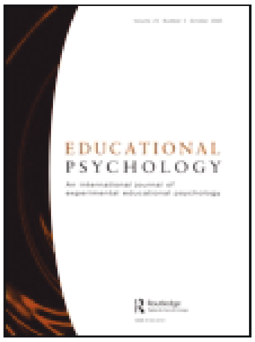
van den Heuvel-Panhuizen, Marja; Elia, Iliada; Robitzsch, Alexander

Effects of reading picture books on kindergartners' mathematics performance

Educational Psychology 36 (2016) 2, S. 323-346

Quellenangabe/ Reference:
van den Heuvel-Panhuizen, Marja; Elia, Iliada; Robitzsch, Alexander: Effects of reading picture books on kindergartners' mathematics performance - In: Educational Psychology 36 (2016) 2, S. 323-346 - URN: urn:nbn:de:0111-pedocs-126872 - DOI: 10.25656/01:12687

https://nbn-resolving.org/urn:nbn:de:0111-pedocs-126872
https://doi.org/10.25656/01:12687



Routledge
Taylor & Francis Group

Educational Psychology
An International Journal of Experimental Educational Psychology

ISSN: 0144-3410 (Print) 1469-5820 (Online) Journal homepage: http://www.tandfonline.com/loi/cedp20

Effects of reading picture books on kindergartners' mathematics performance

Marja van den Heuvel-Panhuizen, Iliada Elia & Alexander Robitzsch

To cite this article: Marja van den Heuvel-Panhuizen, Iliada Elia & Alexander Robitzsch (2016) Effects of reading picture books on kindergartners' mathematics performance, Educational Psychology, 36:2, 323-346, DOI: 10.1080/01443410.2014.963029

To link to this article: http://dx.doi.org/10.1080/01443410.2014.963029

Educational Psychology, 2016
Vol. 36, No. 2, 323–346, http://dx.doi.org/10.1080/01443410.2014.963029

Effects of reading picture books on kindergartners' mathematics performance

Marja van den Heuvel-Panhuizen[a]*, Iliada Elia[b] and Alexander Robitzsch[c]

[a]*Faculty of Science & Faculty of Social and Behavioural Sciences, Freudenthal Institute for Science and Mathematics Education, Utrecht University, Utrecht, The Netherlands;* [b]*Department of Education, University of Cyprus, Nicosia, Cyprus;* [c]*Federal Institute for Education Research, Innovation and Development of the Austrian School System, Salzburg, Austria*

(*Received 8 August 2013; final version received 29 August 2014*)

This article describes a field experiment with a pretest–posttest control group design which investigated the potential of reading picture books to children for supporting their mathematical understanding. The study involved 384 children from 18 kindergarten classes in 18 schools in the Netherlands. During three months, the children in the nine experimental classes were read picture books. Data analysis revealed that, when controlled for relevant covariates, the picture book reading programme had a positive effect (d = .13) on kindergartners' mathematics performance as measured by a project test containing items on number, measurement and geometry. Compared to the increase from pretest to posttest in the control group, the increase in the experimental group was 22% larger. No significant differential intervention effects were found between subgroups based on kindergarten year, age, home language, socio-economic status and mathematics and language ability, but a significant intervention effect was found for girls and not for boys.

Keywords: kindergarten; mathematics; intervention

Introduction

Using picture books for teaching mathematics in kindergarten

For many children between the ages of four and six, kindergarten is the first institutional educational setting in which they come across school subjects, including mathematics. The teaching of mathematics to children of that young age already has a long history (Saracho & Spodek, 2009), dating back to 1631 when Comenius published his book *School of Infancy* and in this way stimulated the creation of mathematics programmes for young children. Comenius emphasised the observation and manipulation of objects and even used a picture book to support children to make impressions in the mind (Schickedanz, 1995). In our times, making use of picture books for supporting children's mathematical understanding has become increasingly popular since the last decade (Haury, 2001; National Council of Teachers of Mathematics [NCTM], 2000). By picture books, we mean books consisting of text and pictures, 'in which the story depends on the interaction between written text and

*Corresponding author. Email: m.vandenheuvel@fi.uu.nl

image and where both have been created with a conscious aesthetic intention' (Arizpe & Styles, 2003, p. 22).

Even though reading picture books might not seem very suitable for teaching mathematics, stories narrated in a book may contain mathematics, and as such they offer children opportunities to face mathematics (Anderson, Anderson, & Shapiro, 2005). In agreement with the comprehensive interpretation of mathematics to be taught to young children that can be found in many mathematics curricula and standard documents worldwide, such as of the NCTM and the National Association for the Education of Young Children (NAEYC) (see NCTM, 2000; NAEYC & NCTM, 2002), the mathematics that children can encounter in picture books should be viewed in a broad sense. This means that in addition to the usual mathematical topics, such as number, measurement and geometry, as well as mathematical processes and dispositions and mathematics-related themes, including, for example, growth, patterns, fairness, and cause and effect can be considered as mathematical content (Van den Heuvel-Panhuizen & Elia, 2012).

A very important reason why reading picture books to children may help them in learning mathematics has to do with the meaningful context of the stories included in picture books (Columba, Kim, & Moe, 2005; Moyer, 2000; Whitin & Wilde, 1992). Also Egan (1985) emphasised that the story form is a basic tool for establishing meaning. Moreover, research about learning word meanings suggests that learning within a story context increases the retention and recall of the learned knowledge (Biemiller & Boote, 2006; Horst, Parsons, & Bryan, 2011). In particular, a number of cognitive scientists consider the story as 'the most natural package of organised knowledge in the cognitive system for acquiring and retaining information' (Casey, Andrews, et al., 2008, p. 276). In agreement with this, Lovitt and Clarke (1992) pointed out that picture books can offer cognitive hooks to explore and construct mathematical concepts and skills.

The theoretical grounding for the necessity of having a meaningful context for developing mathematical understanding is also based on Freudenthal's (1983) ideas of mathematics as a meaningful activity and is supported by the Vygotskian and action psychological approach to learning (Van Oers, 1996). According to Van Oers (1996), learning can endorse the personal and cultural development of a person only when it is meaningful. Such meaningful situations can be created either when a teacher is reading a picture book or when the children are 'reading' a picture book by themselves during free play. A picture book may enable children to encounter problematic situations, may stimulate them to ask their own questions, search for answers, consider different points of view, exchange views with others and incorporate their own findings with existing knowledge. In this way, picture books can support children in attaching personal meaning to the mathematical objects involved in the books.

Apart from the mathematics included in the picture books and the accompanying meaningful contexts that help to understand this mathematics, other so-called 'learning-supportive' characteristics were also identified by Van den Heuvel-Panhuizen and Elia (2012). Based on a literature review and an expert consultation they carried out, it emerged that it is particularly important that picture books offer participation opportunities to children. This implies that children should be triggered to be involved cognitively, emotionally or physically by asking questions, providing explanations and surprising them.

Earlier studies about effect of using picture books on mathematics achievement

Although most studies about reading children picture books focus on the effect on early language and literacy development (see, e.g. Blok, 1999; Collins, 2010; Hindman, Connor, Jewkes, & Morrison, 2008; Mol, Bus, & de Jong, 2009; Mol, Bus, de Jong, & Smeets, 2008), several studies have also been carried out that investigated the effect of reading picture books on young children's learning of mathematics. In most of these studies, the book reading sessions in class were followed by other activities such as playing with story-related (mathematical) materials (Hong, 1996; Jennings, Jennings, Richey, & Dixon-Krauss, 1992; Young-Loveridge, 2004), singing mathematical rhymes (Young-Loveridge, 2004) or composing geometrical puzzles (Casey, Erkut, Ceder, & Mercer Young, 2008). The studies on these combined approaches of book reading and book-related activities generally showed positive results. In particular, studies found a positive effect on kindergartners' mathematics achievement in general (Hong, 1996; Jennings et al., 1992; Young-Loveridge, 2004), their geometric skills (Casey, Erkut, et al., 2008), their attitude towards mathematics (Hong, 1996; Jennings et al., 1992) and their use of mathematical vocabulary (Jennings et al., 1992). However, effects on standardised mathematics tests were not always found (Hong, 1996; Jennings et al., 1992).

In addition to studies that combine book reading with activities in class related to the content of the books, there were also studies which focused on the potential of book reading itself (e.g. Anderson et al., 2005; Van den Heuvel-Panhuizen & Van den Boogaard, 2008). One finding of these studies was that reading picture books to young children resulted in unprompted mathematics-related utterances. In a study by Anderson et al. (2005), parents read a picture story book to their four-year-old child, which resulted in them engaging in mathematical talk with their children spontaneously. The study by Van den Heuvel-Panhuizen and Van den Boogaard (2008) also showed that during book reading sessions children spontaneously came up with mathematics-related utterances, including statements about the numerosity of a collection, using undefined quantifiers such as 'all' or 'everyone' and expressing spatial positions. Although the picture book that was read to the children in the latter study was not specifically written for teaching mathematics, surprisingly, nearly half of their utterances were mathematics-related. In fact, this result confirms the findings of Anderson et al. (2005), because they also used literary trade books which were originally not intended for instruction.

Another research finding to take into account is that children might benefit differently from mathematics programmes that use children's literature. Particularly, with respect to gender, there is some evidence that it could play a role. For example, the study by Casey, Erkut, et al. (2008) revealed that their literature-embedded mathematics intervention, which addressed geometric skills, was more advantageous to girls than to boys. Moreover, Millard (1997) suggests that the focus on narrative books in school would have more appeal for girls than boys, and Wigfield and Guthrie (1997) found that girls' reading motivation was more positive than boys'. Also the finding of Klein, Adi-Japha, and Hakak-Benizri (2010) that the mathematical communication level provided in teacher–child interaction in kindergarten was related to girls' but not to boys' mathematics performance supports the idea that girls would gain more from picture book reading than boys.

Regarding age, several studies found differences in mathematics achievement as early as first grade between children at the same grade level but of different age, favouring the older ones (Gullo & Burton, 1992; Jordan, Kaplan, Locuniak, & Ramineni, 2007; Stipek & Byler, 2001), yet we did not find studies that have shown that older children benefitted more than younger children from an intervention with children's literature.

The socio-economic status (SES) of kindergartners may also be an important factor to consider when investigating the effect of an intervention based on children's literature, although previous research did not indicate what effect to expect from a picture book reading programme. There is some evidence that children with a lower SES would profit more from such a programme. For example, Jennings et al. (1992) and Young-Loveridge (2004) found positive effects of literature-based mathematics programmes for children from low-economic communities. What this means for children with a high socio-economic background is unclear. For example, in Hong's (1996) study, which involved children with highly educated parents, the children's scores were raised in a project test, but not in a standardised mathematics test. More in general, research has shown that, at kindergarten age, children of low SES achieved lower on mathematics tests compared with children of middle SES (Jordan, Kaplan, Nabors Oláh, & Locuniak, 2006; Sarama & Clements, 2009a; Starkey, Klein, & Wakeley, 2004; Wang, 2010). However, in free play few or no differences between low- and middle-income children were found in the amount of mathematics they exhibited (Ginsburg, Chia-Ling, Ness, & Seo, 2003; Seo & Ginsburg, 2004). According to Sarama and Clements (2009b, p. 332), this contradiction may among other things appear because of low-income children's 'lack of opportunities to engage in the language and conversations necessary to bring implicit mathematical ideas to an explicit level'. In other words, the language abilities of children could also be important requirement for the learning of mathematics (Aiken, 1972; Anders et al., 2012; Clarkson, 1992). In line with this, when children speak a different language at home than in school, which can cause that they have difficulties in following the story, this might negatively influence the effect of reading picture books on their mathematics performance.

The present study

The studies described above indicated that using children's literature might be a promising avenue to contribute to the development of children's understanding of mathematics, but they also left us with many unanswered questions. The present study is meant to gain more knowledge about the effect of reading picture books on kindergartners' mathematics performance and to explore how this effect varies with respect to child characteristics. In contrast to most previous effect studies on the use of picture books in mathematics education, our investigation focused on the effect of the book reading itself, i.e. without inclusion of additional (book-related) mathematical activities. Moreover, different from earlier studies, we also examined the relationship between characteristics of children and the effect of the picture book reading. The study was carried out in the Netherlands and was part of the PICO project (PIcture books and COncept development MAthematics).

Our first research question was: *Can an intervention involving picture book reading contribute to children's mathematics performance?* Based on research discussed earlier, our prediction was that kindergartners' performance in a mathematics test would increase due to the picture book intervention, i.e. we hypothesised a positive intervention effect (Hypothesis 1).

Our second research question was: *Is there a relationship between the intervention effect and particular characteristics of children?* In other words: Do differential intervention effects exist? Here, we investigated the role of age, kindergarten year (i.e. the number of years children attended in kindergarten), gender, mathematics and language ability, home language and SES. Based on the findings of Casey, Erkut, et al. (2008) and other studies that provided reasons why picture book reading would be more advantageous to girls than to boys (Klein, Adi-Japha, & Hakak-Benizri, 2010; Millard, 1997; Wigfield & Guthrie, 1997), we only hypothesised a differential intervention effect for gender; expecting that girls would benefit more from the intervention than boys (Hypothesis 2). With respect to the other characteristics, we could not make predictions. One might think that the profit in learning mathematics from a picture book reading programme would be higher for older children and children in a higher kindergarten year, and for children with higher mathematics and language abilities, but we did not find studies that investigated these relationships. For SES and the children's home language, we also could not make predictions. Previous studies did not reach converging results. On the one hand, literature-based mathematics programmes were found to be effective for lower SES children (Jennings et al., 1992; Young-Loveridge, 2004). On the other hand, in general, research has shown that low-SES children achieved lower on mathematics tests compared with children of middle SES (Jordan et al., 2006; Sarama & Clements, 2009a; Starkey et al., 2004; Wang, 2010). Moreover, low-SES children might lack opportunities to engage in classroom conversations (Sarama & Clements, 2009b), because many of them speak a different language at home than in school. However, we did not find studies that showed that having a different home language influenced the learning of mathematics based on reading children picture books.

Method

To investigate the effect of reading picture books on young children's mathematics performance, a field experiment was carried out in kindergarten classes based on a pretest–posttest control group design with a picture book reading programme as an intervention carried out by the classroom teachers.

Participants

Our sample was based on a clustered sampling of matched pairs of schools with random allocation to intervention. We started with a list of all schools in the province of Utrecht. To get a representative sample regarding school location, we classified these schools (about 360 in total) in three categories of urbanisation level, ranging from schools in small towns to schools in large towns. To minimise the influence of schools with non-typical programmes on the effect of the intervention, we then excluded all schools with a special educational approach such as Montessori schools or Peter Petersen schools. To enable a comparison between children with different years of schooling, we also excluded schools that did not have combined kindergarten classes involving both children who are in kindergarten year 1 (K1) and kindergarten year 2 (K2). Because we wanted to control for the children's mathematics and language ability, we also eliminated schools that did not administer the tests developed by the Central Institute for Test Development (Cito) for measuring kindergartners' mathematics and language ability. These tests are very widely used in schools in the Netherlands to monitor children's developments. Both tests have different versions for K1 and K2 children. The reliability of the K1 version of the Cito mathematics test is .85 and of the K2 version it is .81. The reliability of the K1 version of the Cito language test is .86 and for the test in K2 it is .89 (Van Kuyk & Kamphuis, 2001).

From the remaining group of about 80 schools, we first composed 25 pairs of schools that were approximately similar regarding urbanisation level, school size and average SES of their children. The SES data were provided by the Dutch Ministry of Education on 27 February 2007. These data belong to a system that is used in the Netherlands to determine the number of teachers a school can appoint. Children with low-educated parents and/or with immigrant parents get an extra 'weight'. Hereafter, we call the children with a high 'weight', children with a low SES. The children with no extra 'weight' are called children with medium/high SES. However, we should note that this 'SES weight', which is the only information that was available about the SES of the children, is not an accurate measure of their SES as it does not differentiate between the ethnicity background of their parents and the education of their parents. Moreover, this SES weight is a general measure and is not split out for mother and father.

The schools in each pair were assigned randomly to the experimental group or the control group. Next, we invited each of these 50 schools to participate with one kindergarten class. When a school was not willing to participate we searched among the remaining schools in the 80 schools' sample for a comparable school. However, we did not find always a good replacement, or when we found one, schools did not always accept our invitation to participate. After several trials, we ended up with nine pairs of schools willing to participate, which we considered to be a sufficient number of children for the study.

In total, we had a sample of 384 children: 199 in the experimental group and 185 in the control group. As is shown in Table 1 both groups were quite similar. They had about the same average class size, proportions of K1 and K2 children, of girls and boys, of children with non-Dutch and Dutch home language and children with low SES and medium/high SES. Also the age of the children did not differ between the experimental group and the control group. The same is true for the children's scores in the Cito mathematics test and the Cito language test.

Material

Picture books and reading guidelines

The books used in the intervention are picture books which contain text, but in which the illustrations are essential for telling the story. All books are trade books of high literary quality in which the authors unintentionally addressed mathematical topics. Thus, although the books were not purposely written for teaching children mathematics, they have mathematics-related content. To cover a rich variety of mathematical domains, we chose picture books dealing with number, measurement or geometry. Within these domains, we focused respectively on numbers and number relations, growth and perspective. Altogether, eight books were selected within each domain. The 24 picture books were piloted in two schools which were not part of the experiment. In one school, the pilot sessions took place in a small group of children and in the other school, the books were read to a whole kindergarten class.

Table 1. Sample composition.

Child characteristic	Frequency: Experimental group (199 children)		Frequency: Control group (185 children)		d	p
Kindergarten year					−.14	.19
Children in K1	84 (42%)		66 (36%)			
Children in K2	115 (58%)		119 (64%)			
Gender					−.04	.71
Boy	106 (53%)		95 (51%)			
Girl	93 (47%)		90 (49%)			
Home language					−.08	.43
Non-Dutch	28 (14%)		21 (11%)			
Dutch	171 (86%)		164 (89%)			
SES					.04	.67
Low	23 (12%)		24 (13%)			
Medium/high	176 (88%)		161 (87%)			
	M	*SD*	*M*	*SD*		
Age in years at time of pretest						
K1	4.64	.37	4.69	.37	−.14	.39
K2	5.70	.44	5.67	.46	.08	.55
K1 + K2	5.28	.62	5.31	.62	−.05	.65
Cito mathematics at time of pretest						
K1	43.2	13.8	41.8	10.6	.15	.46
K2	57.9	13.7	54.3	10.4	.28	.02
K1 + K2	51.7	15.6	49.8	12.0	.15	.17
Cito language at time of pretest						
K1	58.2	8.9	66.3	6.8	−1.02	.67
K2	73.6	10.6	72.6	10.6	.09	.48
K1 + K2	69.2	12.2	70.7	10.6	−.13	.66
Children's class size	22.6	3.2	22.3	4.8	.07	.47

The pilot sessions were meant to check whether the books are appropriate, i.e. whether they really give opportunities to discuss mathematics-related concepts and provide children with a meaningful context to support their understanding of these concepts. Furthermore, based on the experiences from the pilot sessions, we developed for every book a reading guideline in which we explained to the teacher how to read the book in class. To develop these guidelines, first, each book was piloted in the small-group setting with a preliminary reading guideline. Then we analysed the video recording of the reading session and if necessary this reading guideline was revised. A week later the same book was read in a whole-class setting. If necessary, this was followed by a further revision resulting in the final version of the reading guideline as used in the experiment.

A reading guideline describes for each page of a book how to read it. Next, an example is given for the picture book *Ga je mee?* [Let's go] (Dematons, 2005) which is one of the eight picture books for the domain of geometry, i.e. the topic of perspective. The story in the picture book is about a boy who lives in a house with a large garden. His mother asks him to go and buy apples. Walking through the garden on his way to the greengrocer, in his imagination he undergoes the most exciting adventures. Since the picture book shows the boy walking through the garden from a bird's-eye view, the reader can see all kinds of dangers before the boy is aware of them. The boy in the story knows this and frequently turns towards the reader to ask to help him get through the garden safely.

In Figure 1, the boy is depicted on the left, in a rowing boat on a river (which is actually the garden pond), heading for the sea on the right. He is looking for a safe route and should be careful because of the crocodile at the top of the left page.

The reading guideline states:

> Read the text on page 11. Let a child point out the sea. (The sea is on the extreme right on page 12.) Then read the text on page 12 up to 'You are more able to see it, from above'. Say: 'Yes, we are more able to see the sea.' Ask: 'What else can we see better than the boy?' (The crocodile.) Ask: 'And what can the boy see better?' (The passage under the high rock in the middle of page 11. On most of the drawings in this book, we as readers see more than the boy, but on this page, there is something the boy can see which we cannot see, namely whether or not there is a passage under the rock. The boy must look behind him to see this, because he is rowing backwards.) Ask: 'Which route do you think the boy will take?' (Through the passage). Read the last line on page 12.

In general, the reading guidelines requested the teachers to maintain a reserved attitude and not to take each aspect of the story as a starting point for a class discussion, since lengthy or frequent intermissions could break the flow of being in the story and consequently diminish the story's own power to contribute to the

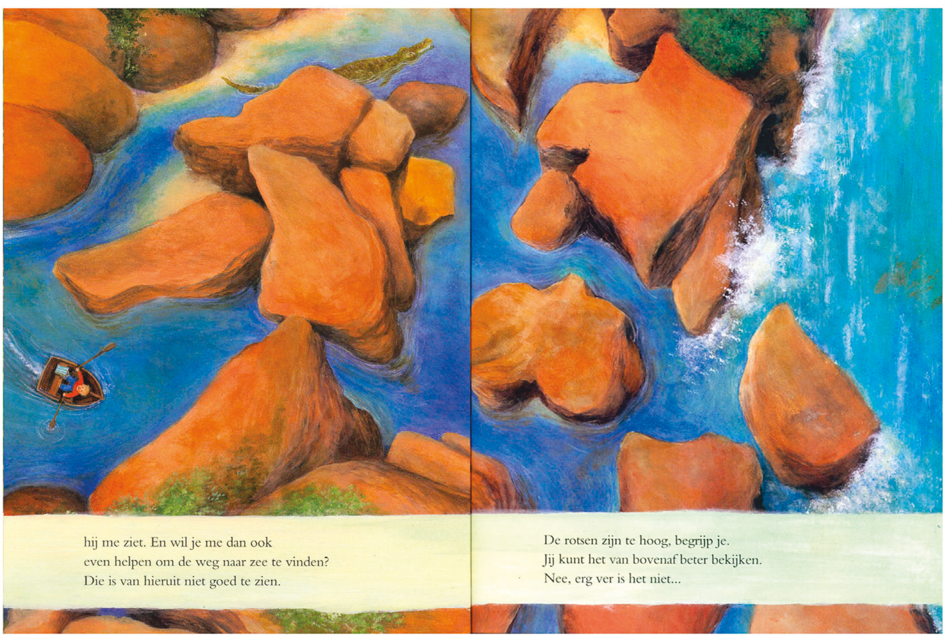


Figure 1. Pages 11 and 12 from *Ga je mee*? [Let's go?] (Dematons, 2005) (illustrations by author, Charlotte Dematons). Translation of text on page 11 (left side): '[…] and would you please help me find the way to the sea? From here it is not easy to see'. Translation of text on page 12 (right side): 'I cannot see it properly from here. The rocks are too high, do you understand? You are more able to see it, from above. No, it is not very far'.

mathematical development of the children. These reading guidelines also ensured that the reading used the books' full potential in creating a supportive learning environment. Furthermore, the keys also enforced that the book reading was carried out similarly across the participating classes.

Table 2 shows a classroom conversation that took place when these pages of the picture book *Ga je mee?* were read to a kindergarten class. Other classroom vignettes can be found in Van den Heuvel-Panhuizen, Van den Boogaard and Doig (2009).

PICO test

To investigate the effect of the picture book reading programme, we developed the so-called PICO test consisting of multiple choice items for the domains of number (with the topic of numbers and number relations), measurement (with the topic of length) and geometry (with the topic of perspective). A sample of the test items is included in the Appendix 1. Every item covers one page and contains an illustration depicting a situation and a number of illustrations that represent the possible answers. After the test item was read aloud to them, the children had to answer by underlining the correct answer or answers. For example, in the Mouse item, like the situation described in the picture book *Let's go* (see Figure 1), the students have to take an imaginary bird's-eye view to identify how the mouse looks from above.

Before the PICO test was used to collect the data in our study, the items were piloted, leading to a revision of some test items. The final test consisted of 42 items split up over two booklets to be administered on different days. Each booklet contained 21 items, which were equally distributed over the three mathematical domains.

Procedure

The field experiment started with administering the PICO test as a pretest in both the experimental and the control group. This was done by trained test administrators. The two test booklets were administered with an interval of a week. At the same time, we asked the teachers of both groups to fill in a questionnaire which provided us with information about each child's age, kindergarten year, gender, SES weight, home language and Cito mathematics and language scores. This questionnaire also contained questions about the teachers' teaching of mathematics, including questions about the mathematical topics addressed by the teacher, the kinds of activities done in class and the materials used for teaching mathematics. At the end of the experiment, this last part of the questionnaire was repeated.

The picture book reading intervention consisted of a three-month programme carried out in the experimental group by the teachers themselves. In addition to the intervention programme activities, the teachers followed their regular mathematics curriculum, provided that the total amount of time spent on mathematics and on picture book reading should not exceed the amount of time usually spent on these activities. Before the intervention started, the teachers received a training consisting of two three-hour sessions, which prepared them for the picture book programme, explaining the set-up of the reading sessions and how to use the reading guidelines. During the programme, the teachers read two picture books in class per week, in accordance with the prescribed schedule and reading guidelines. To verify whether

Table 2. Classroom conversation based on pages 11–12 of picture book Ga je mee? [Let's go] (Dematons, 2005).

	[The teacher turns the page.]
Child A:	'Well, I think it is rather far.'
Child B:	'A crocodile!'
Child A:	'Where?
Child B:	'There!'
	[All children search for the crocodile. Some children point at the crocodile. The teacher continues to read the text and asks child C to point out the sea, which he does.]
Child D:	'And there is a crocodile! And here he cannot pass.'
	[Child D points at the overhanging rock, where the boy is heading to.]
	[The teacher reads the remaining text.]
Teacher:	'Why can we see the sea better?'
Child D:	'Because the boy is looking that way (backwards). He is rowing and then you always have to look that way.'
Teacher:	'Can he look over the rocks?'
Children:	'No.'
Teacher:	'And we can see it, because we are above it, looking down. It is as if we are in an airplane above the book.'
Child E:	'Not!'
Teacher:	'Yes, we are.'
Child F:	'For them we are, for us we are not.'
Teacher:	'What else can we see better?'
Child G:	'The rocks.'
Teacher:	'He can see the rocks, but not all of them.'
Child A:	'The crocodile!'
Teacher:	'Indeed, the boy cannot see the crocodile, while we can.'
Child D:	'Perhaps the crocodile moves this way.'
	[Child D points out a route from the position of the crocodile to the sea.]
Child H:	'The boy cannot pass underneath here.'
	[Child H points at the overhanging rock.]
Child D:	'Yes he can, because that is too high!'
	[Child D also points at the overhanging rock.]
Teacher:	'Perhaps it is a kind of bridge.'
Teacher:	*[Asks child D.]*
	'Can you point out the route that the boy can take, towards the sea?'
	[Child D points out a route that uses the passage under the overhanging rock.]
Child D:	'He can go this way, if he can pass underneath this.'
	[Child D points at the overhanging rock.]
Child A:	'And he can also go this way.'
	[Child A points out an alternative safe route.]
Teacher:	'But he can also go this way.'
	[Teacher indicates the other route, crossing the crocodile's territory.]
Child D:	'But then he has to pass the crocodile.'
Child I:	'And then the crocodile will eat him.'

the picture book reading took place according to the instructions in the reading guidelines, we visited every teacher of the experimental group two times during the intervention to observe and video record two of the reading sessions. Furthermore, while carrying out the book reading programme, teachers kept logs to document how they read the books. Based on our observations and video recordings and the teachers' logs, we concluded that the book reading was done in agreement with the guidelines. Within a period of 2–4 weeks after the intervention, the PICO test was administered again in both the experimental and control group.

The teachers of the control classes were only afterwards informed that the study was aimed at investigating the effect of picture book reading on the mathematics performance of children. When the project started the teachers were just told that a test would be administered at two time points to measure their children's development in mathematics. We communicated to the teachers that the study was meant to gain information about how kindergarten's understanding of mathematics is evolving over a three-month period in normal school practice. Therefore, we asked the teachers to follow for all subjects including mathematics, drawing, playing, book reading and so on, their regular curriculum as if they were not involved in a research. The data from the questionnaires make clear that the teachers in the control groups did not use picture books for teaching children mathematics. In fact, this also applied to the teachers in the experimental group before they enrolled in the field experiment. Also with respect to other aspects of their regular teaching practice the teachers in both groups did not differ.

Data inspection

Psychometric properties PICO test

Based on a calculation of the item discrimination of the PICO test items, we removed two items which had negative item discriminations. This led to a test with 40 items in total that all have a positive correlation with the total score. A calculation of the Cronbach's alpha of this final PICO pretest resulted in a sufficient reliability of $\alpha = .79$ for the whole sample, and $\alpha = .71$ for the sample of K1 children as well as for the sample of K2 children. Furthermore, within the experimental and the control group, we found correlations between the PICO pretest and posttest score ranging from $r = .62$ to $r = .83$, indicating high test stability. We also obtained rather high correlations between the Cito mathematics test and the pretest scores of the K1 and K2 children ($r = .72$ and $r = .60$, respectively), which indicates a sufficient degree of concurrent validity.

To further investigate the properties of the items in the PICO test, we conducted a confirmatory factor analysis at the item level (using WLSMV estimation implemented in Mplus; Muthén & Muthén, 2012) with the three mathematical topics: number, growth and perspective as dimensions. This three-dimensional model resulted in a good fit (CFI = .97, TLI = .97, RMSEA = .02) with only one correlation significantly differing from 1. For number and perspective, the correlation was $r = .85$, which is substantially smaller than 1, whereas the other two correlations were higher than $r = .95$. This result indicated the dominance of one dimension. Coherent with these findings, a one-dimensional factor analysis resulted in an almost equally well-fitting model (CFI = .96, TLI = .97, RMSEA = .02). Based on these findings, we used the total score of the PICO test for the analyses in this study.

The descriptives of the PICO test total score in the pretest and the posttest are presented in Table 3. No indications of any floor or ceiling effect were found; none of the children reached the minimum or the maximum score at the pretest or the posttest. Moreover, the skewness values (pretest K1 = .11, K2 = −.07; posttest: K1 = .08, K2 = −.31) were close to zero, indicating that the pretest score distributions were approximately symmetric.

Table 3. Descriptives of experimental and control group for PICO pretest and posttest.

			Pretest score (total items: 40; max. score: 40)					Posttest score (total items: 40; max. score: 40)				
Kindergarten year	Group	*N*	*M*	*SE*	*SD*	*d*	*p*	*M*	*SE*	*SD*	*d*	*p*
K1	Exp	84	14.0	.54	5.5	.08	.59	18.2	.59	6.0	.25	.11
	Ctr	66	13.6	.61	4.0			16.9	.66	4.5		
K2	Exp	115	20.2	.47	4.8	.02	.92	24.6	.50	5.3	.19	.17
	Ctr	119	20.1	.46	5.1			23.6	.49	5.4		
K1 + K2	Exp	199	17.5	.42	6.0	−.03	.74	21.9	.44	6.4	.11	.26
	Ctr	185	17.7	.43	5.7			21.2	.46	6.0		
Total sample		384	17.6	.30	5.8			21.6	.32	6.2		

Missing data

The collected data contained a non-negligible amount of missing values. Of the 384 children participating in the study, only 308 were present during both the pretesting and the posttesting of the PICO test. Listwise deletion of the data of the children who had at least one missing value would diminish the sample size by about 20%. This might lead to biased estimates and would unnecessarily reduce the power of the statistical tests. Therefore, we applied a multiple imputation procedure (see, e.g. Graham, 2009).

We included all collected variables in the imputation model. Since the relations between the variables involved in this study could be different in the two kindergarten years, the imputation procedure was carried out for the K1 and K2 children separately. Because of the clustering of children within classes we also included two auxiliary variables at the class level (i.e. the class mean of the Cito mathematics test score and the sum of the SES weights of all children in class) in the imputation model. The MICE algorithm (Multivariate Imputation by Chained Equations; Van Buuren & Groothuis-Oudshoorn, 2011) was used to complete the data-set by running the imputation procedure 50 times resulting in 50 imputed data-sets. The analyses were done on each of the imputed data-sets and their results were combined using Rubin's rule (see Graham, 2009).

Statistical analysis

Analysis of the intervention effect and the differential intervention effects

To answer our first research question, we investigated the intervention effect using two linear regression models, namely, two One-Way ANCOVA models. In Model 1, we used the PICO posttest score as a dependent variable and as independent variables the experimental group (as a dummy variable) and the PICO pretest score (as a covariate). In Model 2, further independent variables were included, to increase the statistical power and to obtain an estimate of the intervention effect with a smaller bias. The covariates involved were kindergarten year, age, gender, home language, SES, Cito mathematics and Cito language.

To answer the second research question concerning differential intervention effects, we also used linear regression models, namely, Two-Way ANCOVA models. To test whether the intervention effect differs for particular child characteristics (e.g. age), in these models, we used the posttest score as the dependent variable and the pretest score, the particular child characteristic, the experimental group (dummy variable) and the interaction of the latter two as independent variables. Because some child characteristics depend on kindergarten year, we controlled for this latter child characteristic in the regression models. Furthermore, in case child characteristics were continuous (age, Cito mathematics and Cito language), they were dichotomised by a median split to get information about how the intervention works in the subgroups.

Consequences of clustered sampling

Although a clustered sample – children belonging to classes which belong to schools – generally produces larger standard errors compared to a random sample, in our study, these errors may have been reduced because we applied a procedure of clustered sampling that used matched pairs of schools which were randomly assigned to either the experimental group or the control group. Nevertheless, we could not make the assumption of random sampling of children for statistical inference. Therefore, it would be necessary to correct for a possible underestimation of standard errors through the application of multilevel analysis (MLA). However, despite the nested structure we preferred to apply single-level linear regression (SLA) models, because in our study, children were the units of inference. MLA takes into account the class structure for parameter estimation by which different results could be obtained than in SLA models. To justify that standard errors in SLA are not underestimated due to clustered sampling, we used a multilevel random intercept model in lme4 (Bates, Maechler, Bolker, & Walker, 2013) to calculate the residual intra-class correlation of the PICO posttest score controlling for the pretest score (as a covariate) and the experimental group (as a dummy variable). It turned out that the residual intra-class correlation was .025. This finding supported the conclusion that ignoring the multilevel structure in our analyses did not lead to notable underestimation of standard errors (Hox, 2010). Moreover, with only a few classes multilevel modelling will not lead to valid statistical inferences.

Results

Descriptives of experimental and control group

Table 3 shows the descriptives of all the test scores for the whole sample and specified for the experimental and control group, and for the two kindergarten years. These descriptives were based on the imputed data-sets.

For the PICO pretest score, no differences between experimental and control group were found (whole sample: $d = -.03$, $p = .74$; K1: $d = .08$, $p = .59$; K2: $d = .02$, $p = .92$). For the PICO posttest score, the experimental group scored slightly (but not significantly) higher than the control group (whole sample: $d = .11$, $p = .26$; K1: $d = .25$, $p = .11$; K2: $d = .19$, $p = .17$).

Overall intervention effect

Table 4 shows the results of the two regression models – i.e. One-Way ANCOVA models – we used for investigating the intervention effect on the PICO posttest score. Both models gave comparable results. Model 1 ‘ANCOVA with pretest’, in

Table 4. Results of regression analyses for investigating intervention effect on PICO posttest score.

	Model 1: ANCOVA with pretest					Model 2: ANCOVA with pretest and further covariates			
	B	*SE*	*p*	β		*B*	*SE*	*p*	β
Intervention	.90	.36	.01[a]	.07	Intervention	.76	.37	.02[a]	.06
PICO pretest	.89	.03	<.01[b]	.84	PICO pretest	.69	.05	<.01[b]	.64
					Kindergarten year (K2)[c]	.32	.67	.64	.02
					Age	.05	.04	.22	.06
					Gender (girl)	−.10	.35	.79	−.01
					Home language (Dutch)	1.22	.60	.04	.07
					SES (medium/high)	.20	.80	.80	.01
					Cito mathematics	.07	.02	<.01	.16
					Cito language	.02	.03	.44	.05
R^2 (Explained variance)	.703					.733			

B: unstandardised regression coefficient of the intervention effect; *SE*: standard error of *B*; β: standardised regression coefficient.
[a]Because of Hypothesis 1 the *B* value for the intervention effect was tested in a one-tailed way.
[b]The covariates were only included as control variables in order to get an unbiased estimation of the intervention effect (thus, we were not testing here the interaction of intervention and child characteristics); the *p* values were computed in a two-tailed way.
[c]For the categorical covariates, the dummy variables are placed in parentheses.

which we had only the PICO pretest score as a covariate, revealed a significant intervention effect ($B = .90$, $p = .01$), while Model 2 'ANCOVA with pretest and further covariates', in which we controlled for seven additional covariates, resulted in a similar intervention effect ($B = .76$, $p = .02$). In Model 1, we obtained an explained variance of the PICO posttest score of $R^2 = .703$ and in Model 2, the explained variance was $R^2 = .733$. In Model 2, pretest, home language and Cito mathematics did have a significant influence on the posttest score (for experimental and control group).

To investigate the effect size of the intervention effect, we calculated the effect size *d* for both models by dividing the *B*-values by the standard deviation of the PICO pretest scores. For Model 1 this means that we found an effect size $d = .16$ ($B = .90$ divided by 5.8), for Model 2, the effect size was $d = .13$ ($B = .76$ divided by 5.8).[1]

Comparing these effect sizes with the effect size of the change from pretest to posttest in the control group (gain score: $M = 3.5$, $SD = 3.5$, $d = .60$, $p < .01$) where no intervention took place, we found that the influence of the intervention was substantial. In Model 1, the change in the experimental group was 27% (.16/.60 = .27) larger than the change in the control group and in Model 2, the change was 22% (.13/.60 = .22) larger.

Intervention effects in and between subgroups

Table 5 contains the results of the regression analyses – i.e. One-Way ANCOVA models in the subgroups and Two-Way ANCOVA models in the total group – that were applied for testing the intervention effects for the child characteristics

Table 5. Results of regression analyses for investigating intervention effects on PICO posttest score in subgroups and for investigating differential intervention effects between subgroups.

Child characteristic	Intervention effects in subgroups					Differential intervention effects between subgroups					
	Subgroup	B^{a}	SE	p^{b}	d	Compared subgroups	B	SE	p^{b} one-tailed[c]	p^{b} two-tailed[d]	Δd^{e}
Kindergarten year	K1	1.05	.59	.04	.18	K2 vs K1	−.13	.73		.86	−.02
	K2	.92	.44	.02	.16						
Age	Young	1.00	.55	.03	.17	Old vs Young	−.09	.72		.90	−.01
	Old	.91	.46	.02	.16						
Gender	Boys	.48	.49	.16	.08	Girls vs Boys	1.21	.71	<.05		.21
	Girls	1.66	.51	<.01	.29						
Home language	Non-Dutch	1.37	.98	.08	.24	Dutch vs Non-Dutch	−.37	1.09		.76	−.07
	Dutch	1.00	.38	<.01	.17						
SES	Low	.77	1.04	.23	.13	SES Medium/high vs Low	.19	1.10		.86	.04
	Medium/high	.99	.38	<.01	.17						
Cito mathematics	Low	.99	.52	.03	.17	Cito mathematics High vs Low	−.08	.69		.91	−.02
	High	.89	.46	.03	.15						
Cito language	Low	.69	.61	.13	.12	Cito language High vs Low	.30	.80		.71	.06
	High	1.02	.50	.02	.18						

[a]Because of Hypothesis 1 the B values were tested in a one-tailed way.
[b]Because of multiple testing, we reduced the α level criterion from .05 to .01 (Bonferroni correction).
[c]Because we predicted a positive differential intervention effect, the p values were tested in a one-tailed way.
[d]Because we did not predict a differential intervention effect in a particular direction, the p value was tested in a two-tailed way.
[e]Δd is defined as the difference of the d values in the two subgroups.

kindergarten year, age, gender, home language, SES, Cito mathematics and Cito language. Because we had to carry out multiple tests, we reduced, according to the Bonferroni correction, the α criterion level from .05 to .01.

Using this adjusted significance criterion, we did neither find a significant intervention effect in the subgroup of K1 children nor in the K2 subgroup. Also, no significant differential intervention effect was found between these two kindergarten years. For the two age subgroups, we also did not find a significant intervention effect and the same applied for the differential intervention effect for age.

For the two gender subgroups, we found a significant intervention effect only for the girls ($B = 1.66$, $p < .01$, $d = .29$) and not for boys ($B = .48$, $p = .16$, $d = .08$). Taking into account the adjusted significance criterion, we found a marginally significant differential intervention effect for gender ($B = 1.21$, $p < .05$, $\Delta d = .21$). Moreover, the effect size found in the subgroup of girls ($d = .29$) was more than three times larger than in the boys subgroup ($d = .08$).

Regarding home language, we found a significant intervention effect for the subgroup of children with home language Dutch ($B = 1.00$, $p < .01$, $d = .17$) and not for the non-Dutch subgroup, although for this subgroup the effect was a bit larger ($B = 1.37$, $p = .08$, $d = .24$). Furthermore, no significant differential intervention effect was found for home language. For the children belonging to the medium/high-SES subgroup, we found a significant intervention effect ($B = .99$, $p < .01$, $d = .17$), but not for the low-SES subgroup ($B = .77$, $p = .23$, $d = .13$). Also no significant differential intervention effect was found for SES.

Within the subgroups resulting from the median split of the Cito mathematics test scores, there were neither significant effects for the children belonging to the higher scoring group nor for the children with the lower scores. Furthermore, no significant differential intervention effect was found for the mathematics ability. Within the two subgroups that resulted from the median split of Cito language test scores, we found no significant intervention effects. There was also no significant differential intervention effect.

Discussion

Our study showed that a three-month picture book reading programme, during which the teacher read two picture books that contain mathematics-related content in class each week, had a positive effect on kindergartners' mathematics performance as measured by the PICO test (see Table 4). Although the measured effect size ($d = .16$) of the intervention in the experimental group in Model 1 can be considered as rather small, the increase from pretest to posttest is 27% larger than in the control group. This result is worth speaking of. The same is true for Model 2, where the change was 22% larger. In fact, this gain from a three-month programme is quite a lot taking into account the spurt in cognitive growth children generally make at this age, which is clearly shown by the increase in performance of the children in the control group, and which is also emphasised by other authors (e.g. Bowman, Donovan, & Burns, 2000). Given this spurt, our results indicate a non-negligible effect of the picture book programme. Moreover, a multilevel analysis based on the PICO posttest score controlling for the pretest score and the experimental group led to a very small residual intra-class correlation, indicating that no class-specific intervention effects occurred. Altogether, our findings support Hypothesis 1 and are in line with previous studies (Casey, Erkut, et al., 2008; Hong, 1996; Jennings et al., 1992; Young-Loveridge, 2004) on the positive effect of using picture books on children's mathematics performance. Moreover, in our study, the positive results were found based on picture book reading without additional mathematical activities.

For Hypothesis 2, in which we stated that we expected a significant differential intervention effect for gender, we only found some weak support (see Table 5). Based on the adjusted significance criterion the effect was only marginally significant. Yet, it was nevertheless of a notable size. Moreover, testing the intervention effects in the two gender subgroups revealed, in line with the results from the study of Casey, Erkut, et al. (2008), a significant effect for girls, but not for boys, with an effect size for girls that was more than three times larger than for boys. Based on the above findings, we may conclude that reading picture books can contribute to giving girls a better start in developing mathematical understanding than they, according to several researchers (e.g. Carr & Davis, 2001; Penner & Paret, 2008), presently have on entering first grade.

With respect to the other child characteristics, including kindergarten year, age, home language, SES, mathematics and language ability, we only found significant intervention effects in the subgroup of children with medium/high SES and the subgroup of children with home language Dutch. However, the main finding of our study was that for all these child characteristics there were no significant differential intervention effects (see Table 5).

In sum, we can conclude that our study provided evidence for giving picture book reading a significant place in the kindergarten curriculum for supporting children's mathematical development. Such a picture book reading programme seems to be effective as a whole-class activity in heterogeneously composed kindergarten classes, including children of different kindergarten years, ages, socio-economic background, language abilities and previous knowledge in mathematics. Moreover, it seems to be gainful for girls.

However, the above-mentioned conclusions and recommendations should be considered in the perspective of our study's limitations. A first shortcoming involves our sample of schools. Although we made equivalent school pairs with respect to urbanisation level, school size and SES, and we assigned the schools randomly to either the experimental group or the control group, the sampling as a whole was not completely at random because participation from schools and teachers was on a voluntary basis. As a result of this, it is possible that only motivated teachers were involved in the study. However, because field experiments in educational practice cannot be carried out without the willingness of teachers, this is a flaw in the design that can hardly be avoided.

A further shortcoming is also related to doing the research in educational practice. Although we instructed the teachers about how to carry out the intervention (by means of the reading guidelines and two three-hour sessions) and we checked the intervention by visiting every teacher two times and asking them to make logs, we are not completely sure about the intervention fidelity. A possible way to overcome this weakness would be to carry out a more controlled experiment in which the intervention is delivered by research assistants. By using trained research assistants, it would be possible to ensure that the intervention is in all classes the same. Moreover, in such an approach, the research assistant could have given the intervention by taken children out of the classroom, which would have enabled us to have a sample with control and experimental children in the same classroom. However, as our study intended to provide knowledge about whether picture book reading would have an effect on children's performance when the intervention was carried out in a real-classroom setting, a more controlled experiment was not suitable in our case.

Another restraint of our study is that the sample size was in fact not large enough to optimally investigate the intervention effect within the subgroups based on child characteristics. When splitting the sample into these subgroups this resulted in a considerable reduction of the statistical power. Getting more robust findings for these subgroups requires further research in which a larger number of classes are involved.

Another issue that also needs additional research is related to the picture book reading sessions themselves. Although our findings showed that the reading sessions had an effect on the children's mathematics performance and our analyses made it reasonable to assume that the intervention was conducted similarly in all experimental classes (the variance at the class level was very small), our research design did not include opportunities to identify the effective elements of the picture book reading sessions. To make this identification possible, future research is recommended to monitor classroom interactions in detail. Only then can the full benefit of the potential of picture books for teaching mathematics to young children be revealed.

Acknowledgements

The study reported in this article was carried out in the PICO-ma project that was supported by a programme grant from the Netherlands Organization for Scientific Research (NWO-MaGW/PROO: Project 411-04-072). We are grateful to the teachers who participated in the study. A special thanks goes to Sylvia van den Boogaard who was heavily involved in setting up the study and who worked with the teachers and collected the data.

Note

1. In order to disentangle the explained variance of covariates and intervention we applied a stepwise regression analysis for both Model 1 and Model 2. For Model 1 (including pretest and intervention as covariates), the first step with only pretest included as a covariate led to $R^2 = .698$. This means that the difference with $R^2 = .703$ of Model 1 resulted in an effect size of $\Delta R^2 = .005$. For Model 2, this effect size was $\Delta R^2 = .004$.

References

Aiken, L. R. (1972). Language factors in learning mathematics. *Review of Educational Research, 42*, 359–385.

Anders, Y., Rossbach, H.-G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S., & von Maurice, J. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly, 27*, 231–244.

Anderson, A., Anderson, J., & Shapiro, J. (2005). Supporting multiple literacies: Parents' and children's mathematical talk within storybook reading. *Mathematics Education Research Journal, 16*, 5–26.

Arizpe, E., & Styles, M. (2003). *Children reading pictures: Interpreting visual texts*. London: Routledge/Falmer.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-5*. Retrieved from http://CRAN.R-project.org/package=lme4

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98*, 44–62.

Blok, H. (1999). Reading to young children in educational settings: A meta-analysis of recent research. *Language Learning, 49*, 343–371.
Bowman, B., Donovan, M. S., & Burns, S. (Eds.). (2000). *Eager to learn. Educating our preschoolers*. Washington, DC: National Academy Press.
Carr, M., & Davis, H. (2001). Gender differences in arithmetic strategy use: A function of skill and preference. *Contemporary Educational Psychology, 26*, 330–347.
Casey, B., Erkut, S., Ceder, I., & Mercer Young, J. (2008). Use of a storytelling context to improve girls' and boys' geometry skills in kindergarten. *Journal of Applied Developmental Psychology, 29*, 29–48.
Casey, B. M., Andrews, N., Schindler, H., Kersh, J. E., Samper, A., & Copley, J. (2008). The development of spatial skills through interventions involving block building activities. *Cognition and Instruction, 26*, 269–309.
Clarkson, P. C. (1992). Language and mathematics: A comparison of bilingual and monolingual students of mathematics. *Educational Studies in Mathematics, 23*, 417–429.
Collins, M. F. (2010). ELL preschoolers' English vocabulary acquisition from storybook reading. *Early Childhood Research Quarterly, 25*, 84–97.
Columba, L., Kim, C. Y., & Moe, A. J. (2005). *The power of picture books in teaching math & science. Grades PreK-8*. Scottsdale, AZ: Holcomb Hathaway.
Dematons, Ch. (2005). *Ga je mee?* [Let's go]. Rotterdam: Lemniscaat.
Egan, K. (1985). Teaching as story-telling: A non-mechanistic approach to planning teaching. *Journal of Curriculum Studies, 17*, 39–406.
Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Dordrecht: Reidel Publishing Company.
Ginsburg, H. P., Chia-Ling, L., Ness, D., & Seo, K. H. (2003). Young American and Chinese children's everyday mathematical activity. *Mathematical Thinking and Learning, 5*, 235–258.
Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549–576.
Gullo, D. F., & Burton, C. B. (1992). Age of entry, preschool experience, and sex as antecedents of academic readiness in kindergarten. *Early Childhood Research Quarterly, 7*, 175–186.
Haury, D. (2001). *Literature-based mathematics in elementary school, ERIC Digest*. Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education (ED464807).
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.
Hong, H. (1996). Effects of mathematics learning through children's literature on math achievement and dispositional outcomes. *Early Childhood Research Quarterly, 11*, 477–494.
Horst, J. S., Parsons, K. L., & Bryan, N. M. (2011). Get the story straight: Contextual repetition promotes word learning from storybooks. *Frontiers in Psychology, 2*, 1–11. Article 17. doi:10.3389/fpsyg.2011.00017
Hox, J. (2010). *Multilevel analysis: Techniques and applications*. New York, NY: Routledge.
Jennings, C. M., Jennings, J. E., Richey, J., & Dixon-Krauss, L. (1992). Increasing interest and achievement in mathematics through children's literature. *Early Childhood Research Quarterly, 7*, 263–276.
Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research & Practice, 22*, 36–46.
Jordan, N. C., Kaplan, D., Nabors Oláh, L., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development, 77*, 153–175.
Klein, P. S., Adi-Japha, E., & Hakak-Benizri, S. (2010). Mathematical thinking of kindergarten boys and girls: Similar achievement, different contributing processes. *Educational Studies in Mathematics, 73*, 233–246.

Lovitt, C., & Clarke, D. (1992). *The mathematics curriculum and teaching program (MCTP): Professional development package activity bank*, (Vol. 2). Carlton: Curriculum Cooperation.

Millard, E. (1997). *Differently literate: Boys, girls, and the schooling of literacy*. London: The Falmer Press.

Mol, S. E., Bus, A. G., & de Jong, M. T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research, 79*, 979–1007.

Mol, S. E., Bus, A. G., de Jong, M. T., & Smeets, D. J. H. (2008). Added value of dialogic mother–child book reading: A meta-analysis. *Early Education and Development, 19*, 7–26.

Moyer, P. S. (2000). Communicating mathematically: Children's literature as a natural connection. *Reading Teacher, 54*, 246–255.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

National Association for the Education of Young Children and National Council of Teachers of Mathematics. (2002). Position statement. *Early childhood mathematics: Promoting good beginnings*. Retrieved December 4, 2011, from http://www.naeyc.org/files/naeyc/file/positions/psmath.pdf

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research, 37*, 239–253.

Saracho, O., & Spodek, B. (2009). Educating the young mathematician: A historical perspective through the nineteenth century. *Early Childhood Education Journal, 36*, 297–303.

Sarama, J., & Clements, D. H. (2009a). *Early childhood mathematics education research. Learning trajectories for young children*. New York, NY: Routledge.

Sarama, J., & Clements, D. H. (2009b). Building Blocks and cognitive Building Blocks. Playing to know the world mathematically. *American Journal of Play, 1*, 313–337.

Schickedanz, J. (1995). Early education and care: Beginnings. *Journal of Education, 177*(3), 1–7.

Seo, K.-H., & Ginsburg, H. P. (2004). What is developmentally appropriate in early childhood mathematics education? Lessons from new research. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics* (pp. 91–104). Mahwah, NJ: Lawrence Erlbaum Associates.

Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly, 19*, 99–120.

Stipek, D., & Byler, P. (2001). Academic achievement and social behaviors associated with age of entry into kindergarten. *Journal of Applied Developmental Psychology, 22*, 175–189.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). MICE: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), 1–67.

Van den Heuvel-Panhuizen, M., & Elia, I. (2012). Developing a framework for the evaluation of picturebooks that support kindergartners' learning of mathematics. *Research in Mathematics Education, 14*, 17–47.

Van den Heuvel-Panhuizen, M., & Van den Boogaard, S. (2008). Picture books as an impetus for kindergartners' mathematical thinking. *Mathematical Thinking and Learning, 10*, 1–33.

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*, 30–39.

Van Kuyk, J., & Kamphuis, F. (2001). *Verantwoording van de toetsen uit de pakketten Ruimte en Tijd, Taal voor Kleuters en Ordenen* [Account for the tests Space and Time, Language for Kindergartners, and Ordering]. Arnhem: Citogroep.

Van Oers, B. (1996). Learning mathematics as a meaningful activity. In P. Nesher, L. P. Steffe, P. Cobb, G. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 91–113). Mahwah, NJ: Lawrence Erlbaum.

Wang, A. H. (2010). Optimizing early mathematics experiences for children from low-income families: A study on opportunity to learn mathematics. *Early Childhood Education Journal, 37*, 295–302.

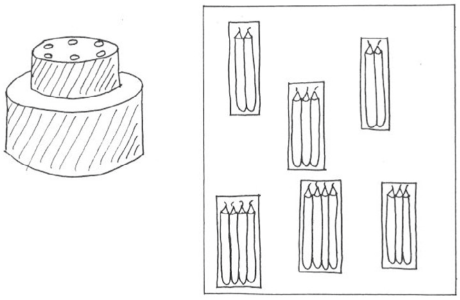
Whitin, D. J., & Wilde, S. (1992). *Read any good math lately? Children's books for mathematics learning, K-6*. Portsmouth, NH: Heinemann.

Wigfield, A., & Guthrie, J. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology, 89*, 420–432.

Young-Loveridge, J. M. (2004). Effects on early numeracy of a program using number books and games. *Early Childhood Research Quarterly, 19*, 82–98.

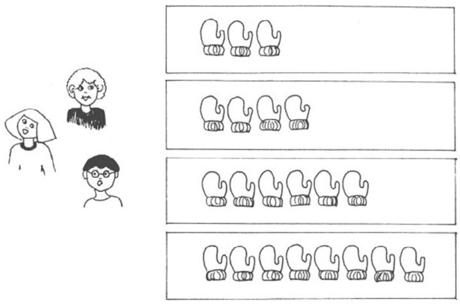
Appendix 1. Sample of PICO test items

Number: Cake



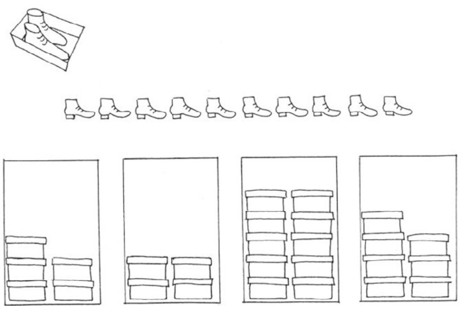
Buy exactly 6 candles to put on the cake. Put a line under the boxes that you take

Number: Mittens



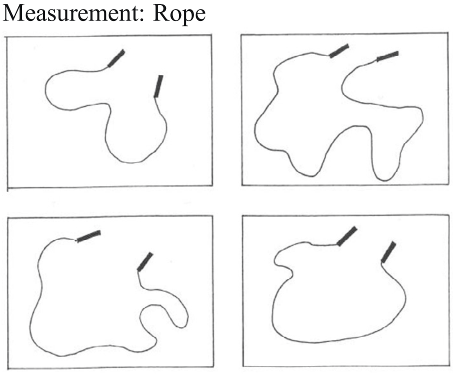
Underline the amount of mittens these children need in total

Number: Shoe boxes



Two shoes fit into one box. Underline the number of boxes you need for the other shoes

Measurement: Rope

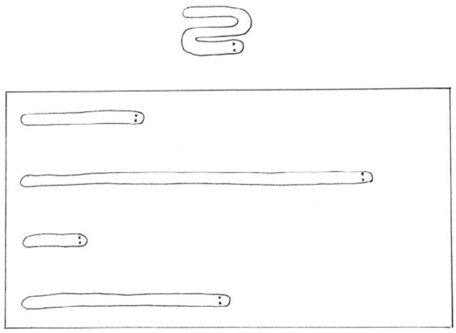


Which skipping rope is the longest? Put a line under the longest rope

(*Continued*)

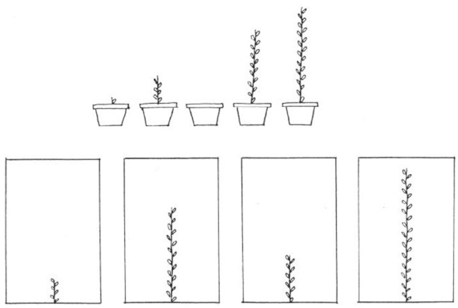
Appendix 1. (*Continued*)

Measurement: Snake



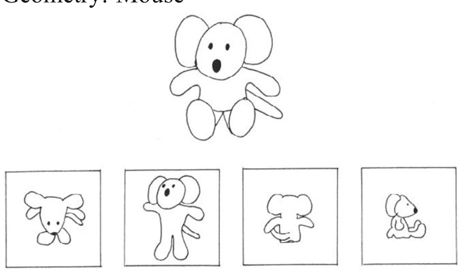
Which snake is as long as the writhing snake? Put a line under the snake that is just as long

Measurement: Plant



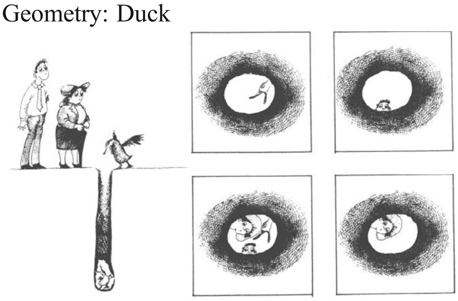
The plant gets taller and taller. Put a line under the plant that belongs in the empty flowerpot

Geometry: Mouse



There is Mouse. How would Mouse look if you looked down on him like a bird? Underline the way Mouse looks from above

Geometry: Duck



In the hole there is a little duck. What does the little duck see when it looks up? Put a line under what the little duck sees

(*Continued*)

Appendix 1. (*Continued*)

Geometry: Soccer

Two children are playing soccer. Put a line under what you see if you look at them from above like a little bird