroepskracht-gestuurde academische activiteiten in kleine groepen is problematisch, omdat dat juist een van de belangrijkste middelen van het programma Taalrijk is om de ontwikkeling van kinderen te stimuleren. Ook zagen we in hoofdstuk 3 dat het VVE-programma niet doorslaggevend is in het aanbod van deze activiteiten; beroepskrachten die geen programma gebruikten slaagden er net zo goed of soms zelfs beter in om een behoorlijk gedeelte van de tijd te besteden aan dit type activiteiten.

De vraag is of alle effecten van deelname aan voor- en vroegschoolse educatie makkelijk te vertalen zijn naar uitkomstmaten zoals die ook in het onderhavige onderzoek zijn gebruikt. Vaak wordt gekeken naar uitkomstmaten als woordenschat, maar misschien zijn er effecten die niet eenvoudig meetbaar zijn. Dankzij de voorschool worden ontwikkelingsachterstanden bijvoorbeeld vaak eerder gesignaleerd en krijgen kinderen en ouders al vroeg passende hulp. Ook het toerusten van ouders met kennis en vaardigheden om in hun handelen beter aan te sluiten bij het ontwikkelingsniveau van hun kind is een belangrijke functie van de voorschool.

Effecten van beroepskracht-gestuurde academische activiteiten

Hoewel in het huidige onderzoek geen effect werd gevonden van de implementatie van een VVE-programma op de ontwikkeling van deelnemende kinderen, zijn er wel aanwijzingen dat beroepskracht-gestuurde activiteiten een effect hebben op de ontwikkeling van kinderen. Dit zijn activiteiten die van begin tot einde worden begeleid door de beroepskracht en ze vormen de kern van VVE-programma's. Het effect van dit type activiteiten op de ontwikkeling van kinderen is in twee studies onderzocht. In hoofdstuk 2 zagen we dat doelgroep- en niet-doelgroepkinderen, vooral peuters, zich na een jaar sneller ontwikkeld hadden in taal- en beginnende lees- en rekenvaardigheid in groepen waar de beroepkracht relatief veel academische activiteiten initieerde en begeleidde (voor kleuters werd alleen een effect gevonden voor de beginnende rekenvaardigheid).

In hoofdstuk 3 is alleen gekeken naar de effecten van dit type activiteiten op de ontwikkeling van doelgroepkinderen. Hier zagen we dat beroepskracht-gestuurde academische activiteiten vooral een positief effect hadden op de ontwikkeling van kleuters; na twee jaar deelname bleken zij zich sneller in taal- en beginnende lees- en rekenvaardigheid ontwikkeld te hebben dan kinderen in groepen waar de beroepskracht minder van deze activiteiten aanbood (voor peuters werd alleen een effect gevonden voor de beginnende rekenvaardigheid). Hoewel beroepskracht-gestuurde activiteiten over het algemeen weinig en onafhankelijk van het VVE-programma voorkwamen, was er variatie tussen groepen en bleken de activiteiten samen te hangen met de ontwikkeling van kinderen.

Effecten van gemengde groepen

Hoofdstuk 3 laat zien dat doelgroepkinderen, vooral kleuters, zich sneller ontwikkelden in gemengde groepen dan kinderen in niet-gemengde groepen (voor peuters werd alleen een effect gevonden voor de beginnende rekenvaardigheid). In hoofdstuk 4 zijn de effecten van gemengde groepen verder ontrafeld door te laten zien dat de daadwerkelijke interactie met Nederlandse groepsgenootjes samenhangt met ontwikkeling in de taal- en beginnende leesvaardigheid van kleuters; als kinderen meer interacties hadden met Nederlandse groepsgenootjes lieten zij een snellere ontwikkeling zien in taal- en beginnende leesvaardigheid. Voor peuters werd geen effect gevonden.

Tot slot is in hoofdstuk 5 gekeken of de effecten van gemengde groepen mogelijk via het aanbod van de beroepskracht verlopen. Er zijn geen noemenswaardige verschillen gevonden tussen gemengde en niet-gemengde groepen in het type vragen dat werd gesteld en het abstractieniveau van uitingen. Dit proefschrift lijkt er dus op te wijzen dat de effecten van gemengde groepen via de aanwezigheid van groepsgenootjes verlopen en niet via een aanpassing van het aanbod van de beroepskracht.

Een verklaring voor de over het algemeen positievere resultaten van gemengde groepen voor kleuters is dat zij meer interacties hebben met hun groepsgenootjes dan peuters; in de kleutergroepen komt een groter gedeelte van de taalinput van groepsgenootjes. Een andere verklaring is dat de gemengde kleutergroepen 'gemengder' waren dan de gemengde peutergroepen; gemengde peutergroepen bestonden voor een relatief kleiner gedeelte uit Nederlandse kinderen. Tot slot zijn peuters vergeleken met kleuters veel minder aanwezig op de voorschool; peuters zijn er doorgaans 10 à 12 uur, terwijl kleuters ongeveer 25 uur per week naar school gaan.

Aanbevelingen

Voor beroepskrachten

Gezien de beperkte mogelijkheden voor kinderen om een actieve (verbale) bijdrage te leveren aan activiteiten in de grote groep, wordt aanbevolen om het aantal activiteiten in de grote groep tot een minimum te beperken en in plaats daarvan activiteiten zoveel mogelijk in kleine groepjes te doen. Activiteiten in kleine groepjes hebben twee belangrijke voordelen vergeleken met de grote groep. Ten eerste biedt de kleine groep veel meer mogelijkheden voor kinderen om zelf een actieve (verbale) bijdrage te leveren. Ten tweede is het in een kleine groep voor de beroepskracht makkelijker om het aanbod aan te passen aan het ontwikkelingsniveau van de kinderen in het groepje. In de grote groep is het door het grote aantal kinderen vrijwel onmogelijk om een activiteit aan te bieden die voor alle kinderen passend is.

Verder wordt aanbevolen om de ontwikkeling van kinderen goed te volgen en deze informatie te gebruiken om betekenisvolle passende kleine groepsactiviteiten te ontwerpen. Daarbij moet overigens opgemerkt worden dat het belangrijk is om het aanbod te plannen, maar dat het grijpen van kansen in spontane situaties minstens zo belangrijk is. Kinderen leren tenslotte het meest wanneer zij geïnteresseerd en gemotiveerd zijn.

Uit de onderzoeksgegevens bleek dat veel tijd 'verloren' gaat aan wachten en opruimen. Om optimaal te profiteren van een VVE-programma is het belangrijk om de tijd zo effectief mogelijk te besteden, zodat er veel gelegenheid is voor betekenisvolle interessante activiteiten. Maar ook tijdens routines zijn er kansen te grijpen, soms door heel eenvoudige ingrepen zoals kinderen zich bij het naar buiten gaan op laten stellen in een rij van groot naar klein.

Tot slot is het heel belangrijk dat de taalleeromgeving een rijke is, ook wat betreft de taalinput van de beroepskrachten. Het taalgebruik moet niet versimpeld worden; kinderen leren nieuwe woorden, uitdrukkingen en zinsconstructies alleen als ze er herhaaldelijk mee in aanraking worden gebracht. Taalgebruik en daar feedback op krijgen is minstens zo belangrijk voor de taalontwikkeling als het horen van taal, de taalinput. Daarom is het belangrijk om open en uitdagende vragen te stellen en kinderen daarbij tijd geven om een antwoord op de vraag te formuleren. Ook het begeleiden en stimuleren van interacties tussen kinderen, vooral als er grote verschillen zijn in taalvaardigheid, is een belangrijke taak van de beroepskracht.

Voor beleid

Een mogelijke sleutel om de opvang en educatie van jonge kinderen verder te verbeteren is continue professionalisering van beroepskrachten, waarbij coaching een van de middelen kan zijn. Dit proefschrift laat niet als eerste studie zien dat de implementatie van een programma geen garantie is voor handelen conform dat programma. Er lijkt meer nodig te zijn dan een eenmalige training om er voor te zorgen dat beroepskrachten conform het programma (blijven) handelen en coaching is daarbij, zoals gebleken uit internationaal onderzoek, een beproefd middel. Kernbestanddelen van continue professionalisering zijn reflectie op het eigen handelen en dat van collega's, monitoren van effecten van het aanbod op de ontwikkeling van kinderen en de bereidheid om te experimenteren met nieuwe handelswijzen. Kortom, het is belangrijk dat er op alle niveaus binnen organisaties een lerend klimaat is. Alleen dan kunnen beroepskrachten zich ontwikkelen naar kritische en reflecterende professionals.

De resultaten van dit proefschrift laten zien dat gemengde groepen een voordeel hebben voor deelnemende doelgroepkinderen, met name voor kinderen in kleutergroepen. Mogelijk zouden er meer positieve effecten van menging zijn als het voorschoolse systeem, net als de basisscholen, universeel zou zijn, dat wil zeggen dat alle kinderen deelnemen. Op dit moment zijn de voorzieningen voor kinderen jonger dan vier jaar versnipperd en 'gescheiden'; voorschool wordt veelal bezocht door doelgroepkinderen, terwijl kinderen van hoger opgeleide werkende ouders veelal naar kinderdagverblijven gaan. Deze scheiding hangt samen met de hogere kosten van kinderdagverblijven, maar ook met het gegeven dat werkende ouders hele dagopvang voor hun kinderen nodig hebben, die alleen wordt aangeboden door kinderdagverblijven (voorschool biedt doorgaans alleen halve dagopvang). Een integratie van de verschillende voorzieningen, aan de ene kant kinderdagverblijven die over het algemeen meer gericht zijn op verzorging, en aan de andere kant voorscholen, die meer gericht zijn op educatie en ontwikkeling, zou wel eens positieve effecten kunnen hebben op de ontwikkeling van alle deelnemende kinderen, onafhankelijk van hun achtergrond.

Conclusie

In dit proefschrift werd geen effect gevonden van de implementatie van een VVE-programma op de ontwikkeling van deelnemende doelgroepkinderen. Een andere belangrijke bevinding, die meteen een verklaring vormt voor de afwezigheid van een effect van het geïmplementeerde VVE-programma, is dat het programma niet werd uitgevoerd zoals het bedoeld was. Deze conclusie werd getrokken op basis van het gegeven dat de observaties lieten zien dat er over het algemeen weinig beroepskracht-gestuurde actviteiten in kleine groepjes werden gedaan

en dat beroepskracht-gestuurde activiteiten over het algemeen weinig voorkwamen. Verder zagen we dat beroepskrachten in groepen waar geen VVE-programma was geïmplementeerd er net zo goed of soms zelfs nog beter in slaagden om een behoorlijk gedeelte van de tijd te besteden aan beroepskracht-gestuurde activiteiten. We zagen echter dat dit type activiteiten wel samenhangt met de ontwikkeling van kinderen. Kortom, de implementatie van een VVE-programma is niet doorslaggevend voor het aanbod van beroepskracht-gestuurde activiteiten, maar deze activiteiten zijn wel belangrijk voor de ontwikkeling van deelnemende kinderen. Dit proefschrift laat zien dat het niet de vraag is of er een programma wordt gebruikt, maar eerder hoe de beroepskracht het programma hanteert.

Een andere belangrijke bevinding van dit proefschrift is dat de ontwikkeling van doelgroepkinderen, met name in de kleutergroepen, gebaat is bij deelname in gemengde groepen; doelgroepkinderen laten een snellere ontwikkeling zien in gemengde groepen dan in niet-gemengde groepen. Het positieve effect van gemengde groepen lijkt te verlopen via interactie met groepsgenootjes en niet via het aanbod van de beroepskracht. Om de kwaliteit van voor- en vroegschoolse educatie verder te verbeteren wordt continue professionalisering, waarbij coaching een beproefd middel is, aanbevolen. De ontwikkeling van beroepskrachten naar kritische en reflecterende professionals komt het leren en ontwikkelen van de kinderen in hun groepen waarschijnlijk ten goede.

Dankwoord
(Acknowledgements)

Dit proefschrift heeft met name in de laatste jaren, toen ik niet meer in Utrecht werkte, veel van mijn doorzettingsvermogen gevergd. Het advies, de steun en aanmoediging van veel mensen hebben ervoor gezorgd dat ik de eindstreep heb gehaald. Als ik iedereen die een steentje heeft bijgedragen persoonlijk moet bedanken, kan ik met gemak nog een boekje vullen. Een aantal mensen wil ik op deze plaats graag persoonlijk bedanken.

Allereerst mijn promotoren Ed Elbers en Paul Leseman. Ik waardeer het dat jullie mij, ondanks dat ik uit een heel andere hoek kwam, de kans hebben gegeven om dit project uit te voeren. Ook bedankt voor het geduld dat jullie met me hebben gehad, ik heb veel van jullie geleerd.

Het onderzoek naar de gemengde groepen was er zonder de inspanning van Ina Vugts niet geweest. Ina, ik heb enorme bewondering voor je inzet en vastberadenheid om het 'onderwijs' voor jonge kinderen te verbeteren. De bijeenkomsten met de expertgroep en met de deelnemende locaties waren altijd inspirerend, alle deelnemers daaraan bedankt voor jullie bijdrage. Een speciaal woord van dank gaat uit naar Jarise Kaskens, Angelique Hoogstrate en Karolien Schmidt van Eduniek, die een belangrijke rol speelden bij de organisatie van de diverse bijeenkomsten en meedachten over het verloop van het onderzoek en interpretatie van resultaten.

Laura Justice and all the people at the CLRC, thanks for having me in Columbus. My visit was very inspiring and fun, a great combination!

Ik wil ook graag alle (oud-)collega's van de vakgroep bedanken voor de samenwerking en gezelligheid. In het bijzonder Annika, Ludger, Meijke en Sylke. Als 'lachen' een criterium is voor hoe leuk het is geweest, dan was mijn periode als aio ontzettend leuk en dat voor een groot gedeelte dankzij jullie! Ik herinner me verhalen over kabouters, of waren het clowns, waar ik letterlijk om gierde van het lachen! Niet alleen was het vaak gezellig, ik heb ook veel gehad aan jullie steun en advies. Ik hoop op nog meer bier, wijn, cassis, fanta en bitterballen met de bende! Met andere Annie en Ludger is dat de laatste jaren al goed gelukt en wat heb ik vaak spierpijn in mijn wangen na die avonden. Geen grap is te hard en dat is te gek! Natte tranen van het lachen, mooier kan het leven niet worden ;-).

Een speciaal woord van dank is voor mijn kamergenoot Meijke, niet alleen konden we altijd bij elkaar terecht voor advies en vragen, ik heb ook veel respect voor het doorstaan van mijn 'dolle kwartiertjes' met slechte zang, foute grappen en andere onzin. Ik ben blij dat je de spannendste drie kwartier achter me staat!

En dan André, wat ben ik blij dat ik je heb leren kennen: mattie van de overkant, gast, Sint Oedenrode of tegenwoordig meestal gewoon Miranda: bedankt!

Ook wil ik mijn – inmiddels niet meer zulke – nieuwe collega's bij Impuls bedanken voor de aanmoediging. Omdat jullie bleven vragen, verdween het proefschrift nooit helemaal uit

mijn gedachten en heb ik uiteindelijk met de nodige horten en stoten de eindstreep gehaald. Bedankt Aaf, Anita, Elly, Marion, Mascha, Naima en alle andere collega's.

Mijn ouders wil ik bedanken voor de mogelijkheid die ze mijn zus, broer en mij hebben gegeven om ons eigen pad te kiezen (dat niet altijd het meest voor de hand ligt). Door de kansen die jullie mij hebben gegeven, hebben jullie indirect hopelijk bijgedragen aan de schoolloopbanen van minder bevoorrechte kinderen.

De belangrijkste rol bij de totstandkoming van dit proefschrift was voor de pedagogisch medewerkers, leerkrachten en VVE-coördinatoren van een groot aantal peuterspeelzalen, voorscholen en basisscholen in Utrecht en omgeving. Niet alleen was dit proefschrift er zonder jullie niet geweest, jullie zijn onmisbaar voor de ontwikkeling van de kinderen in jullie groepen en klassen!

Curriculum vitae

Annika de Haan (1982) was born in Assen, The Netherlands. After moving from 'the land of Bartje' at the age of five, she spent her childhood strolling around farmer Pelleberg's lands, leaping over ditches, playing tennis and attempting to start a fire. She attended high school at the Isendoorn College in Warnsveld. She unsuccessfully pursued a career in every imaginable sport (except curling and archery) in order to accomplish her dream: an Olympic gold medal. She completed the first year of law studies at the University of Groningen and subsequently obtained a Bachelor's degree in Educational Science from the Radboud University in Nijmegen. In 2007 she obtained a Research Master's degree in Behavioral Science from that university. In October 2007 she started working as a PhD-candidate at Utrecht University. She presented her work at several national and international conferences. In addition to her dissertation work, Annika was also a workgroup teacher for the BA-course Cultural Diversity and she supervised bachelor and master theses. During her work as a PhD-candidate she was representative for the PhD-candidates of the former Langeveld Institute in the Faculty of Social Sciences' PhD-council. In 2011 Annika spent a month at professor Laura Justice's lab at Ohio State University. After finishing this dissertation Annika will pursue dreams such as travelling the Trans Mongolia Express, climbing the Kilimanjaro in Tanzania, walking to Santiago de Compostela, and finally starting a fire herself.

For further information, see nl.linkedin.com/in/annikadehaan