Early Childhood Research Quarterly 44 (2018) 231–241

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

Effectiveness of the Building Blocks program for enhancing Ecuadorian kindergartners' numerical competencies

Gina Bojorque[a,b,*], Joke Torbeyns[b], Jo Van Hoof[b], Daniël Van Nijlen[b], Lieven Verschaffel[b]

[a] Universidad de Cuenca, Facultad de Filosofía, Av. 12 de Abril, Cuenca, Ecuador
[b] KU Leuven, Center for Instructional Psychology and Technology, Dekenstraat 2, Post box 3773, B-3000 Leuven, Belgium

ARTICLE INFO

Article history:
Received 9 March 2017
Received in revised form
28 November 2017
Accepted 8 December 2017

Keywords:
Building blocks program
Early numerical abilities
SFON
Intervention
Kindergarten

ABSTRACT

We investigated the effectiveness of the Building Blocks program for enhancing Ecuadorian kindergartners' early numerical abilities and spontaneous focusing on numerosity (SFON), after controlling for working memory, intelligence, age, and SES. Following a pretest-intervention-posttest design, 18 classes comprising 355 children from varied SES backgrounds were randomly assigned to either an experimental (BB program) or a control (regular mathematics program) condition. Results showed that the children from the experimental group made more progress in their early numerical competencies than those from the control group. Furthermore, the BB program was associated with higher quality mathematics education. We discuss the theoretical and educational implications for early numeracy development in general, and for the Ecuadorian situation in particular.

1. Introduction

Early numerical competencies are of major importance for children's further numerical and mathematical development (Aunio & Niemivirta, 2010; De Smedt, Verschaffel, & Ghesquière, 2009; Hannula-Sormunen, Lehtinen, & Räsänen, 2015; Hannula, Lepola, & Lehtinen, 2010; Hannula, Rasanen, & Lehtinen, 2007; Jordan, Kaplan, Ramineni, & Locuniak, 2009). It is increasingly emphasized that these competencies include both children's early numerical abilities (e.g., their ability to count, to compare numerical magnitudes or to decompose numbers) and their numerical dispositions (e.g., their spontaneous inclination to focus on and make sense of the numerical magnitudes in the situation) (Bojorque, Torbeyns, Hannula-Sormunen, Van Nijlen, & Verschaffel, 2017; Mulligan et al., n.d.). Children's early numerical abilities (Aunio & Niemivirta, 2010; De Smedt et al., 2009; Jordan et al., 2009) as well as their early numerical dispositions — more specifically, their spontaneous focus on numerosities or SFON, defined as children's natural tendency to spontaneously focus attention on the aspect of the exact number of items or incidents when exact numerosity is utilized in action (Hannula-Sormunen et al., 2015; Hannula & Lehtinen, 2005; Hannula et al., 2010, 2007) — were shown to contribute to children's later mathematical performance at school.

1.1. Contribution of background and domain-general cognitive characteristics

Children's early numerical abilities have been shown to be moderated by several domain-general cognitive characteristics and background characteristics. With respect to the domain-general cognitive characteristics, working memory and intelligence have shown to play a central role in the acquisition of early numerical abilities (Bull, Espy, & Wiebe, 2008; Friso-van den Bos, van der Ven, Kroesbergen, & Van Luit, 2013; Geary, Hoard, & Nugent, 2012; Passolunghi, Lanfranchi, Altoè, & Sollazzo, 2015; Swanson, Jerman, & Zheng, 2008). As reported in the review of Friso-van den Bos et al. (2013), children's working memory capacity is closely related to their mathematics performance and predicts proficiency in mathematics achievement at seven years of age. In the same proposition, Passolunghi et al. (2015) found that kindergarteners' verbal intelligence was directly associated with their early numerical abilities.

* Corresponding author at: Universidad de Cuenca, Facultad de Filosofía, Av. 12 de Abril, Cuenca, Ecuador.
E-mail addresses: gina.bojorque@ucuenca.edu.ec (G. Bojorque), joke.torbeyns@kuleuven.be (J. Torbeyns), jo.vanhoof@kuleuven.be (J. Van Hoof), daniel.vannijlen@kuleuven.be (D. Van Nijlen), lieven.verschaffel@kuleuven.be (L. Verschaffel).

https://doi.org/10.1016/j.ecresq.2017.12.009
0885-2006/

Furthermore, intelligence and working memory measured at Grade 1, 2, and 3 were shown to contribute to children's accuracy in mathematical word problem solving two years later (Swanson et al., 2008). Similarly, Geary et al. (2012) found that intelligence and working memory assessed in kindergarten were associated with the complexity and accuracy of children's addition strategies at the beginning of first grade. Thus, stronger intelligence and working memory capacity measured in kindergarten are associated with more sophisticated arithmetic and mathematical abilities. To the best of our knowledge, evidence on the relation between children's SFON and their intelligence or working memory capacity is currently missing.

With respect to background characteristics, age and SES play an important role in the acquisition of children's early numerical abilities. It was observed that children who started kindergarten at an older age had an advantage over younger children in early numerical abilities (Jordan et al., 2009). Similarly, previous studies reported that young children from disadvantaged SES backgrounds, on average, have weaker mathematical knowledge and skills than their peers from middle SES backgrounds (Clements & Sarama, 2011a; Jordan et al., 2009; Siegler & Ramani, 2008; Starkey, Klein, & Wakeley, 2004). These early difficulties in the acquisition of early numerical abilities are of great concern, as these children are at risk to continue to perform weakly during formal schooling (Jordan et al., 2009). Although prior SFON studies have included children from various SES backgrounds (e.g., Hannula et al., 2010), they did not systematically investigate the relation between SES and SFON.

1.2. Early mathematics programs

Given the difficulties of children from disadvantaged SES backgrounds, several mathematics programs have been developed to stimulate the development of the early numerical abilities of (especially these) children, including Number Worlds (Griffin, 2007), Pre-K Mathematics (Klein, Starkey, & Ramirez, 2002), Big Math for Little Kids (Ginsburg, Greenes & Balfanz, 2003), and Building Blocks (Clements & Sarama, 2013). Intervention studies evidence the effectiveness of these programs for enhancing the early numerical abilities of 4- and 5-year-old children from low SES backgrounds (Clements, Sarama, Spitler, Lange, & Wolfe, 2011; Griffin, 2005; Lewis Presser, Clements, Ginsburg, & Ertle, 2015). Additionally, Ramani, Siegler, and Hitti (2012) demonstrated that playing linear board games (e.g., The Great Race) in small groups during 20–25 min for 3–4 weeks improved early numerical abilities of 3- to 5-year-old children from low SES backgrounds from Head Start classrooms. Finally, it is worth noting that all the intervention studies reported here have been performed, so far, only in developed countries where the primary language is English.

In this study, we will focus on one of these programs, namely the Building Blocks (BB) program,[1] for the following reasons. First, it was developed based on a comprehensive Curriculum Research Framework and structured in research-based learning trajectories (Clements, 2007). Second, it also includes an appropriate professional development program for teachers, emphasizing teaching for understanding via these learning trajectories. By using an observational instrument that measures the quality of the mathematics environment and activities, namely the Classroom Observation of Early Mathematics Environment and Teaching (COEMET; Sarama & Clements, 2009b), Clements et al. reported that the professional development program helped teachers to increase the quality of their mathematics classroom environment and teaching practices (Clements & Sarama, 2008; Clements et al., 2011). Third, empirical evidence supports the effectiveness of the BB program for enhancing young children's early numerical abilities. In this respect, previous studies demonstrated that prekindergarten children who received the BB program outperformed their peers not involved in this program in general mathematics achievement and early numeracy tasks (Clements & Sarama, 2007, 2008; Clements et al., 2011). These positive effects of the BB program on children's mathematics achievement persisted in kindergarten (Sarama, Clements, Wolfe, & Spitler, 2012) and first grade (Clements, Sarama, Wolfe, & Spitler, 2013). Furthermore, children in the BB classes performed better than the control children in oral language subtests (Sarama, Lange et al., 2012). Fourth, and finally, as a complement of the BB program, the same research team created the Fidelity of Implementation (Fidelity) instrument (Sarama & Clements, 2012), a measure of implementation fidelity that evaluates the degree to which teachers are accurately teaching the BB program.

1.3. The Building Blocks program

The BB program's basic approach is ". . . to find the mathematics in, and develop mathematics from children's experiences and interests" (Clements & Sarama, 2013, p. T13). Its activities are based on the developmental levels of mathematics learning trajectories and are carefully designed and sequenced to address each level of the learning trajectory. Learning trajectories refer to children's natural developmental progression in learning mathematics. They include three important elements: (1) a mathematical goal, defined as an aspect of a mathematical domain that children should learn; (2) a developmental path, describing the development of children's levels of thinking to reach that mathematical goal; and (3) a set of instructional activities, indicating how to help children move along that developmental path (Clements & Sarama, 2004).

The program addresses five mathematical areas: (1) number, (2) geometry, (3) measurement, (4) patterns and early algebra, and (5) classifying and analyzing data. In this study we focused on the area of number. This area includes (1) counting, (2) comparing and ordering, (3) recognizing numbers and subitizing, (4) composing numbers, (5) adding and subtracting, and (6) numerals. The program consists of daily lessons, in which children are guided to explore, represent and discuss mathematics through activities and games in the whole group, in small groups, in free-choice learning centers, and during reflection time. Important components of the program are the use of technology, permanent assessments, family involvement, and the inclusion of so-called "mathematics throughout the year" activities (i.e., activities that help to integrate mathematics into daily classroom practices). An example of the latter activities is called *I see numbers* where teachers try to help children see groups of one, two, and three everywhere and in every opportunity they have along the day, such as three Trees – not just a group of Trees – "helping them form the habit of quantifying small collections" (Clements & Sarama, 2013, p. 3), a concept closely related to children's SFON as acknowledged explicitly by the authors (Sarama & Clements, 2009a).

1.4. The present study

Although important, all previously mentioned intervention studies (Clements et al., 2011; Griffin, 2005; Lewis Presser et al., 2015; Ramani et al., 2012) have some limitations. Firstly, they focused on systematically evaluating children's early numerical abilities, leaving aside their dispositions to attend to and make sense of numerical magnitudes, including SFON. As mentioned above, SFON has shown to play a pivotal role in the development of young children's early numerical competencies and has demon-

[1] As explained in Section 2, we were only able to implement the core BB program but not the software activities. Moreover, in our measurements we only looked at its effectiveness for children's early numerical competencies and, thus, not at their broader mathematical development.

strated to have predictive power in explaining children's later mathematical achievement at school. Secondly, although these intervention studies included children from low SES backgrounds, they have been carried out in developed countries (mainly in the US), which differ from developing countries in terms of general cultural, societal, and educational characteristics. Therefore, the effectiveness of the BB program in less developed countries, such as Ecuador, remains an open question. Thirdly, previous intervention studies have not controlled for cognitive variables that are well known to influence children's early numerical abilities, especially intelligence and working memory (Geary et al., 2012; Swanson et al., 2008). Finally, many prior investigations have been conducted only by the same research team that designed the program, which, according to some authors, might jeopardize the validity of the findings on the program's efficacy (Putnam, 2003; Schoenfeld, 2007). Therefore, we aimed at evaluating the effectiveness of the BB program for enhancing young children's early numerical competencies – including both early numerical abilities and SFON – in Ecuador, a developing country (United Nations, 2016). We included 5–6 year olds from various SES backgrounds and controlled for children's intelligence and working memory.

We tested two hypotheses in this study. As mentioned above, several studies suggest that children who follow the BB program outperform children who do not follow the BB program in early numeracy tasks (Clements & Sarama, 2007, 2008; Clements et al., 2011). Consequently, our first hypothesis was children who follow the BB program will make more progress in their acquisition of early numerical abilities than children from the control group, as indicated by the differences in their gain on early numerical abilities tests at the end of the school year (Hypothesis 1).

The fact that the BB program includes activities that try to help children "see small groups of objects everywhere" (see above) may encourage children to see the amount of something as opposed to only seeing the 'something' (Clements & Sarama, 2013). Based on the study of Hannula, Mattinen, and Lehtinen (2005), in which the personnel of a day care center intentionally directed 3-year-olds' attention towards small numbers of items, resulting in children's enhancement of their initial SFON tendency, we expected that these kinds of activities from the BB program (see Sarama & Clements, 2009a) may also have an influence on children's SFON development. Furthermore, Hannula and Lehtinen (2005) reported a reciprocal relationship between SFON and early numerical abilities, suggesting that the development of SFON promotes the development of numerical abilities, and vice versa. Accordingly, we expected that the different numerical activities proposed in the BB program would not only enhance children's early numerical abilities but also their SFON. Therefore, our second hypothesis was children who follow the BB program will make more progress in their SFON than children from the control group, as indicated by the differences in their gain on SFON tasks at the end of the school year (Hypothesis 2).

In addition to our two hypotheses, we also formulated one research question for which specific hypotheses could not be raised. Prior intervention programs have proved to be an effective way to promote teachers' professional development, and thus improve the quality of mathematics education (Clements & Sarama, 2008; Clements et al., 2011; Griffin, 2004, 2005). Accordingly, our research question was phrased as follows: Do teachers who follow the BB program offer higher quality mathematics education than the teachers from the control group, as indicated by their scores on the COEMET (Research question 1)? This question was not phrased in terms of a hypothesis because the rather low number of teachers involved in the experimental and control condition did not allow a proper statistical test of such a hypothesis. By also exploring this question in the present study, we hoped that this could act as a starting point for future research to examine the impact of the BB program on children's learning outcomes, with the quality of teachers' mathematics education as a mediator.

2. Method

2.1. Overall design

This study followed a pretest-intervention-posttest design. Before the beginning of the study, a cluster randomized controlled trial design was utilized in which 18 kindergarten schools were randomly assigned to either an experimental or control condition. Half of the schools implemented the BB mathematics program (Clements & Sarama, 2013; experimental group) for 30 weeks during the school year (i.e., from October until May), whereas the other half followed the regular mathematics program (control group). The school year lasts a total of 40 weeks, not including four weeks of holidays. Teachers in both groups maintained their typical schedule including 40 min of mathematics classes every day. The teachers in the experimental group implemented the BB activities during these 40 min of mathematics classes. In line with the program designers (e.g., Sarama & Clements, 2009a), the teachers in the experimental group followed a slightly adapted version of the BB professional development program before (i.e., August) and during (i.e., November and February) the intervention. The adaptation of the program consisted in replacing some songs and rhymes that did not have a Spanish version or were difficult to translate in Spanish by available songs and rhymes with similar content (e.g., counting from one to ten) from the Ecuadorian culture. With respect to the manipulatives proposed by the program, some of them were exactly reproduced from the originals while others were created by the teachers (e.g., foamy pizzas), or adapted using the materials available in the classrooms (e.g. farm animals).

2.2. Participants

Participants were 355 Ecuadorian 5–6 year-olds (182 boys). To maximize the representativeness of our sample, we recruited children from the three major school types in Ecuador (i.e., public urban, public rural, and private), six schools per type, one class per school, about 20 children per class. At the beginning of the study, the mean age of the children was 5 years 2 months (*SD* = 3.7 months). Parents' informed consent forms were collected from all participating children. From this original sample, two children were lost because they changed schools.

Children's SES was calculated via the mothers' educational level (e.g., Aunio & Niemivirta, 2010; Starkey et al., 2004). The level of maternal education was organized into nine categories: (1) no education; (2) pre-primary education; (3) primary education; (4) lower secondary education; (5) upper secondary education; (6) lowest level tertiary education; (7) lower-degree level tertiary education; (8) higher-degree level tertiary education; and (9) doctorate level degree. These nine categories were afterwards reorganized into three categories: (1) low SES level, when the highest level of the mother's education was primary education (i.e., former categories 1–3); (2) middle SES level, when the highest level of the mother's education was secondary education (i.e., former categories 4 and 5); and (3) high SES level, when the highest level of the mother's education was higher education (i.e., former categories 6–8). There were no mothers within the ninth category (i.e., doctorate level degree).

By using a blocked randomized design, in which randomization occurred within each type of school setting, schools were randomly assigned to one of the two conditions, resulting in nine schools (i.e., three public urban, three public rural, and three private) belonging to the experimental group and nine schools (i.e., three public

Table 1
Number of children, mean, age, and SES in experimental and control condition.

Condition	Children			Mean age (SD)	Mean SES (SD)	Number of children per SES category		
	Boys	Girls	Total			Low	Middle	High
Experimental	91	86	177	5 y 2 m (3.8 m)	4.92 (1.70)	53	53	71
Control	91	87	178	5 y 2 m (3.7 m)	4.80 (1.65)	50	70	58

urban, three public rural, and three private) to the control group. *T*-tests confirm that the experimental and the control condition were similar with respect to children's gender, $t(353) = -.27$, $p = 0.79$, age, $t(353) = .53$, $p = 0.59$, and SES (based on the nine categories), $t(353) = .67$, $p = 0.51$. Table 1 presents the descriptive statistics of the sample.

A total of 18 teachers participated in the study. All these teachers had education-related degrees ranging from technical degrees in childhood education, to bachelor degrees in educational psychology or primary education, up to master degrees in early childhood education. The mean years of service were set at 22 years.

The educational system in Ecuador is organized in three levels: (1) beginning level, for children aged 3–5 years; (2) basic education, for children aged 5–14 years; and (3) high school, for students aged 15–17 years. The first year of basic education, for children aged 5–6 years, corresponds to kindergarten. The Ministry of Education issues a mandatory national curriculum in both public (urban and rural) and private sectors. During the kindergarten year, children spend five days per week at school, from 7:30 in the morning till 12:30 in the afternoon.

2.3. Materials

To analyze children's early numerical competencies development, they were offered a test battery focusing on early numerical abilities and SFON at both the start (i.e., September) and the end (i.e., May/June) of the school year. Children's intelligence and working memory were also assessed at the start of the school year. We observed twice (i.e., January and May) the quality of mathematics education children received in all participating classes, using the COEMET (Sarama & Clements, 2009b). Finally, we controlled for the fidelity of implementation of the BB program by the teachers, from the experimental group, three times (i.e., January, March, May) using the Fidelity instrument (Sarama & Clements, 2012). These materials are described in the following paragraphs.

2.3.1. Early numerical abilities

Children's early numerical abilities were measured using two different instruments: the Test of Early Number and Arithmetic (TENA; Bojorque, Torbeyns, Moscoso, Van Nijlen, & Verschaffel, 2015) and the Tools for Early Assessment in Math (TEAM; Clements & Sarama, 2011b). The TENA is a reliable and valid instrument based on the Ecuadorian national standards for kindergarten number and arithmetic (Bojorque et al., 2015). The test consists of 54 items divided over nine subscales (with 6 items per subscale), namely (1) quantifiers, (2) one-to-one correspondence, (3) order relations more than/less than, (4) counting, (5) quantity identification and association with numerals, (6) ordering, (7) reading and writing numerals, (8) addition, and (9) subtraction. The administration of the TENA involves an individual as well as a collective part. The individual part consists of an individual interview with each child in a separate room outside his/her classroom. This part has 29 items that require the child to respond in a physical and/or oral way. The collective part consists of a paper-and-pencil test comprising 25 items that are administered collectively to the whole class and that require a written response from the children. The maximum score on the test is 54 (see Bojorque et al., 2015, for a more detailed description). Cronbach's alpha for the TENA scores for the sample from the present study was .89. Next, the TEAM is an international reliable and valid test that evaluates children's mathematical knowledge and skill (Clements & Sarama, 2011b). The TEAM is organized in two parts, number (Part A) and geometry (Part B). For the purpose of this study we administered the Spanish version of Part A. Part A consists of 93 items that measure (1) recognition of numbers and subitizing, (2) verbal and object counting, (3) number comparison and number sequencing, (4) number composition and decomposition, (5) adding and subtracting, (6) place value, and (7) multiplication and division. It also includes the ability to connect numerals to quantities. The TEAM uses an individual interview format. The maximum score on Part A of the TEAM is 104. For more details about this instrument, see Clements and Sarama (2011b). Cronbach's alpha for the TEAM scores for the sample from this study was .93.

2.3.2. SFON

We used the Elsi Bird Imitation task (Hannula & Lehtinen, 2005) at the start of the school year, and the Mailbox Imitation task (Hannula & Lehtinen, 2005) at the end of the school year. In these two versions of the SFON tasks, children are requested to feed a parrot with differently-colored berries and post differently-colored envelopes into a mailbox, respectively. Both versions consist of four trials, with two differently-colored numerosities per trial, ranging from one to three. For a detailed description of these two tasks, their administration, and their coding, see Hannula and Lehtinen (2005). We used a different version of the SFON task at the second measurement to prevent children from associating the task with a quantitative situation based on their memories of the first measurement (Bojorque et al., 2017). As indicated by a recent study (Hannula-Sormunen et al., in preparation), the SFON Elsi Bird Imitation Task and the SFON Mailbox Imitation Task are of equivalent difficulty. A group of 87 4–7-year-old children who were offered both variants of the SFON Imitation Task at the same measurement time, received an overall mean score of $M = 1.98$ (SD = 1.54) on the Elsi Bird Imitation Task and of $M = 2.15$ (SD = 1.64) on the Mailbox Imitation Task, the difference between the two task scores being not statistically significant, $t(86) = 1.62$, $p > .05$. Furthermore, Hannula and Lehtinen (2005) reported stability in children's SFON tendency assessed at the age of 4, 5, and 6 years (the average intraclass correlation was $r = 0.59$), using these two Imitation tasks, among others, which can be considered as a further indication of the equivalence of these two SFON tasks.

The correlation between SFON Test 1 and SFON Test 2 in our sample was $r = .41$, $p < .001$. Cronbach's alpha for the two SFON scores from the present sample was .76 for the Elsi Bird Imitation task and .79 for the Mailbox Imitation task.

2.3.3. Intelligence

The Vocabulary and Block Design subtests from the Spanish edition of the Wechsler preschool and primary scale of intelligence – III (WPPSI-III; Wechsler, 2002; Spanish Edition) were administered as indicators of children's verbal and non-verbal intelligence. More information on the WPPSI-III, including its reliability and validity, can be found in Wechsler (2002). Cronbach's alphas for the vocabulary subtest and the block design subtest scores from our sample were, .83 and .69, respectively.

2.3.4. Working memory

We used the Spanish version of the Odd One Out task from the Automated Working Memory Assessment Battery (AWMA; Alloway, 2007) to assess the visuospatial subsystem of children's Central Executive (CE). CE has shown to be significantly associated with mathematical skills (Friso-van den Bos et al., 2013). The tasks of the AWMA are reliable and valid assessments for measuring visuo-spatial short-term working memory (see Alloway, 2007, for further information).

2.3.5. Nature and quality of mathematics education

The COEMET (Sarama & Clements, 2009b) is a half-day administration instrument specifically designed to assess the quality of mathematics education in early education settings. Although the COEMET was developed by the same authors as the BB program, and, therefore, is based on the same guiding principles, it is not connected to any specific curriculum, allowing for intervention-control condition contrasts (Sarama, Lange, Clements, & Wolfe, 2012). The instrument is divided in two sections: (1) classroom culture (CC) and (2) specific math activities (SMA). It has 28 items, all but four of which are 5-point Likert scale (ranging from strongly disagree to strongly agree). The other four items are scored in terms of percentage of occurrence on a 5-point scale (0%; 1–25%; 26–50%; 51–75%; 76–100%). Maximum possible scores for each section and for the total scale are, CC = 45; SMA = 95; and COEMET total = 140. As stated above, the COEMET was administered two times per classroom (i.e., January and May). For each of these two observation moments, two observers spent a half-day in the classrooms from the beginning of the activities until lunch time. The observers took field notes and videotaped the lessons. To compute inter-rater reliability, two observers completed 10% of the COEMET scoring forms based on the notes and videos of those lessons. Inter-rater reliability (on this 10% of the data) was $K = .88$, $p < .001$. Next, one of the observers scored the rest of the COEMET forms. Cronbach's alpha for the COEMET scores in our sample was .94.

2.3.6. Fidelity

The Fidelity (Sarama & Clements, 2012) documents how the mathematics activities prescribed in the BB program are implemented by teachers. This instrument includes one section for each component of the implemented program (described in more detail below), namely (1) general curriculum (GC), (2) hands on center activities (HCA), (3) whole group activities (WGA), (4) small group activities (SGA), and (5) computer activities (CA). As mentioned below, we were not able to implement computer activities in the experimental schools due to the absence of computers in the classes, which is typically the case in Ecuadorian kindergarten. Consequently, the corresponding (fifth) part of the Fidelity instrument was not administered. The Fidelity instrument contains 39 items (without the computer activities part), and responses to all but seven of them are coded on 5-point Likert scale (ranging from strongly disagree to strongly agree). These items are scored as follows: −2 = strongly disagree, −1 = disagree, 0 = neutral, +1 = agree, +2 = strongly agree. The remaining seven items are "no" or "yes" items scored as −2 = no, +2 = yes. Maximum possible scores for each section are: GC = 10; HCA = 12; WGA = 14; SGA = 42. For a detailed description of this instrument, see Sarama and Clements (2012). As reported above, in the present study, Fidelity observations were made at three different moments. To complete each section of the instrument, two observers visited the experimental classrooms two times per moment: first, they observed the implementation of a complete lesson including hands-on center activities; second, they observed only the small group activities part. Observers took field notes and videotaped the lessons. To compute inter-rater reliability, two observers completed 10% of the Fidelity scoring forms based on the notes and videos of those lessons. Inter-rater reliability was $K = .92$, $p < .001$. Next, one of the observers scored the rest of the Fidelity forms. Cronbach's alpha for these Fidelity scores was .97.

2.4. Intervention

Although we focused only on children's development of competence with number and operations in this study, we implemented the whole intervention program to keep its integrity. Teachers in this study completed the 30 weeks of the program. The complete program was translated into Spanish and slightly adapted to the Ecuadorian context by the first author of this study and by an expert in the English and Spanish languages. The program has two components: program materials and professional development.

2.4.1. Program materials

The program materials include the teacher's edition, the teacher's resource guide, assessments, manipulatives, big books, and software activities. However, as the teachers involved in the present study had access to neither computers nor internet, it was not possible to include the software activities in the study.

The teacher's edition contains the complete daily lesson plans for 30 weeks, and suggestions about how to develop mathematical concepts. Each daily lesson is organized into a 40-min lesson and follows a consistent plan that includes *whole group* activities (warming-up activity to get children ready to do mathematics); *work time* (free hands-on math center on Monday, Wednesday, and Friday, and small group activities with the teacher on Tuesday and Thursday); *reflection* (questions encouraging children to talk about their thinking and reasoning); and *assessment* (informal assessment opportunities to record children's progress). The teacher's resource guide provides teachers with key tools (e.g., counting cards, puzzles, etc.) to help them deliver the program. It also includes weekly family letters to inform parents about what their children are doing in school and how to support their children at home. The assessment consists of simple record sheets that enable teachers to record and monitor children's participation and progress. The manipulative kit includes key manipulatives (e.g., connecting cubes, counters, number cubes, etc.) used for hands-on activities. Finally, the four big storybooks provide children with mathematics-related literature that they can use as much as they want.

2.4.2. Professional development

The teachers in the experimental group followed a five-day professional development training, with a duration of 40 h in total, before the start of the intervention, and two additional professional training days, with a duration of eight hours each, during the school year. These training days were organized and conducted by four researchers who extensively studied the program.

The first five days of professional development, focused on: (1) the theoretical framework of the program; (2) the learning trajectories for each mathematical topic associated to specific instructional activities; and (3) the BB program materials. During the professional development sessions, teachers had plenty of opportunities (around 15 h in total) to manipulate these materials, practice the implementation of the lessons for the first five weeks, set up learning centers, organize small group activities, and conduct assessments. Teachers also had the opportunity to watch a video of one lesson being implemented in a local kindergarten classroom. The sessions included hands-on experience in implementing the program, with an emphasis on interaction and communication among teachers. In the two additional days of the professional development, teachers observed and discussed videos of themselves enacting the activities of the program in their classrooms, shortly revised the learning trajectories, and had the opportunity to practice with the program materials for the two coming weeks. Teachers' attendance to the professional development sessions was 100%.

The professional development sessions were complemented by weekly in-class coaching visits during the mathematics lessons. Coaches visited the teachers' classrooms once per week (for a total of 30 weeks) and provided them with constructive feedback as to maximize the effective implementation of the program. The procedure followed during the coaches' visits was: (1) the coaches observed the implementation of the lesson and made comments on both the positive aspects of the lesson and those aspects that needed to be improved, (2) the teachers had the opportunity to ask questions, to talk about their classes, and to plan activities for the coming days, with the assistance of the coaches, and (3) the coaches reminded the teachers of their commitment with the program implementation, but also of the fact that they could count on the support and help of the research team.

2.5. Control group

The teachers in the control condition continued using the regular national mathematics curriculum given by the Ministry of Education (Ministerio de Educación, 2010). This curriculum focuses on five areas, namely (1) relations and functions of objects (examples of learning goals of relations and functions of objects are "Describe the features of surrounding objects" and "Reproduce, describe, and make patterns of objects"); (2) number (examples of number learning goals include "Identify quantities and associating them with the numerals 8, 9 and 0" and "Add and subtract using whole numbers from 0 to 10"); (3) geometry (examples of geometry learning goals are "Identify geometric solids in surrounding objects" and "Classify geometric shapes using surrounding objects"); (4) measurement (examples of measurement learning goals include "Recognize and compare objects according to their size" and "Recognize and compare objects according to their length"; and (5) statistics and probability (examples of statistics and probability learning goals are "Identify likely and unlikely events in everyday situations" and "Collect and represent information about the environment in pictograms"). The major numerical topics the control teachers addressed during the school year were: (1) counting, (2) quantity identification and association with numerals, (3) comparing and ordering, (4) adding and subtracting; and (5) reading and writing numerals. The program materials include the national curriculum, a national textbook, and the accompanying teachers' guide. The use of the national textbook and its accompanying teacher's guide is compulsory for public kindergartens. Private kindergartens use other publically available textbooks and accompanying teachers' guides aligned to the national curriculum. The teachers in the control group were (just as the teachers from the experimental group at the start of the study) used to the national curriculum as it has been implemented in the Ecuadorian kindergarten classrooms since the year 2010. The mathematics lessons were organized and conducted by the teacher on a daily basis, with a duration of 40 min each. Given that in Ecuador professional development for teachers is limited, the teachers in the control group did not receive any professional development on early mathematics education before or during the intervention. To motivate the teachers from the control schools, we offered them the same professional development program at the completion of the research study.

2.6. Analyses

To take into account the nested structure of the data (i.e., children nested within schools), child outcome data were analyzed using multilevel regression analyses. More specifically, we conducted multilevel regression analyses in IBM SPSS 24 using the Mixed Models technique (Hayes, 2006). To analyze children's gain in early numerical abilities (Hypothesis 1), we included children's pretest score (Test 1) as independent variable while their posttest score (Test 2) was the dependent variable. The same method was followed concerning children's SFON (Hypothesis 2). We controlled for the contribution of children's working memory, verbal and nonverbal intelligence, age, and SES in all analyses. The order of the covariates differed between the analyses depending on their correlation with the dependent variables. We always used the raw test scores in our analyses except for the Working Memory test, for which we used the standard scores provided by the computerized program utilized to assess working memory. Finally, given the small number of schools in our study, we used a non-parametric test, i.e., Mann Whitney *U* test (Mann & Whitney, 1947) to compare the quality of early mathematics education between the experimental and control condition (Research question 1).

Table 2
Means, standard deviations, and range scores for the experimental and control condition.

Subscales[a]	Experimental (*N* = 176)			Control (*N* = 177)		
	M	SD	Range	*M*	SD	Range
Working memory	80.32	16.15	60–131	77.72	13.37	60–120
Verbal intelligence	10.38	4.85	2–27	10.40	5.43	1–27
Non-verbal intelligence	19.69	4.71	7–36	20.22	5.00	1–35
TENA	25.32	8.71	7–49	25.37	8.50	7–49
TEAM	14.85	8.74	0–43	16.59	8.66	0–49
SFON	0.76	1.24	0–4	0.52	0.98	0–4

[a] We reported the raw scores for all the tests except for the Working Memory test, for which we reported the standardized scores.

3. Results

The results are organized in five sections. First, we present the compatibility of the experimental and control condition at the start of the school year. Second, we report on the fidelity of implementation of the BB program. Third, we examine the effectiveness of the BB program for children's early numerical abilities development, compared to the control condition (Hypothesis 1). Fourth, we report the effectiveness of the BB program for children's SFON, compared to the control condition (Hypothesis 2). Finally, we examine the quality of mathematics education offered by the teachers in the experimental group compared to the teachers in the control group (Research question 1).

3.1. Compatibility of the experimental and control condition

Table 2 presents the descriptive statistics on all variables measured at the start of the school year, namely working memory (verbal and non-verbal), intelligence, early numerical abilities, and SFON. Multilevel analyses (see Table 3) revealed that the experimental and control condition did not differ on any variable at the start of the study.

3.2. Fidelity of implementation

To measure the fidelity of implementation of the BB program, we computed the mean scores on total Fidelity on the three observations as well as the means on the four subscales of the Fidelity. With responses ranging from −2 (strongly disagree) to +2 (strongly agree), the mean on total Fidelity scores was 1.17, averaging near agree. The means per subscale were (1) GC = 1.71, averaging near strongly agree, (2) HCA = 0.88, averaging near agree, (3) WGA = 1.28, averaging near agree, and (4) SGA = 0.79, averaging near agree.

Table 3
Multilevel models of the initial comparison between experimental and control condition.

Variable	Coeff.	SE	df	*t*	Sig.
Working memory					
Intercept	80.35	2.06	18.14	39.02	.001
Control group	−2.68	2.91	18.11	−0.92	.369
Verbal intelligence					
Intercept	10.42	0.86	18.01	12.08	.001
Control group	−0.05	1.22	17.99	−0.04	.970
Non-verbal intelligence					
Intercept	19.73	0.71	17.95	27.70	.001
Control group	0.49	1.01	17.92	0.49	.630
SFON Test 1					
Intercept	0.76	0.12	18.16	6.59	.001
Control group	−0.24	0.17	18.11	−1.47	.159
TENA Test 1					
Intercept	25.33	1.06	18.08	23.84	.001
Control group	0.02	1.50	18.04	0.02	.988
TEAM Test 1					
Intercept	14.84	1.10	18.09	13.54	.001
Control group	1.74	1.55	18.06	1.13	.275
SES					
Intercept	5.10	0.87	15.98	5.85	.001
Control group	−0.15	0.55	15.98	−0.27	.793

These results indicate that the teachers from the experimental group implemented the BB program with adequate fidelity.

3.3. Effectiveness of the BB program for children's early numerical abilities

Children's early numerical abilities were measured using two tests: a standards-based early numerical test (i.e., TENA) and an international test for early assessment in mathematics (i.e., TEAM). Given the high correlations between children's scores on these two tests at both the start (Test 1) and the end (Test 2) of the kindergarten year (respectively, $r = .82$, $p = .01$, and $r = .76$, $p = .01$), we decided to create one global score for early numerical abilities (ENA).[2] We did this by standardizing (calculating *z*-scores) the scores of the TENA and the TEAM and then calculating the mean of those standardized scores.

To test our first hypothesis, namely that children who follow the BB program would make more progress in their acquisition of early numerical abilities than children from the control group as indicated by the differences in their gain on early numerical abilities tests at the end of the school year, we computed a multilevel regression model. In this model, ENA Test 2 was the dependent variable. As control variables, the grand-mean centered scores of children's working memory, SES, verbal and non-verbal intelligence, age, and the scores of ENA Test 1 were entered in the model as independent variables. The order of the independent variables is based on their correlation with the dependent variable. The dichotomous intervention variable (i.e., control, experimental) was also added as independent variable to this model. Table 4 summarizes the results of these analyses. As displayed in Table 4, only children's ENA pretest scores and the intervention variable were significant. This means that, given that the children in both conditions (experimental, control) started at the same level, children in the experimental group not only had significantly higher early numerical abilities scores at the posttest, but also gained more in early numerical abilities between pretest and posttest than children in the control group did.

Table 5 presents the variance partitioning for the null and the full model. Table 5, shows that the multilevel regression model as presented in Table 4 explained 66.29% of the total variance, with 64.18% explained variance at the student level and 72.72% explained variance at the school level. These results allowed us to confirm our first hypothesis stating that children from the experimental group would outperform those from the control group for early numeracy. This difference was moreover characterized by a medium to large effect size (*Hedge's g* = 0.73).

[2] We also analyzed the data for TENA and TEAM separately, which resulted in the same findings as combining both measures.

Table 4
Multilevel model of the BB program's impact on children's early numerical abilities.

Variable	Coeff.	SE	df	*t*	Sig.	−2LL[a]
Intercept	1.07	0.20	17.06	5.40	0.001	894.42
Working memory	0.01	0.01	346.98	0.94	0.351	829.74
SES	0.02	0.02	328.98	0.85	0.395	819.51
Verbal intelligence	−0.01	0.01	352.98	−0.19	0.850	802.82
Non-verbal intelligence	0.01	0.01	352.80	0.28	0.780	781.50
Age	−0.01	0.01	344.42	−0.90	0.367	778.53
ENA Test 1	0.71	0.04	344.59	18.79	0.001	541.65
Control group	−0.72	0.13	17.07	−5.69	0.001	523.24

Note. R^2 = .66.
[a] Including additional predictor. Given that the predictor variables working memory and intelligence do not contain a meaningful value of zero (i.e., nobody has a working memory or intelligence of zero), we grand-mean centered the scores of these variables to help the interpretation of the parameter estimates (coefficients).

Table 5
Variance partitioning of the null and the full model.

	Variance null model	Variance full model
Student level	0.67	0.24
School level	0.22	0.06
Total variance	0.89	0.30

Table 6
Multilevel model of the BB program's impact on children's SFON.

Variable	Coeff.	SE	df	*t*	Sig.	−2LL[a]
Intercept	2.44	0.31	18.30	7.80	0.001	1282.23
Working memory	0.01	0.01	344.99	1.16	0.248	1268.73
Non-verbal intelligence	0.04	0.02	321.24	2.52	0.012	1256.82
SES	0.09	0.05	176.30	1.82	0.070	1254.51
Verbal intelligence	−0.01	0.02	315.22	−.082	0.412	1253.77
Age	−0.02	0.02	352.57	−0.93	0.352	1253.47
SFON Test 1	0.46	0.07	351.25	6.73	0.001	1209.60
Control group	−0.69	0.19	17.00	−3.55	0.002	1199.99

Note. R^2 = .24.
[a] Including additional predictor.

3.4. Effectiveness of the BB program for children's SFON

To test our second hypothesis, namely that children who follow the BB program would make more progress in their SFON than children from the control group as indicated by the differences in their gain on SFON tasks at the end of the school year, we computed a multilevel model with SFON Test 2 as dependent variable, and the grand-mean centered scores of working memory, verbal intelligence, SES, non-verbal intelligence, age, the scores of SFON Test 1, and the dichotomous intervention variable (i.e., control, experimental) as independent variable. As in the previous analysis, the order of the independent variables was based on their correlation with the dependent variable. After controlling for the effects of working memory, verbal intelligence, SES, non-verbal intelligence, age, and SFON Test 1, the impact of the BB program on children's SFON at the end of kindergarten was significant (see Table 6). Children in the experimental group not only had significantly higher SFON scores at the end of the school year, but also gained more in SFON than children in the control group between the start and the end of the school year.

Table 7
Variance partitioning of the null and the full model.

	Variance null model	Variance full model
Student level	2.06	1.70
School level	0.33	0.08
Total variance	2.39	1.78

The variance partitioning for the null and the full model is reported in Table 7. This analysis indicates that the multilevel regression model as displayed in Table 6 explained 25.52% of the total variance. Of this 25.52% explained variance, 17.48% was explained at the student level and 75.76% at the school level.

These results confirmed our second hypothesis stating that children in the experimental group would outperform those from the control group in SFON. This difference was moreover characterized by a medium effect size (*Hedge's* $g = 0.53$).

3.5. Quality of mathematics education

Our research question was whether teachers who follow the BB program offer higher quality mathematics education than the teachers from the control group, as indicated by their scores on the COEMET. Table 8 presents the means and standard deviations for the COEMET. As can be observed, the experimental teachers outperformed the control teachers on the two observations for the COEMET as a whole as well as for the two subscales. To test the significance of the observed differences in COEMET scores, we conducted Mann Whitney *U* tests with the mean total COEMET score for the two observations as well as with the means per subscales (i.e., CC total scores and SMA total scores) also for the two observations. Because the scores for the two observation moments were highly correlated, $r = .93$, $p < .001$, we used the mean COEMET score for these two observations in our analyses.

The results of these analyses revealed that there was a significant difference in the total COEMET scores between the experimental ($M = 99.86$) and the control teachers ($M = 63.98$), $U = 75.00$, $z = 3.05$, $p < .01$, $r = 0.72$. There were also significant differences at the subscale level. The experimental teachers ($M = 29.06$) performed better than the control teachers ($M = 17.28$), $U = 79.50$, $z = 3.45$, $p < .01$, $r = 0.81$ on the CC section. The former also ($M = 70.81$) scored higher than the latter ($M = 46.70$), $U = 72.00$, $z = 2.782$, $p < .01$, $r = .66$ on the SMA section. The teachers in the experimental group thus offered higher quality mathematics education than the teachers in the control group as measured by the COEMET.

4. Discussion

The main objective of our study was to evaluate the effectiveness of the BB program for enhancing Ecuadorian kindergartners' early numerical abilities and SFON. Following a pretest-intervention-posttest design, the participating schools were randomly assigned to either an experimental (BB program) or a control (regular mathematics program) condition. At the beginning of the study, both groups were comparable in working memory, verbal and non-verbal intelligence, age, SES, early numerical abilities, and SFON. Consistent with prior studies (Clements & Sarama, 2008; Clements et al., 2011), the teachers in the experimental group implemented the BB program with adequate fidelity. Hereafter, we summarize and discuss the major findings of the study. We first discuss the findings related to the two hypotheses and the research question. We then reflect on the theoretical and educational implications.

4.1. Effectiveness of the BB program for young children's early numerical abilities

A first major finding of this study is that children who received the BB program made more progress in both their early numerical abilities than their peers who did not follow the program. Although it is not possible to identify which aspect(s) of the BB program contributed to children's greater gain in early numerical competencies, there are at least three aspects that may have played an important role. The first aspect relates to teachers' participation in the professional development program, with a specific focus on children's mathematical thinking and learning trajectories. This kind of intensive and high-quality professional development is unusual in Ecuador. Furthermore, professional development on early mathematics education for teachers is limited. We suggest that the BB professional development program enabled and motivated the teachers in the experimental group to involve their children in more and higher-quality early numerical experiences. Given that we did not assess teaching quality before the start of the study, we unfortunately cannot test this suggestion. Therefore, we recommend that future studies should also include a COEMET measure before the start of the intervention study.

The second aspect may have to do with two early numerical components that are included in the BB program but not in the Ecuadorian national curriculum and thus also not in the regular (control) mathematics classes, namely subitizing and composing numbers. Given the important role of these two components in young children's numerical development (Baroody, 2004; Le Corre, Van de Walle, Brannon, & Carey, 2006), we hypothesize that the inclusion of subitizing and composing activities in the BB classes additionally stimulated children's early numeracy development.

The third aspect refers to the nature of the BB program activities, and more concretely (1) children's active participation in daily games, (2) the learning centers, which promoted hands-on mathematics activities and mathematical discussion, and (3) reflection time, encouraging children to talk about their thinking and reasoning (Clements & Sarama, 2013). These three features are opposed to what usually happens in Ecuadorian kindergarten classes. As documented by a recent study (Bojorque, Torbeyns, Van Nijlen, & Verschaffel, 2018), Ecuadorian kindergartners typically complete written worksheets and have hardly any opportunity to interact with their peers and to reflect on their own strategies. Future research, involving several experimental groups that systematically vary on these three major features of the BB program (and also other variables that might enhance children's early numeracy development) in combination with more intensive qualitative analyses of what actually happened in the experimental and control classes, are needed to disentangle the relative contribution of these and possible other important features of the BB program to children's early numeracy development.

4.2. Effectiveness of the BB program for young children's SFON

A second major result of the study is that children who received the BB program had significantly greater gain in SFON compared to children who did not receive the program. This finding is particularly noteworthy given that the BB program does not deliberately and explicitly focus on stimulating children's SFON development. However, some activities included in the program such as *I see numbers* may have prompted children's SFON development. This finding provides support for the claim that SFON can be enhanced through meaningful guided activities (Hannula-Sormunen, 2015; Hannula et al., 2005). It is also consistent with the finding of Hannula et al. (2005) that directing children's attention towards small numbers of items enhances SFON. As children's SFON and early numerical abilities are reciprocally related (Hannula & Lehtinen,

Table 8
Means and standard deviations of COEMET scores.

	Experimental group						Control group					
	Observation 1[a]		Observation 2[b]		Mean 2 Obs.		Observation 1		Observation 2		Mean 2 Obs.	
	M	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD
Classroom culture	28.78	5.14	29.33	2.74	29.06	3.14	18.11	7.85	16.44	3.05	17.28	5.14
Specific math activities	70.58	17.88	71.03	14.12	70.81	15.79	46.19	14.97	47.21	10.93	46.70	12.80
Total COEMET	99.36	22.66	100.36	15.63	99.86	18.60	64.31	22.07	63.65	12.97	63.98	17.31

Note.
[a] First observation = January.
[b] The second observation = May.

2005; Hannula et al., 2010), children's progress in early numerical abilities might also have contributed to their progress in SFON and vice versa. As reported by Hannula-Sormunen et al. (2015), SFON is a separate and domain-specific attentional process within children's existing early numerical competencies. This attentional process differs from children's early numerical abilities. Accordingly, these authors found that SFON and verbal counting skills are distinct, but correlated aspects of early mathematical development (Hannula-Sormunen et al., 2015). Notwithstanding the theoretical and empirical arguments of Hannula-Sormunen et al., further evidence on the differences between SFON and children's early numerical abilities is needed. In view of addressing their divergent validity, future studies offering different types of SFON and early numerical abilities tasks and applying confirmatory factor analysis are required. In addition to these studies aiming at addressing the conceptual and methodological challenges related to the construct of SFON, future intervention studies that combine the implementation of the BB program with a program aiming at intentionally stimulating children's SFON development are welcome to optimally enhance children's early numerical competencies. Moreover, these intervention studies can help to scrutinize more deeply the relative contribution of and interaction between the ability and disposition aspects of children's early numerical competencies.

4.3. Quality of mathematics education

A third major finding was that the quality of early mathematics education provided by the teachers in the experimental group was higher relative to the teachers in the control group. This finding is consistent with Clements and et al.'s findings (e.g., Clements & Sarama, 2008), and is in line with what one would expect. In general, the teachers in the BB group offered richer classroom environments, including more opportunities for children to connect mathematical ideas with daily-life situations, to interact with peers, to reason about and share their mathematical ideas, to communicate their strategies, all of which are considered effective teaching practices (Epstein, 2007). In contrast, the teachers in the control group applied a more direct teaching approach, emphasizing individual work, and offered limited opportunities for the children to interact with their peers and to reflect on and communicate their own strategies. These teaching practices are more in line with practices that negatively impact on children's learning (see e.g., Marcon, 2002). Importantly, the higher quality of early mathematics education provided by the experimental teacher remained stable between January and May, probably due to our efforts to encourage the teachers to apply what they learned during the first intensive professional development. Future studies are needed to empirically address this hypothetical explanation.

Unfortunately, the small number of schools in our study did not allow us to examine whether teachers' fidelity of implementation of the BB program had a positive effect on the quality of mathematics education offered. Prior studies of Sarama and colleagues reported that higher levels of fidelity of implementation of the BB program resulted in higher scores on the quality of education in the intervention classes, as well as in greater gains in experimental children's mathematics achievement compared to the control children (Sarama, Clements, Starkey, Klein, & Wakeley, 2008). By including a larger sample of schools, future studies can address this limitation. In these studies, it would be important to examine whether the fidelity of implementation of the BB program has a direct positive effect on the quality of the mathematics education offered, and also, whether the fidelity of implementation of the BB program has an indirect positive effect on children's progression in early numerical abilities and in SFON.

Future research including a larger sample of children and schools is also needed in developing countries to evaluate whether the implementation of the BB program is effective when implemented on a large scale (e.g., Clements et al., 2011), providing less opportunity to establish the intensive contacts with and permanent coaching of teachers in the experimental group. These large-scale studies should also follow children through the initial years of formal education as to examine the persistence of the observed positive effects.

4.4. Theoretical implications

Our findings on the effectiveness of the BB program for children's early numeracy development add, in three important ways, to the existing body of research in the domain. First, as previous early mathematics intervention studies have been conducted in developed countries, mainly in the US (Clements et al., 2011; Griffin, 2005; Lewis Presser et al., 2015; Starkey et al., 2004), our findings complement this body of research by demonstrating the effectiveness of an early mathematics program (i.e., the BB program), for enhancing children's early numerical competencies in developing countries (i.e., Ecuador). Given the differences in political, cultural, economic, and educational characteristics of developed versus developing countries, our findings suggest that the same essential features of such intervention programs are effective in such largely differing countries.

A second important contribution to the available research is that the BB program proved highly effective for enhancing children's early numeracy development even when controlling for two variables (in addition to age and SES) that are known to importantly contribute to early numeracy development but have not been used in most intervention studies so far, namely working memory and intelligence.

Third, as previous early mathematics intervention studies have focused only on the ability component of children's early numerical competencies, leaving aside the dispositional component (i.e., SFON), our investigation also points to the importance of this dispositional component in current theoretical models of early numeracy and in the development, implementation, and evaluation of early mathematics programs.

Although our results empirically support the effectiveness of the BB program for children's early numeracy development, we were not able to implement one of its essential features, namely the BB software activities (Clements & Sarama, 2013). A prior study aiming at improving the mathematics abilities of young children through computer-assisted instruction, implemented the BB software in kindergarten classrooms, reporting positive gains in young children's numerical abilities, (Foster, Anthony, Clements, Sarama, & Williams, 2016; Sarama, 2004). Therefore, it is important that future research in developing countries also include the BB software activities to further unravel the contribution of also this program feature to children's early numeracy development.

4.5. Educational implications

This study provides additional empirical evidence for the effectiveness of early mathematics programs in general (Klein, Starkey, Clements, Sarama, & Iyer, 2008; Lewis Presser et al., 2015; Starkey et al., 2004), and of the BB program in particular (e.g., Clements et al., 2011; Clements & Sarama, 2008), for young children's early numeracy development. Importantly, children from different SES levels appeared to benefit from the research-based early mathematics programs addressed in the present study. This finding reveals the need for focused interventions aimed at improving the early numerical development of children from all SES levels during as well as before kindergarten. Focused intervention programs before the onset of kindergarten and in out-of-school contexts can additionally enhance the early numerical competencies of children from different SES levels, and as such provide the necessary stepping stones towards a higher mastery and more fluent development of early numerical competencies during the first years of kindergarten and formal mathematics education. Given that we operationalized SES via only the mother's educational level, our findings need to be confirmed in future studies using also other SES indicators such as family income or a composite score of both parents' education and family income (e.g., Galindo & Sonnenschein, 2015).

Finally, this is the first early mathematics intervention study conducted in a developing country, i.e., Ecuador (United Nations, 2016). As such, our study addresses the current gap in our knowledge of Ecuadorian young children's early numerical competencies. The fact that children in the experimental group (taught with the BB program) outperformed children in the control group (taught with the regular national curriculum) raises serious concerns regarding, on the one hand, the adequacy of the national curriculum for promoting children's early numerical competencies, and, on the other hand, the preparedness of the kindergarten teachers to teach early numeracy. Therefore, a revision of the Ecuadorian kindergarten curriculum for mathematics and the corresponding instructional materials in light of the relevant international literature, as well as the incorporation of well-designed professional development aiming at improving current teaching practices is more than necessary as a means to support the development of Ecuadorian children's numerical competencies.

Acknowledgements

This study was supported by the Bilateral Scientific Cooperation Project "Early math development in Ecuadorian Kindergartners: The role of child and context characteristics" sponsored by the Research Foundation Flanders, Project G0C8715N and the Secretaría de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT), Project 20150008CI.

References

Alloway, T. P. (2007). *Automated working memory assessment*. London: Psychological Corporation.

Aunio, P., & Niemivirta, M. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences*, *20*(5), 427–435. http://dx.doi.org/10.1016/j.lindif.2010.06.003

Baroody, A. J. (2004). The developmental bases for early childhood operations and number standards. In D. H. Clements, & J. Sarama (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 173–219). Mahwah, NJ: Lawrence Erlbaum Associates.

Bojorque, G., Torbeyns, J., Hannula-Sormunen, M., Van Nijlen, D., & Verschaffel, L. (2017). Development of SFON in Ecuadorian kindergartners. *European Journal of Psychology of Education*, *32*(3), 449–462. http://dx.doi.org/10.1007/s10212-016-0306-9

Bojorque, G., Torbeyns, J., Moscoso, J., Van Nijlen, D., & Verschaffel, L. (2015). Early number and arithmetic performance of Ecuadorian 4–5-year-olds. *Educational Studies*, *41*, 565–586. http://dx.doi.org/10.1080/03055698.2015.1090302

Bojorque, G., Torbeyns, J., Van Nijlen, D., & Verschaffel, L. (2018). Ecuadorian Kindergartners' numerical development: Contribution of SES, quality of early mathematics education and school type. *Educação e Pesquisa*, *44*, 1–21. http://dx.doi.org/10.1590/S1678-463444201801164156

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, *33*(3), 205–228. http://dx.doi.org/10.1080/87565640801982312

Clements, D. H. (2007). Curriculum research: Toward a framework for research-based curricula. *Journal for Research in Mathematics Education*, *38*(1), 35–70. http://dx.doi.org/10.2307/30034927

Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning*, *6*(2), 81–89. http://dx.doi.org/10.1207/s15327833mtl0602

Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. *Journal for Research in Mathematics Education*, *38*(2), 136–163. http://dx.doi.org/10.2307/30034954

Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal*, *45*(2), 443–494. http://dx.doi.org/10.3102/0002831207312908

Clements, D. H., & Sarama, J. (2011a). Mathematics knowledge of low-income entering preschoolers. *Far East Journal of Mathematical Education*, *6*, 41–63.

Clements, D. H., & Sarama, J. (2011b). *Tools for early assessment in math (TEAM)*. Columbus, OH: McGraw-Hill.

Clements, D. H., & Sarama, J. (2013). . *Building Blocks* (Vols. 1–2) Columbus, OH: McGraw-Hill Education.

Clements, D. H., Sarama, J., Spitler, M. E., Lange, A. A., & Wolfe, C. B. (2011). Mathematics learned by young children in an intervention based on learning trajectories: A large-scale cluster randomized trial. *Journal for Research in Mathematics Education*, *42*(2), 127–166.

Clements, D. H., Sarama, J., Wolfe, C. B., & Spitler, M. E. (2013). Longitudinal evaluation of a scale-up model for teaching mathematics with trajectories and technologies: Persistence of effects in the third year. *Journal of Research on Educational Effectiveness*, *5*(4), 812–850. http://dx.doi.org/10.1080/19345747.2011.627980

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, *103*(4), 469–479. http://dx.doi.org/10.1016/j.jecp.2009.01.010

Epstein, A. S. (2007). *The intentional teacher: Choosing the best strategies for young children's learning*. Washington, DC: National Association for the Education of Young Children.

Foster, M. E., Anthony, J. L., Clements, D. H., Sarama, J., & Williams, J. M. (2016). Improving mathematics learning of kindergarten students through computer-assisted instruction. *Journal for Research in Mathematics Education*, *47*(3), 206–232. http://dx.doi.org/10.5951/jresematheduc.47.3.0206

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & Van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. http://dx.doi.org/10.1016/j.edurev.2013.05.003

Galindo, C., & Sonnenschein, S. (2015). Decreasing the SES math achievement gap: Initial math proficiency and home learning environments. *Contemporary Educational Psychology*, *43*, 25–38.

Geary, D. C., Hoard, M. K., & Nugent, L. (2012). Independent contributions of the central executive, intelligence, and in-class attentive behavior to developmental change in the strategies used to solve addition problems. *Journal of Experimental Child Psychology*, *113*(1), 49–65. http://dx.doi.org/10.1016/j.jecp.2012.03.003

Ginsburg, H. P., Greenes, C. R., & Balfanz, R. (2003). *Big math for little kids*. Parsippany, NJ: Dale Seymour Publications, Pearson Learning Group.

Griffin, S. (2004). Number Worlds. A research-based mathematics program for young children. In D. H. Clements, J. Sarama, & A.-M. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 325–342). Mahwah, NJ: Lawrence Erlbaum Associates.

Griffin, S. (2005). Fostering de development of whole-number sense: Teaching mathematics in the primary grades. In M. S. Donovan, & J. D. Bransford (Eds.),

How Students Learn: History, mathematics, and science in the classroom (pp. 257–308). Washington, DC: The National Academies Press.
Griffin, S. (2007). *Number Worlds: A mathematics intervention program from Grades preK-6*. Columbus, OH: SRA/McGraw-Hill.
Hannula, M. M., & Lehtinen, E. (2005). Spontaneous focusing on numerosity and mathematical skills of young children. *Learning and Instruction*, *15*(3), 237–256. http://dx.doi.org/10.1016/j.learninstruc.2005.04.005
Hannula, M. M., Lepola, J., & Lehtinen, E. (2010). Spontaneous focusing on numerosity as a domain-specific predictor of arithmetical skills. *Journal of Experimental Child Psychology*, *107*(4), 394–406. http://dx.doi.org/10.1016/j.jecp.2010.06.004
Hannula, M. M., Mattinen, A., & Lehtinen, E. (2005). Does social interaction influence 3-year-old children's tendency to focus on numerosity? A quasi-experimental study in day care. In L. Verschaffel, E. De Corte, G. Kanselaar, & M. Valcke (Eds.), *Powerful learning environments for promoting deep conceptual and strategic learning (Studia Pae)* (pp. 63–80). Leuven: Leuven University Press.
Hannula, M. M., Rasanen, P., & Lehtinen, E. (2007). Development of counting skills: role of spontaneous focusing on numerosity and subitizing-based enumeration. *Mathematical Thinking and Learning*, *9*(1), 51–57. http://dx.doi.org/10.1207/s15327833mtl0901_4
Hannula-Sormunen, M. M. (2015). Spontaneous focusing on numerosity and its relation to counting and arithmetic. In R. Cohen Kados, & A. Dowker (Eds.), *The Oxford handbook of numerical cognition* (pp. 1689–1699). Oxford, UK: Oxford University Press. http://dx.doi.org/10.1017/CBO9781107415324.004
Hannula-Sormunen, M. M., Lehtinen, E., & Räsänen, P. (2015). Preschool children's spontaneous focusing on numerosity, subitizing, and counting skills as predictors of their mathematical performance seven years later at school. *Mathematical Thinking and Learning*, *17*(2–3), 155–177. http://dx.doi.org/10.1080/10986065.2015.1016814
Hannula-Sormunen, M. M., Torbeyns, J., Kyttälä, M., Simms, V., De Smedt, B., Batchelor, S. (in preparation). Action and verbally based measures of spontaneous focusing on numerosity.
Hayes, A. F. (2006). A primer on multilevel modeling. *Human Communication Research*, *32*(4), 385–410. http://dx.doi.org/10.1111/j.1468-2958.2006.00281.x
Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, *45*(3), 850–867. http://dx.doi.org/10.1037/a0014939
Klein, A., Starkey, P., Clements, D., Sarama, J., & Iyer, R. (2008). Effects of a pre-kindergarten mathematics intervention: A randomized experiment. *Journal of Research on Educational Effectiveness*, *1*(3), 155–178. http://dx.doi.org/10.1080/19345740802114533
Klein, A., Starkey, P., & Ramirez, A. B. (2002). *Pre-K mathematics curriculum: Early childhood*. Glenview, IL: ScottForesman.
Le Corre, M., Van de Walle, G., Brannon, E. M., & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology*, *52*(2), 130–169. http://dx.doi.org/10.1016/j.cogpsych.2005.07.002
Lewis Presser, A., Clements, M., Ginsburg, H. P., & Ertle, B. (2015). Big Math for Little Kids: The effectiveness of a preschool and kindergarten mathematics curriculum. *Early Education and Development*, *26*(3), 399–426. http://dx.doi.org/10.1080/10409289.2015.994451
Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, *18*, 50–60.
Marcon, R. (2002). Moving up the grades: Relationship between pre-school model and later school success. *Early Childhood Research and Practice*, *4*(1), 517–530.
Ministerio de Educación. (2010). *Actualización y fortalecimiento curricular de la educación general básica*. Quito–Ecuador: Primer año.
Mulligan, J., Verschaffel, L., Baccaglini-Frank, A., Coles, A., Gould, P., Obersteiner, A.,. . ., Roberts, N. (n.d.). Whole number thinking, learning and development. In M. G. Bartolini Bussi, & X. Sun (Eds.), *Building the foundation: Whole numbers in the primary grades*. The ICMI Study 23. New York: Springer.
Passolunghi, M. C., Lanfranchi, S., Altoè, G., & Sollazzo, N. (2015). Early numerical abilities and cognitive skills in kindergarten children. *Journal of Experimental Child Psychology*, *135*, 25–42. http://dx.doi.org/10.1016/j.jecp.2015.02.001
Putnam, R. (2003). Commentary on four elementary mathematics curricula. In S. Senk, & D. Thompson (Eds.), *Standards oriented school mathematics curricula: What does research say about student outcomes* (pp. 161–178). Mahwah, NJ: Erlbaum.
Ramani, G. B., Siegler, R. S., & Hitti, A. (2012). Taking it to the classroom: Number board games as a small group learning activity. *Journal of Educational Psychology*, *104*(3), 661–672. http://dx.doi.org/10.1037/a0028995
Sarama, J. (2004). Technology in early childhood mathematics: building blocks as an innovative technology-based curriculum. In A.-M. Clements, D. H. Sarama, & J. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 361–375). Mahwah: Lawrence Erlbaum Associates.
Sarama, J., & Clements, D. H. (2009a). *Early childhood mathematics education research: Learning trajectories for young children*. New York: Routledge.
Sarama, J., & Clements, D. H. (2009b). *Manual for classroom observation (COEMET)-Version 3. Unpublished version*.
Sarama, J., & Clements, D. H. (2012). *Manual for TRIAD fidelity of implementation (near fidelity). Unpublished version*.
Sarama, J., Clements, D. H., Starkey, P., Klein, A., & Wakeley, A. (2008). Scaling up the implementation of a pre-kindergarten mathematics curriculum: Teaching for understanding with trajectories and technologies. *Journal of Research on Educational Effectiveness*, *1*(2), 89–119. http://dx.doi.org/10.1080/19345740801941332
Sarama, J., Clements, D. H., Wolfe, C. B., & Spitler, M. E. (2012). Longitudinal evaluation of a scale-up model for teaching mathematics with trajectories and technologies. *Journal of Research on Educational Effectiveness*, *5*(2), 105–135, 10.10801193457 47.2011.627980.
Sarama, J., Lange, A. A., Clements, D. H., & Wolfe, C. B. (2012). The impacts of an early mathematics curriculum on oral language and literacy. *Early Childhood Research Quarterly*, *27*(3), 489–502. http://dx.doi.org/10.1016/j.ecresq.2011.12.002
Schoenfeld, A. (2007). Method. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 69–107). Charlotte, NC: Information Age Publ.
Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science*, *11*(5), 655–661. http://dx.doi.org/10.1111/j.1467-7687.2008.00714.x
Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly*, *19*(1), 99–120. http://dx.doi.org/10.1016/j.ecresq.2004.01.002
Swanson, H. L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, *100*(2), 343–379. http://dx.doi.org/10.1037/0022-0663.100.2.343
United Nations. (2016). *World economic situation and prospects 2016*. New York: United Nations.
Wechsler, D. (2002). *Wechsler preschool and primary scale of intelligence – third edition*. New York: Psychological Corporation.