Empirical Research

Check for updates

Improving Mathematics Performance in 7-Year-Old Children: Training the Mapping From Estimated Quantities to Arabic Digits

Nuria Ferres-Forga [1] , Justin Halberda [2] , Ariadna Batalla-Ferres [3] , Luca L. Bonatti [1,4,5]

[1] *Center for Brain and Cognition, Universitat Pompeu Fabra, Barcelona, Spain.* [2] *Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD, USA.* [3] *Applied Statistics and Sociology, Universitat Autònoma de Barcelona, Barcelona, Spain.* [4] *ICREA (Catalan Institution for Research and Advanced Studies), Barcelona, Spain.* [5] *Università Ca' Foscari, Venezia, Italy.*

Corresponding Author: Nuria Ferres-Forga, Universitat Pompeu Fabra, Center for Brain and Cognition, Ramon Trias Fargas, 25-27, 08005 Barcelona, Spain. E-mail: nuriaferresforga@gmail.com

Supplementary Materials: Data, Materials [see Index of Supplementary Materials]

Abstract

Exact arithmetic abilities require symbolic numerals, which constitute a precise representation of quantities, such as the Arabic digits. Numerical thinking, however, also engages an intuitive non-linguistic number sense, the Approximate Number System (ANS). The ANS allows us to discriminate quantities, approximate arithmetic transformations, and estimate quantities, all without counting individual items. Although the ANS does not require language, estimations made by means of the ANS can be expressed with number words or Arabic digits. A connection between the ANS and school math performance has been established. A child's accuracy in mapping from approximate quantities to Arabic digits is associated with children's symbolic math abilities and can also predict their success at learning new arithmetic skills. Here, we explore whether directly training the mapping between estimated quantities and Arabic digits transfers to better math proficiency. The control training was based on discriminating quantity representations, without involving digits. Each of these three-week computer-based trainings were added to the school schedule. We measured improvements in approximate and exact arithmetic after training. Both the experimental and the control group improved in approximate arithmetic performance. However, in exact arithmetic, results show that strengthening the digit-quantity relation improved the 7-year-olds' competence in symbolic additions and subtractions over and above the improvement measured in the control group. Our results speak to the complexity of the factors involved in developing mathematical abilities, making the case that training the mapping from estimated quantities to digits can be particularly effective in improving children's mathematical performance.

Keywords

approximate number system, Arabic digits, mapping, mathematics education, numerical cognition, symbolic and nonsymbolic, arithmetic

The construction of mathematical competence is a complex challenge which we still do not fully understand, although much progress has been made in the last 20 years to reveal the complexity of the mental processes involved. The translation of such scientific understanding into practical ways to improve mathematical skills during the crucial ages when basic mathematical abilities are being acquired is still in its infancy. This is not surprising, considering that this acquisition may be affected by many factors, such as gender (Stoet, Bailey, Moore, & Geary, 2016), motivation (Simzar, Domina, & Tran, 2016), socioeconomic status (Langfus et al., 2019; Thien & Ong, 2015; Valle-Lisboa et al., 2016; Verdine et al., 2014), language development (Moll, Snowling, Göbel, & Hulme, 2015), math anxiety (Pletzer, Kronbichler, Nuerk, &

Kerschbaum, 2015), lack of predisposition to mathematics (Cerda et al., 2015), or the effects of teachers' biases (Demaray & Elliot, 1998; Tournaki, 2003). These are examples of important factors which contribute to success in mathematics at school. Another factor that may have a relation with mathematical performance is the intuitive sense of number, also called the Approximate Number System (ANS). The ANS is an approximate and non-linguistic sense that allows us to compare quantities, to approximate basic arithmetic operations, and to estimate, all without counting the items (Dehaene, 1997, 2001; Feigenson, Dehaene, & Spelke, 2004; Gallistel & Gelman, 1992; Odic & Starr, 2018). At the same time, mathematics is expressed in language, and its comprehension requires the understanding of the relations between quantities and the symbols we use to express them. In the present study we focused on training the estimation of quantities and its mapping to their linguistic counterparts, Arabic digits (Carey, 2009; Odic & Starr, 2018). The aim was to foster the comprehension of the relation between a digit and the quantity it represents, with the goal of enhancing arithmetic skills in children.

Many studies show that the ANS is present in non-linguistic animals as well as in preverbal infants and adults from different cultures (Brannon, 2006; Cantlon, Platt, & Brannon, 2009; Feigenson et al., 2004; Libertus & Brannon, 2009; Pica, Lemer, Izard, & Dehaene, 2004; Xu & Spelke, 2000). Although the ANS and its innateness are a subject of debate (Leibovich, Katzin, Harel, & Henik, 2017), according to some studies the ability to approximately estimate a quantity of items appears to be present at birth (Coubart, Izard, Spelke, Marie, & Streri, 2014; Izard, Sann, Spelke, & Streri, 2009) or shortly after birth (Xu & Spelke, 2000). Moreover, results suggest that these early abilities are continuous with later abilities to estimate and discriminate quantities throughout childhood and across the lifespan (Halberda & Feigenson, 2008; Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Libertus & Brannon, 2009).

When processing non-symbolic representations, two phenomena that are hallmarks of ANS engagement emerge: the distance effect and the size effect (Barth, Kanwisher, & Spelke, 2003; Cordes, Gelman, Gallistel, & Whalen, 2001; Dehaene, 1993; Dehaene, Dehaene-Lambertz, & Cohen, 1998; van Oeffelen & Vos, 1982). The distance effect is the phenomenon that the ability to discriminate between two quantities improves as the numerical distance between them increases. The size effect is the phenomenon that, for equal numerical distance, the ability to discriminate between two quantities improves as their numerical size decreases. These two phenomena are also observed when we process symbolic representations (Dehaene & Akhavein, 1995; Dehaene, Dupoux, & Mehler, 1990; Gilmore, McCarthy, & Spelke, 2007; Holloway & Ansari, 2008; Moyer & Landauer, 1967; Temple & Posner, 1998), and can be seen in neural activity when we discriminate digits (Núñez-Peña & Suárez-Pellicioni, 2014; Temple & Posner, 1998) or number words (Lussier & Cantlon, 2017). These results suggest that the ANS may be involved when symbolic numerical representations are processed, in accord with a format-independent coding of numbers (Butterworth & Walsh, 2011; Dehaene, 1992; Dehaene et al., 1990; Dehaene, Piazza, Pinel, & Cohen, 2003; Hinrichs, Yurko, & Hu, 1981; Moyer & Landauer, 1967; Piazza, Pinel, Le Bihan, & Dehaene, 2007; Pinel, Dehaene, Rivière, & LeBihan, 2001). However, other studies attest to a format-dependent coding of numbers (Bulthé, De Smedt, & Op de Beeck, 2014; Lyons, Ansari, & Beilock, 2012; Piazza & Eger, 2016; Shum et al., 2013). For sure, how the brain codes numbers is a complex issue that still needs to be fully investigated. The fact remains that the distance and size effects have been observed behaviorally for both symbolic and non-symbolic representations (Bulthé et al., 2014; Dehaene et al., 1998; Gilmore et al., 2007; Holloway & Ansari, 2009; Schneider et al., 2017; Zbrodoff & Logan, 2005).

Many studies have revealed a connection between the ANS and school math performance (Amalric & Dehaene, 2016; Feigenson, Libertus, & Halberda, 2013; Halberda & Feigenson, 2008; Halberda et al., 2012; Halberda, Mazzocco, & Feigenson, 2008; He et al., 2016; Libertus, Feigenson, & Halberda, 2013a; Mazzocco, Feigenson, & Halberda, 2011a, 2011b; Shusterman, Slusser, Halberda, & Odic, 2016; Starr, Libertus, & Brannon, 2013b). Although this relation is not always found (Butterworth, 2010; Libertus, Feigenson, & Halberda, 2013b), several meta-analyses support it (Chen & Li, 2014; Fazio, Bailey, Thompson, & Siegler, 2014; Schneider et al., 2017). This variation in results may be due to the fact that the relation between the ANS and school mathematics may be affected by several factors. Among them, some may be particularly relevant. First, the development of the relation between the ANS and mathematical language may be non-linear across childhood (Purpura & Logan, 2015). Second, math anxiety can induce fluctuations; indeed, the ANS may be less precise in high math-anxious individuals (Núñez-Peña & Suárez-Pellicioni, 2014). Third, the differences in school math scores; for example, the correlation between ANS precision and mathematical competence is stronger in children with lower mathematical scores (Bonny & Lourenco, 2013); furthermore, some studies showed that training

the ANS did not benefit every child equally, but tended to be more beneficial to children with low math scores (Ferres-Forga & Halberda, 2020; Langfus et al., 2019; Szkudlarek & Brannon, 2018). Fourth, different ANS abilities used as outcome measures may lead to different results: quantity comparison/discrimination (Halberda et al., 2012, 2008; Lyons & Beilock, 2011; Odic et al., 2016; Pinheiro-Chagas et al., 2014), approximate arithmetic (Gilmore, McCarthy, & Spelke, 2010; Pinheiro-Chagas et al., 2014), or quantity estimation (Pinheiro-Chagas et al., 2014; Siegler & Booth, 2004). Fifth, different ANS abilities used in training may lead to different results: several researchers trained quantity comparison/discrimination (DeWind & Brannon, 2012; Ferres-Forga & Halberda, 2020; Hyde, Khanum, & Spelke, 2014 *in training task "c"*); while others trained approximate arithmetic (Au, Jaeggi, & Buschkuehl, 2018; Hyde et al., 2014 *in training task "a"*; Park, Bermudez, Roberts, & Brannon, 2016; Park & Brannon, 2013, 2014; Szkudlarek & Brannon, 2018; *see* Szkudlarek, Park, & Brannon, 2021 *for a failure to replicate in adults*; *and see* Bugden, Szkudlarek, & Brannon, 2021 *for a training failure in 8-10 year olds*); and yet others trained quantity estimation (Sella, Tressoldi, Lucangeli, & Zorzi, 2016, *in Pre-syntactic subscale*). Lastly, the aspect of math that is being measured is also crucial; indeed some aspects appear to be more related to ANS than others. For example, informal math seems to be more related to the ANS than formal math (Libertus et al., 2013b; Szkudlarek & Brannon, 2018), and approximate arithmetic seems to be more related to the ANS than exact arithmetic (Ferres-Forga & Halberda, 2020). In Ferres-Forga and Halberda (2020), who trained the ANS using a quantity discrimination task, training benefits were visible only in their Operations Test, consisting of symbolic approximate arithmetic equations, which did not require exact calculations (Ferres-Forga & Halberda, 2020; Figure 1c). In these, the child had to mark the operation sign corresponding to the result of the equations (e.g., 8 □ 6 = 2). With this format, by estimating whether the solution number on the right side of the equal sign was slightly larger, much larger or smaller than the operator numbers on the left side of the equal sign, children could decide if the correct operation sign was, respectively, an addition (5 □ 3 = 8), a multiplication (4 □ 5 = 20) or a subtraction (8 □ 6 = 2). In this way, the Operations Test was sensitive to children's approximate arithmetic level.

Although the ANS has been related to mathematical achievement across many studies, as reported above, transfer to arithmetic after training has returned inconsistent results, depending on the trained ANS ability. For example, training quantity discrimination leads to small (Ferres-Forga & Halberda, 2020; Langfus et al., 2019) or no (Park & Brannon, 2014) benefits to arithmetic. Training approximate arithmetic gives inconsistent results in adults (Park & Brannon, 2013; Szkudlarek et al., 2021) and in third-to-fourth graders (Bugden et al., 2021), and provides little benefits in preschoolers (Szkudlarek & Brannon, 2018). We argue that these limited advantages depend on the fact that the contribution of the ANS is not sufficient to explain exact arithmetic skills nor other advanced mathematical abilities. These require a symbolic and accurate mathematical language that offers a precise representation of quantities, such as Arabic digits (Bonny & Lourenco, 2013; Dehaene, 2001; Gordon, 2004; Lemer, Dehaene, Spelke, & Cohen, 2003; McCrink, Spelke, Dehaene, & Pica, 2013; Pica et al., 2004). Although the ANS allows us to roughly estimate if there are many or few items in a collection, we often need to express these estimations accurately with a number word, or an Arabic digit. Indeed, estimating numerosity requires translating a non-symbolic quantity into a number (Siegler & Booth, 2005). To this end, we create a linguistic mapping of non-symbolic quantities onto symbols. Therefore, even the simple ability of activating such mapping requires knowledge of an arbitrary numerical symbol system -- knowledge that goes beyond the number sense. Nevertheless, training interventions based on the mapping from an estimated quantity to a number have been poorly studied. Furthermore, when they have, the focus has been on how ANS accuracy itself can be trained, instead of how such training could lead to potential transfers to arithmetic- (Russo, MacDonald, & Russo, 2021). In other cases, training of the mapping between the ANS and symbolic representations was administered together with other numerical training techniques, and limited to children with learning difficulties or low socioeconomic status (Räsänen, Salminen, Wilson, Aunio, & Dehaene, 2009; Sella et al., 2016; Wilson, Dehaene, Dubois, & Fayol, 2009; Wilson, Revkin, Cohen, Cohen, & Dehaene, 2006).

Mapping from quantity to numbers, that is, expressing an estimated quantity in mathematical language, can be more or less accurate. Importantly, the accuracy of this mapping relates to children's math abilities, both in number word format (Libertus, Odic, Feigenson, & Halberda, 2016; Marinova, Reynvoet, & Sasanguie, 2021; Mazzocco et al., 2011a) and, mainly, in Arabic digit format (Booth & Siegler, 2006; Brankaer, Ghesquière, & De Smedt, 2014; Holloway & Ansari, 2009; Mundy & Gilmore, 2009). Several studies show that, in 6-8-year-olds, a more accurate mapping between nonsymbolic numerical representations and Arabic digits is related to mathematical achievement (Booth & Siegler, 2006;

Brankaer, Ghesquière, & De Smedt, 2014; Holloway & Ansari, 2009; Mundy & Gilmore, 2009). Moreover, Booth and Siegler (2008) found that not only was the quantity-digit mapping related to 6-7-year-old's mathematical performance, but it also predicted their ability to learn new arithmetic skills. Indeed, it has recently been shown that Arabic numeral knowledge at 4 years of age is the sole independent predictor of arithmetic skills at 6 years, and is also a strong predictor at 5 years (Habermann, Donlan, Göbel, & Hulme, 2020). This study extends the findings on the importance of knowing the numerical meaning of Arabic digits for children's arithmetic development (Göbel, Watson, Lervåg, & Hulme, 2014; Sasanguie, De Smedt, Defever, & Reynvoet, 2012). Similarly, symbolic and mapping skills were found to be important predictors for math performance in 4- to 6-year-olds (Kolkman, Kroesbergen, & Leseman, 2013), in early grades (Lyons, Price, Vaessen, Blomert, & Ansari, 2014; Marinova et al., 2021) and in Grades 1 to 5 (Moore & Ashcraft, 2015). Even in adults, Castronovo and Göbel (2012) found that the ability to accurately map non-symbolic and symbolic quantities was significantly better in high-level math-educated adults. In line with these studies, difficulty in the acquisition of the meanings of number symbols may offer a clue to the presence and etiology of mathematics learning disabilities in children (De Smedt & Gilmore, 2011; Noël & Rousselle, 2011; Rousselle & Noël, 2007), and training this mapping may improve mathematics achievement in children at risk for developing difficulties in math (Tobia, Bonifacci, & Marzocchi, 2021). Also at a larger age, as shown by Mazzocco et al. (2011a), 14- and 15-year olds with mathematical learning disabilities are impaired in their mapping ability from non-symbolic representations to number words.

Mapping between non-symbolic and symbolic numerical representations is an ability that children develop over time (Brankaer et al., 2014; Mundy & Gilmore, 2009), along with accuracy, which also improves (Booth & Siegler, 2006) during development. Likewise, the mapping becomes more robust as children learn the cardinal meanings of number words (Carey, 2004; Le Corre & Carey, 2007; Odic, Le Corre, & Halberda, 2015; Slusser, Ditta, & Sarnecka, 2013; Slusser & Sarnecka, 2011; Wynn, 1990, 1992), even before they master the verbal count list (Pinhas, Donohue, Woldorff, & Brannon, 2014; Wagner & Johnson, 2011). And, learning the meanings of number words may correspond with an increase in ANS precision (Shusterman et al., 2016). While the early stages of how the mapping between the approximation of sets and number words is established remain to be fully articulated (Marinova et al., 2021), it seems that by the age of 6, children have formed a functional mapping between them, although the precision and biases involved will likely continue to undergo development even into the adult years (Izard & Dehaene, 2008; Siegel, Goldsmith, & Madson, 1982; Sullivan & Barner, 2014).

Mapping can be trained (Russo et al., 2021; Tobia et al., 2021). Importantly, calibration by means of informative trials can improve the accuracy of the estimations. For example, humans have a tendency to underestimate the number of items in a collection (Hollingsworth, Simmons, Coates, & Cross, 1991; Kemp, 1984; Krueger, 1982; Revkin, Piazza, Izard, Cohen, & Dehaene, 2008) but are able to renormalize their estimates based on feedback, as showed by Izard and Dehaene (2008). They found that adults, when exposed to a few trials in which they were informed of the number of dots on display, improved their accuracy in subsequent estimations. Likewise, Laski and Siegler (2007) showed that giving categorical information to calibrate "very small", "small", "medium", "big" and "very big" numbers promoted linear and accurate estimation in kindergartners.

As reviewed above, a more accurate mapping between estimated quantities and Arabic digits is related to higher mathematical achievement. Mapping develops over time, and experimental results suggest that its accuracy can be calibrated and trained. However, the development of these findings into concrete training programs to help children has yet to be robustly explored. In the current study, we present a training regime of the estimation of quantities and their mapping to Arabic digits in 7-year-old children during a 3-week period. Our aim is to assess change and transfer from such training to improvements in arithmetic. We named this training regime Numerical Estimation Training (NET). Because the accuracy of the estimations can be improved by calibrating with informative trials (Izard & Dehaene, 2008; Laski & Siegler, 2007), we decided to include such trials in this training, which we call *passive learning trials*.

Our control group replicated the Quantity Discrimination Training introduced by Ferres-Forga and Halberda (2020) (hereafter, QDT). In that study, the training was based on the ANS component of quantity discrimination: the ability to determine which of two collections is greater in quantity without counting their items. The training yielded some benefits compared to a business as usual control group in which no ANS activity or other numerical tasks were trained. Such improvement was found in approximate arithmetic, but not in exact arithmetic, and only in children with low-starting arithmetic scores. The choice of replicating the QDT for our control group, instead of using a business

as usual approach, allowed us to determine if training 7-8 year old children with the NET could transfer into an improvement in arithmetic over and above the potential improvement obtained by the QDT. Therefore, we administered a training regime to the control group that was known to generate some advantages, within the limits indicated above, and hence was potentially positive for these participants as well. This design allowed us to better estimate the relevance of the advantage that the NET may generate.

The study comprised pre- and post-training tests and an intervention of 3 weeks of training. For pre- and post-testing, we used the exact Arithmetic test and the Operations test used in Ferres-Forga and Halberda (2020), because they were apt to measure exact and approximate arithmetic respectively, as previously explained.

We also considered the behavioral presence of distance and size effects for both symbolic and non-symbolic representations, as discussed above (Bulthé et al., 2014; Dehaene et al., 1998; Gilmore et al., 2007; Holloway & Ansari, 2009; Schneider et al., 2017; Zbrodoff & Logan, 2005). Thus, in the present study we measured both phenomena in NET. Regarding the QDT, in Ferres-Forga and Halberda (2020), they measured the ratio effect (the closer the ratio of the compared sets of items is to 1, the more difficult is to discriminate the sets of items) which is conceptually the distance and size effects combined and is typically the measure of performance in non-symbolic comparison tasks. Therefore, in the current study we measured the ratio-effect for the QDT.

In summary, we replicated both the QDT (here, the training for the control group) and the pre- and post-math tests as in Ferres-Forga and Halberda (2020), whereas the experimental group was trained with NET.

The main motivation of this study is the suggestion in the literature that in the learning of mathematics, strengthening the mapping between nonsymbolic and symbolic representations of number is a key factor that is often overlooked. We explore the hypothesis that training children to map estimated quantities to their corresponding digits can better calibrate the quantitative meaning of numbers and result in an overall better understanding of the digit-quantity relation, with the potential of manifesting in improvements in their arithmetic abilities.

Method

Participants

Ninety-one children from the second grade of primary (38 girls; average age = 7 yrs 9 mos, range = 6 yrs 4 mos - 8 yrs 9 mos) participated in the study. The children mostly came from middle-to-high socioeconomic status families. The study was conducted at the Hamelin International Laie School (http://www.hamelininternacionallaie.com/school/), at the premises of the school. The study took place in the same school and during the same month but one year after the study of Ferres-Forga and Halberda (2020), with different participants. This offered the best control of external factors, notably of math level during school year, for comparison with Ferres-Forga and Halberda (2020).

Participants attended four different classrooms, taught by two different teachers. Two classes were randomly assigned to the Numerical Estimation Training group (experimental; $n = 45$; 18 girls; average age = 7 yrs 10 mos), and the other two classes to the Quantity Discrimination Training group (control; $n = 46$; 20 girls; average age = 7 yrs 9 mos), with the constraint that the assignment was counterbalanced across the two teachers. We explain the differences in training below. For the NET group, all participants completed the full training, except for one child who completed the initial test but did not complete the whole training and the final assessment, and was excluded from analysis. All participants in the QDT group completed the total of 24 runs over the course of three weeks, except for six participants who completed 21 runs, one who completed 22 runs and one who completed 20. Considering the minimal amount of training that they lost, these children were kept in the data analysis.

Materials

Mathematical Competence Assessment

In order to determine participants' mathematical competence, we administered the three pencil-and-paper tests administered by Ferres-Forga and Halberda (2020). They were contained in three separate test booklets: an Additions Test (Figure 1a); a Subtractions Test (Figure 1b); and an Operations Test (Figure 1c). When working with the Additions or

Subtractions Tests, children had to write the exact answer to the addition or subtraction problems presented. These two tests were entirely composed of problems that children were accustomed to in their standard mathematical activity at school. The Operations Test contained problems of a novel kind, to which children had not been exposed to before. In them, each equation already contained the result of the computation, but the operation sign itself was omitted. Children had to write the operation sign required by the problem (an addition, a subtraction, or a multiplication) to answer correctly.

Two versions of each test were prepared. Each version contained different problems although, due to the limited amount of possible combinations to form the equations, a few of them were present in both versions. The order of the problems was randomized. In this way, both versions of the booklets could be used for the pre- and post-training tests, and could be counterbalanced across the sessions, so as to control for tests effects. We created a large number of problems for each test, so that the children could not complete all of the problems during the allotted time (6 minutes).

The Additions Test included 210 problems presented in a columnar operation algorithm form, printed on 10 pages. The maximum number that each addendum could reach was 18, with the highest sum being 18 + 18 and the lowest sum being 0 + 0 (Figure 1a).

Figure 1

Examples of Problems in Additions Test (a), Subtractions Test (b) and Operations Test (c)
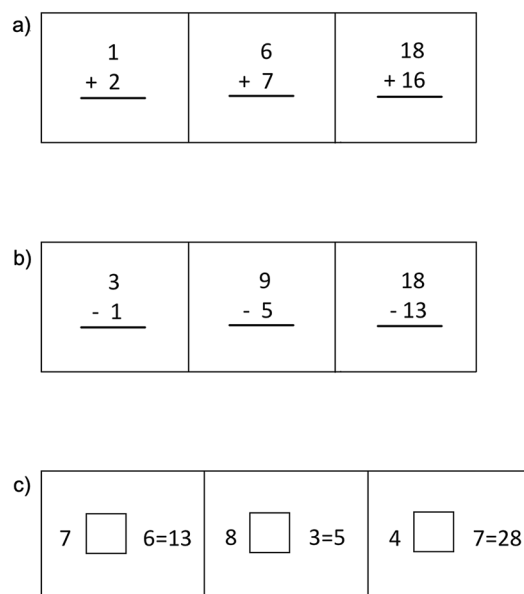
The Subtractions Test included 190 problems presented in a columnar operation algorithm form, printed on 10 pages. Both the minuend and subtrahend ranged between 0 and 18; the results of the subtractions were always positive (Figure 1b).

The Operations Test included 117 problems presented in horizontal format, printed on 3 pages (Figure 1c). The problems were equations whose unknown operation could be an addition, a subtraction or a multiplication. In the addition equations, the first addendum had a maximum value of 7 and the second of 10. In the subtraction equations, both the minuend and subtrahend ranged between 1 and 11, with the result being always positive. In the multiplication equations, only the timetables of 1 to 5 and of 10 were used, because those were the timetables children had been exposed to. The $2 \square 2 = 4$ problem was omitted due to possible confusion. The three types of operations (addition, subtraction, multiplication) and the problem difficulties were presented in random order. One of the versions of the operations test contained forty-two additions, thirty-eight subtractions and thirty-seven multiplications while the other version contained forty-three additions, thirty-nine subtractions and thirty-five multiplications.

It is important to note the different nature of the tests. Although all three tests require the previously acquired knowledge of what the basic arithmetic operations do, the Additions and Subtractions Tests also require the ability to make exact calculations, while the Operations Test only requires a comprehension of how the three arithmetic operations roughly change quantities. Stating it in a simple manner, the result is generally "more" for addition, "less" for subtraction, and "much more" for multiplication, allowing for a comparison between both sides of the equal sign.

Computer Activities

Twenty-eight computers (model: clon PCs, Intel(R) Core(TM) i3-4170 CPU @ 3.70 GHz, 4GB RAM, 64-bit; monitor: 17" LCD 16/9 from ASUS; operative system: Windows 7 Professional) were used for the training activities. The children wore headphones during training. The stimuli that formed the collections of items for the training part were the same (blue dots, yellow dots, cars, bears, birds, dogs...) in both training groups. In order to maintain children's interest, these items varied along the trials. Both training regimes had the same number of runs and trials: 24 runs with each composed of 35 consecutive trials. In each trial, the stimuli appearing on-screen remained visible too short of a time for children to count the items. This was meant to engage the ANS in the task. Trials were presented in increasing order of difficulty (details below), a procedure which is known to facilitate learning (Odic, Hock, & Halberda, 2014; Wang, Odic, Halberda, & Feigenson, 2016). Participants received a feedback sound of ≈ 330 ms, with a high-pitch beep for correct answers and a low-pitch beep for incorrect answers. After the feedback, a screen with the indication to press the space bar to continue appeared, and the successive trial started when the child pressed the bar.

Experimental Group: Numerical Estimation Training (NET) — Participants in this group were presented with the "Digits" game, written in PsychoPy v1.83.01. The Digits game program generated two types of trials: the *passive learning trials* and the *active training trials*. Within each run of 35 consecutive trials, the first seven trials were passive learning trials and the next 28 were active training trials. In the passive learning trials, which lasted 1200 ms, a collection of items was presented on-screen while a prerecorded voice named the exact number of items (Figure 2a). No action was needed on the part of the participant. In the active training trials (Figure 2a), a collection of items was first presented for 1 s in silence, so that its numerosity could be estimated. Then, the collection disappeared and three digits were presented on the screen (Figure 2b). These three digits remained on-screen until the participant indicated their numerical estimation by clicking the mouse on one of them, which immediately triggered the feedback sound. In both types of trials, the collections of items ranged from 1 to 21, all items with the same size and orientation.

Figure 2

Digits Game
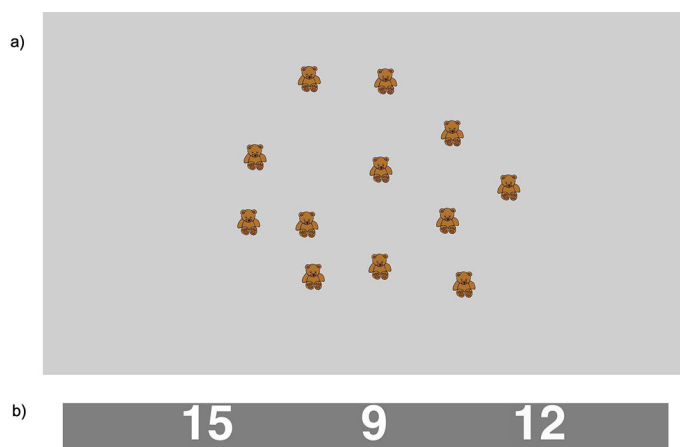
Note. (a) Example of a learning trial. This display could appear as either a passive trial (accompanied by a spoken numerical label, such as "twelve") or as an active learning trial (with no auditory information given). (b) Response screen for the active learning trials. In active learning trials, the collection of items (a) would appear for 1 s, in silence. Afterwards, the collection would disappear and the answer options (b) would be presented on-screen until the participant made their choice by clicking on one of them.

The purpose of the passive learning trials was to provide children with opportunities to directly calibrate their estimation system before the active training trials began (Izard & Dehaene, 2008; Krueger, 1989). This was particularly important because active trials presented three choices and, therefore, feedback for a wrong choice could not indicate which of the two remaining choices was correct. Thus, feedback provided during active trials could only serve as partial guidance. In this regard, passive trials could compensate for this partial guidance before the active trials started. The passive trials were presented in decreasing order of difficulty (from bigger to smaller sets of items), so that the easier passive learning trials were always presented last. This increased children's confidence before the active learning trials started. Although we could not control automatically if participants were attentive to the passive trials, because no action was required, these trials were placed at the beginning of each run to help children to calibrate before the active trials.

In the NET, there was a range of distances across trials from the target value to the distractor values. For example, if the target number was 15, a decision between 18, 21 and 15 (i.e., +3, +6, and 0 distances from the correct choice) would be easier than the decision between 15, 13 and 17 (0,-2,+2) because of the differences in distance between the target value (15) and the distractor values. We defined three difficulty levels based on the range that the trials could have. For *Span 6*, the easiest one, distances could be (0,3,6), (-6,-3,0) or (-3,0,3). For *Span 4*, distances could be (0,2,4), (-4,-2,0) or (-2,0,2). Finally, for *Span 2*, the hardest one, distances could be (0,1,2), (-2,-1,0) or (-1,0,1). By manipulating the value of the target and the distances of the choices, difficulty was increased within each run and throughout the session. Within each run, increasing the value of the targets would increase the difficulty. Additionally, within each training session, decreasing the *Span*, so that the distances between the correct digit and the other two distractor digits decreased at each run, would also increase difficulty.

All set sizes, from 1 item to 21 items, had to be estimated in each of the three spans. The distances between possible answers were maintained irrespective of the target answer (rather than scaling the distractor answers relative to the correct answer by some ratio). Thus, trial difficulty increased with target answer (Figure 4). The position of the correct digit (left, middle, right) as well as its numerical relation with respect to the other two digits (the smallest, the middle one, the largest) were balanced across trials.

Control Group: Quantity Discrimination Training (QDT) — For this group, we used a modified version of the computer game Panamath (Halberda et al., 2008), written in Java SE6. In this version, two collections of items inside two rectangles would appear to the right and the left of the screen. For example, 12 teddy bears could appear inside the left rectangle and 6 blue dots inside the right rectangle. The number of items within each rectangle was always between 5 and 21. On each trial, the two sets of items were displayed simultaneously for 1382 ms. After this interval, the screen turned blank until an answer was typed. The game consisted in choosing which of the two sets on either side of the screen had more items (Figure 3a-c), and typing the answer on the keyboard ("f" and "j" keys for left or right side respectively). If participants took longer than ≈1.3 seconds to answer, the objects disappeared and the screen remained blank until they typed their choice. Their answer immediately triggered the feedback sound.

The items were presented in seven different ratios (larger set ÷ smaller set). The ratios could be 3, 2, 1.5, 1.25, 1.17, 1.14, and 1.1. For example, on a 3-ratio trial children might see 21 blue dots on the right side of the screen and 7 yellow dots on the left side. Smaller ratios correspond to more difficult trials. Always, the first five trials of each run presented the easiest ratio. Then, every five trials the game increased in difficulty, with the ratios becoming closer to 1 (without ever reaching 1), until the seven different ratios were presented. This method of presentation uses "confidence hysteresis" and tends to return best possible performance and successful transfer to symbolic mathematics (Odic et al., 2014; Wang et al., 2016).

To vary the relationship between surface area and number, the Panamath game implemented three different models controlling for object size: size-confounded (42% of the trials in each run; Figure 3a), size-controlled (42% of the trials; Figure 3b), and stochastic size-control (16% of the trials; Figure 3c). In size-confounded trials, the average size of the items was equal for both sets, so that the cumulative surface area occupied by the objects was congruent with the number of objects. In size-controlled trials, the average size of the objects was smaller for the larger set, so that the ratio of the cumulative area occupied by the objects in each set was equated. In stochastic size-control trials, object sizes were stochastically varied to give children no consistent size cue for number. In these, the average size of the objects varied

randomly between size-anti-correlated pictures (in which the numerically larger set occupied less total area on screen), size-controlled pictures, and size-confounded pictures (where the numerically larger set occupied more total area on the screen).

Figure 3

Examples of the Three Models of Object Size Controlling for Surface Area of QDT Trials
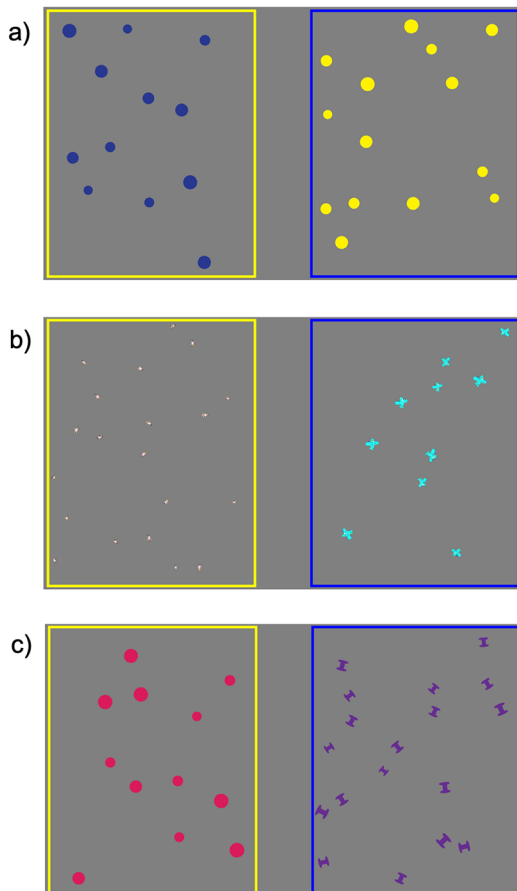
Note. (a) size-confounded; (b) size-controlled; (c) stochastic size-controlled.

Passive learning trials, which have a quantity-to-digit calibrating role, were not included in the QDT. This is because the QDT trains the comparison of quantities without any symbolic representation, and this renders such calibration antithetical. Moreover, because the trials always presented a binary choice, the two-sound feedback provided complete guidance for the correct answer. Therefore, in the QDT, all trials are *active learning trials*. The total number of trials presented was the same as that of the NET.

Procedure

The experiment was run in three phases: *Pre-training Assessment*, *Training*, and *Post-training Assessment*. The teachers and the experimenter did not mention any relation between the computer training and the assessments, or mathematics in general, to the students. To further separate these, the training and assessments were run at differently scheduled times and in different classrooms.

Pre-Training and Post-Training Assessments

The pre- and post- assessments were intended to measure the mathematical competence of the participants before and after training. We followed the same procedure as was applied in Ferres-Forga and Halberda (2020). The Pre-training assessment was administered three days before the first training session. The Post-training assessment was administered three days after training completion. These assessments were conducted in the children's respective classrooms, in the sole presence of their math teacher. Each of the Pre-training and Post-training tests required children to solve as many equations as possible in 6 minutes. All tests contained more problems than could be solved within the 6-minute interval, so that potential speed and accuracy improvements induced by the training could be assessed by comparing the number of correct answers in Pre- and Post-tests.

The assessments always began with the Additions Test. After completing this 6-minute test, children had to stop answering, return the additions booklet, and wait until they were given the subtractions booklet. The Subtractions Test followed, with the same 6-minute procedure. For the third and final test, the Operations Test, the teacher explained the task to the children in more details, given that this kind of problem was new to them. The teacher briefly described the equation structures and completed 3 examples on the board (one each for addition, subtraction and multiplication) in front of the class. Then children began the 6-minute assessment of the Operations Test.

The Pre- and Post-Training assessments were completed in the students' regular classrooms, during their regular math class time. The teacher and the experimenter carefully avoided drawing any attention to possible connections between the training schedule and the Pre- and Post- assessments.

Training Phase

Training was administered in six different sessions within a three-week period, at a pace of two sessions per week. Both groups were trained for the same amount of time and with the same number of trials. Each session took approximately 25-30 minutes to complete. The training sessions were integrated into the school schedule of each class, during their computer class-time. Because of this procedure, any effect due to potential deviations of the experiments from class routine was minimized. Training sessions for both conditions were carried out in the same computer classroom and always in the presence of the computer class teacher and the experimenter. Before each training session the experimenter prepared the computer classroom. All material was checked by the experimenter, including the volume of the headphones. At each training session, children worked individually after receiving instructions. They were informed that both speed and accuracy were important. Children were also told that the game would vary in difficulty. They were informed that two different sounds would provide them with feedback about the correctness of their answer, and that this feedback would be provided after every response. Lastly, it was explained to them how to answer. Children of both groups appeared to like the games of their training sessions.

Experimental Group: Numerical Estimation Training (NET) — To train children's ability to map from estimated quantities to digits, children belonging to the Numerical Estimation Training group practiced the Digits game. When introduced to the game, children were told they would play a game where they would first see a collection of objects for a very short time, while an audio recording would tell them how many objects were in the collection. They were told that they would have to pay attention to these trials, but not take any action. They were also informed that they would then see many trials where a collection would be shown for a short period of time, after which they would have to estimate their numerosity by choosing one of three digits that would appear on-screen immediately after the disappearance of the items.

Control Group: Quantity Discrimination Training (QDT) — In our control group children practiced the Panamath quantity discrimination game, replicating the procedure of the experimental group in Ferres-Forga and Halberda (2020). This game does not require any understanding of the relation between digits and quantities.

When introducing this QDT game, children were told that they would play a game where they would see two sets of objects on two sides of the screen – for example, blue dots on the left and yellow dots on the right – and that they would have to choose the set that had more items.

Results

We first present the results of the ANS hallmarks and the efficiency of the training sessions for each group. We then turn to the main results of interest, that is, the effects of the training on symbolic mathematics.

Experimental Group: Numerical Estimation Training (NET)

In the NET task, children viewed a quantity of briefly flashed items and had to estimate the amount by choosing the correct match among three possible digits. On average, children responded correctly on 51.3% of trials (SD = 6.3%; chance level = 33%). Participants' accuracy was above chance for all runs at all three levels of difficulty (*Spans 6, 4,* and *2*; Figure 4) as revealed by planned *t*-tests: Easy, $t(43)$ = 19.1, p < .001; Medium, $t(43)$ = 20.1, p < .001; Hard, $t(43)$ = 12.7, p < .001.

Figure 4

Percent Correct on the Numerical Estimation Training Task as a Function of the Number of Objects of the Collections (Size Effect)
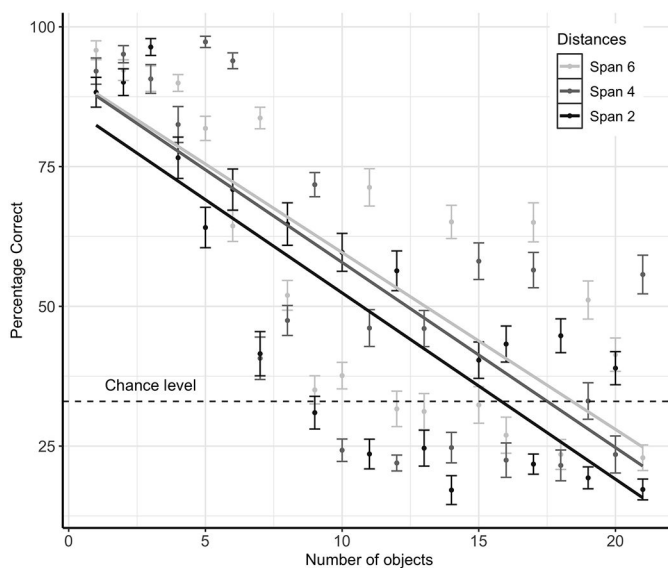

Note. The three linear regression lines refer to the distance levels of difficulty (distance effect): Easy (Span 6), Medium (Span 4), Hard (Span 2). The dashed horizontal line represents random performance (chance level = 33%).

Consistent with predictions of the distance effect, children's responses were better as the runs were easier (percentage of correct answers on average: M_{Easy} = 58.23%, SD = 1.32% and M_{Medium} = 50.86%, SD = 0.89%, M_{Hard} = 44.86%, SD = 0.93%. A repeated measures ANOVA, with Difficulty (distance effect) as an independent factor and participants' percent correct as the dependent variable revealed a main effect of Difficulty, $F(2,86)$ = 139.9, p < .001, η^2 = 0.38. Bonferroni *post hoc* tests revealed differences between the Easy runs and the Medium runs (p < .001), the Easy runs and the Hard runs (p < .001) and the Medium runs and the Hard runs (p < .001). The distance effect is also accompanied by a size effect. The manipulation of the size of the collections presented within each run, through incrementing the number of its items, allowed us to analyze the size effect on children's performance. As expected, for equal *Spans*, performance decreased as the quantity of items to estimate increased (Figure 4). The percentage of correct answers was significantly correlated with the number of items presented, r = -.91, p < .001. A simple regression analysis was used to test if the size of the collections significantly predicted participants' performance. The results of the regression indicated that the number of items accounted for the 82% of the variance in performance, R^2 = .82, $F(1,19)$ = 88.67, p < .001. Children's percentage of correct answers could be predicted from the size of the collections by the following formula: -3.38 × number of items + 90.23. In sum, both distance and size effects appeared during in the NET.

Next, we analyzed children's response time. Average RT was 2234.6 ms (SD = 615 ms; ternary choice trials). A logarithmic training slope was computed for each child. Children had significantly negative training slopes for response time across training sessions: $t(43)$ = -3.38, p = .002; Figure 5.

Figure 5

Response Time (RT) in the Numerical Estimation Training Task as a Function of the Progress in Training Runs (24 Runs in Six Sessions Across Three Weeks)
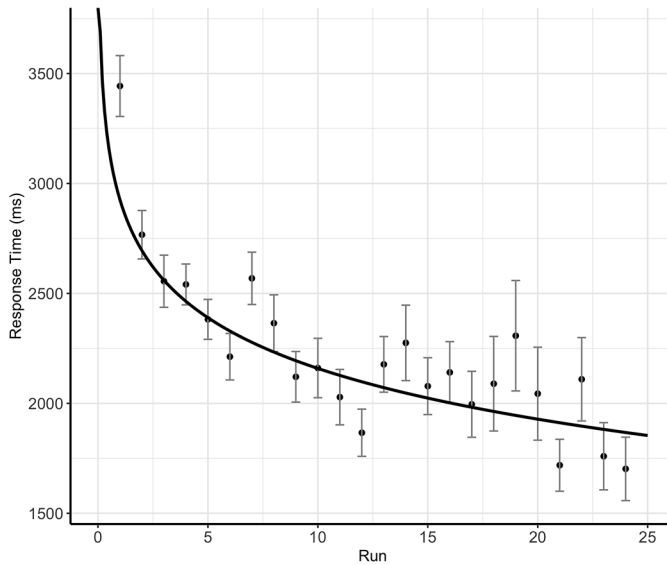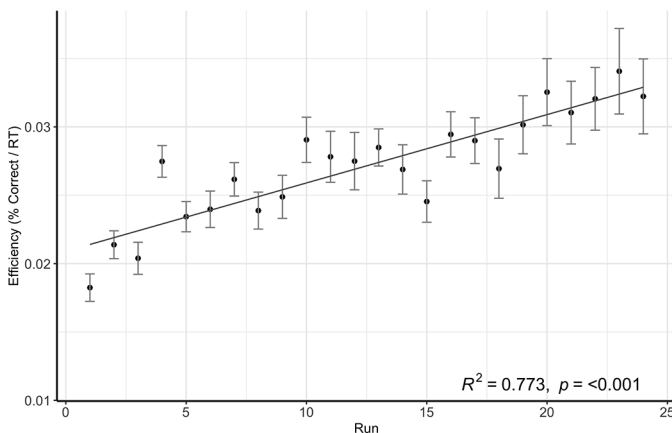

Note. Error bars represent ± *SE* for the group performance. The logarithmic regression line represents the average slope and intercept across participants (n = 44).

The efficiency, operationalized as the percentage of correct responses divided by RT, increased as children progressed throughout the sessions of the NET (Figure 6). A linear training slope was computed for each child. Children had significantly positive slopes for efficiency across training sessions, indicating that efficiency improved throughout training: $t(43)$ = 2.76, p = .008; Figure 6.

Figure 6

Efficiency in the Numerical Estimation Training Task as a Function of Training Runs

$R^2 = 0.773$, p = <0.001

Note. Error bars represent ± *SE* for the group performance at each run. The linear regression line represents the average slope and intercept across participants (n = 44).

## Control Group: Quantity Discrimination Training (QDT)

In the QDT, children briefly viewed two non-symbolic sets of items and had to choose the biggest set. On average, children responded correctly on 72.2% of trials (SD = 7.1%; chance level = 50%). Participants' performance can be modeled by a specific curve of percentage of correct answers as a function of the ratio between the two quantities. This ratio-dependence is predicted by Weber's law and is the main signature of the Approximate Number System (Feigenson et al., 2004; Halberda & Feigenson, 2008; Libertus & Brannon, 2009; Piazza & Izard, 2009; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Starr, Libertus, & Brannon, 2013a). Specifically, the percentage of correct answers increases as the ratio between the two approximate numerosities increases. We checked that performance on the QDT showed this signature (Figure 3). The ratio-dependent performance curve is observed for all three trial types of size control for the stimuli. That is, as the numerical ratio between the two collections became easier (e.g., ratio 3 versus ratio 1.2), children's percentage of correct responses improved, regardless of the type of size control for the trial. Also, children chose the numerically greater collection well above chance as shown by planned t-tests (size-confounded, $t(45)$ = 20.34, p < .001; size-controlled, $t(45)$ = 22.76, p < .001; stochastic size-controlled, $t(45)$ = 12.47, p < .001).

In all cases, children's performance exhibits the smooth curve of the Approximate Number System. That is, even if size contributes somewhat to children's decisions, children's numerical decisions were likely based on the ANS. The curves in Figure 7 are generated by fitting a model of Weber's law to the mean performance of children in each ratio for each size control type, with each child contributing equally to the curves.

Using the standard psychological model and fitting methods (e.g., Halberda et al., 2008; Odic et al., 2014) we estimated the ANS acuity (i.e., Weber fraction or w) for the numerical discrimination task on the first day for the whole control group. The best fit model returned a w of 0.284 (SEM = 0.064). Other studies have found similar Weber fraction; for example Halberda et al. (2008) found a w of 0.279 (SEM = 0.012) for 14-yr-old children. Quantitatively similar results for w were obtained in Piazza and Izard (2009) and in Halberda and Feigenson (2008). In Ferres-Forga and Halberda (2020), the experimental group, trained with the QDT, scored a w of 0.192 (SEM = 0.013). This value was slightly better (higher ANS acuity) than the one obtained here for our control group, which replicated the QDT.

Figure 7

Accuracy in the Quantity Discrimination Training Task as a Function of the Ratio Between Numerical Values
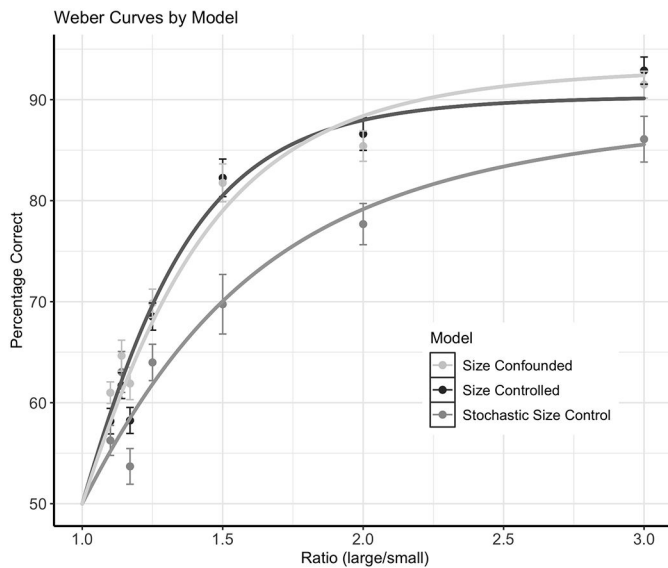

Note. The x-axis shows the seven different ratios presented during the training (1.1, 1.14, 1.17, 1.25, 1.5, 2, 3). The Weber curve of the three surface-area manipulations are plotted separately. Error bars represent ± *SE* for the group performance.

Children's average RT across training sessions was 952.8 ms (SD = 325.4 ms; binary choice trials). Figure 8 represents the mean RT for the group at each run (± SE). A logarithmic training slope was computed for each child. Children had significantly negative training slopes for response time across training sessions, $t(45)$ = -7.3, $p < .001$.

Figure 8

Response Time in the Quantity Discrimination Training Task as a Function of the Progress in Training Runs (24 Runs in Six Sessions Across Three Weeks)
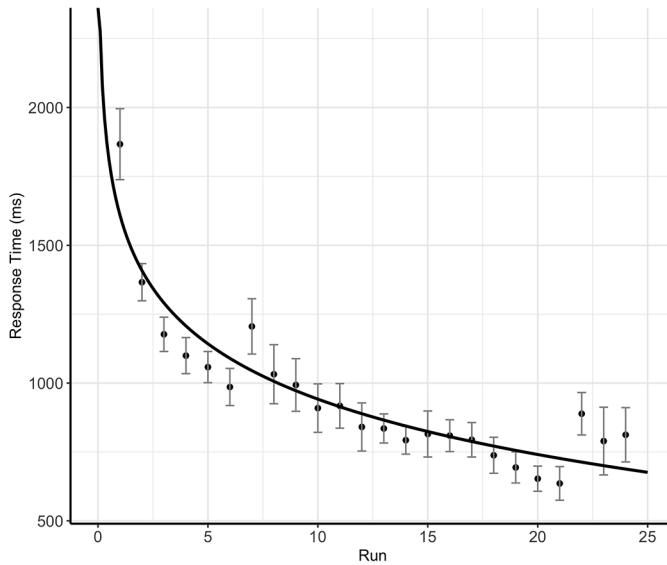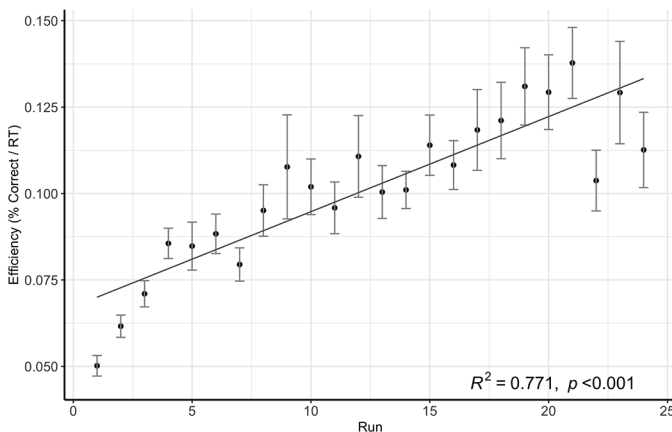

Note. Error bars represent ± SE for the group performance. The logarithmic regression line represents the average slope and intercept across participants (n = 46).

We analyzed the efficiency of the QDT, operationalized as the percentage of correct responses divided by the RT. A linear training slope was computed for each child. Children had significantly positive slopes for efficiency across training sessions, indicating efficiency improvement during the training, $t(45)$ = 7.3, $p < .001$; Figure 9.

Figure 9

Efficiency in the Quantity Discrimination Training Task as a Function of Training Runs

Note. Error bars represent ± SE for the group performance. The linear regression line represents the average slope and intercept across participants (n = 46).

In sum, both the NET and the QDT showed signatures of engaging the ANS and had the appropriate level of difficulty, as shown by the fact that children's task efficiency improved during the 3 weeks of training.

Effects of Training on Symbolic Mathematics

The main analysis of interest concerns the effect of training on children's mathematical performance, in the Additions, Subtractions and Operations Tests. A 2 x 2 mixed ANOVA with Training Condition (NET, QDT) as a between-participant factor, Phase (Pre-, Post-training) as a within-participant factor, and total correct answers as the dependent variable was run for each test. We report the results below.

Additions Test

For the Additions Test, the analysis revealed a main effect of Phase, $F_{(1,88)}$ = 10.03, p = .002, η^2 = 0.011, no effect of Training Condition, $F_{(1,88)}$ = 0.177, p = .67, η^2 = 0.001, and a significant Training Condition by Phase interaction, $F_{(1,88)}$ = 7.03, p = .009, η^2 = 0.008. Bonferroni *post hoc* tests revealed that this interaction was uniquely driven by the fact that, in the NET group, the total correct answers in the Post-training phase were higher than in the Pre-training phase (respectively, M_{Post} = 43.34, SD = 3.33 and M_{Pre} = 35.6, SD = 3.15, p < .001; Figure 10), while there were no improvement in the QDT group (respectively, M_{Post} = 37.74, SD = 3.52 and M_{Pre} = 37, SD = 2.68, p = .72; Figure 10). There was no difference in total correct answers between the groups in Pre-training (p = .63, Figure 10) nor in Post-training (p = .25, Figure 10). Thus, the NET group significantly improved additions performance while the QDT group did not.

Subtractions Test

For the Subtractions Test, the ANOVA revealed a main effect of Phase, $F(1,88)$ = 34.32, p < .0001, η^2 = 0.03, no effect of Training Condition, $F(1,88)$ = 3.72, p = .057, η^2 = 0.037, and a significant Training Condition by Phase interaction, $F(1,88)$ = 6.58, p = .012, η^2 = 0.006. With respect to the significant interaction, Bonferroni *post hoc* tests revealed that in both training conditions the total correct answers increased from Pre-training to Post-training. Specifically, in the NET group the total correct answers increased more than in the QDT group (for NET: M_{Post} = 47.81, SD = 3.86, M_{Pre} = 37.22, SD = 3.18, p < .0001; for QDT: M_{Post} = 35.98, SD = 3.26, M_{Pre} = 31.69, SD = 2.54, p = .027). While both training groups improved performance, the improvement was higher in the NET group: indeed, while there was no difference in total correct answers between the groups in Pre-training (p = .21; Figure 10), children in the NET group scored significantly higher than those in the QDT group in the Post-training (p = .021; Figure 10). Results show that the NET group significantly improved subtractions performance over and above the QDT group.

Operations Test

For the Operations Test, the ANOVA revealed a main effect of Phase, $F_{(1,88)}$ = 29.33, p < .001, η^2 = 0.04, no effect of Training Condition, $F_{(1,88)}$ = 0.74, p = .39, η^2 = 0.007, and a non-significant Training Condition by Phase interaction, $F_{(1,88)}$ = 0.10, p = .75, η^2 = 0.0001, (Figure 10). There was no difference in total correct answers between the groups in Pre-training (p = .39, Figure 10) nor in Post-training (p = .43, Figure 10). Results show that both groups had a similarly large improvement.

Figure 10

Total Correct Answers in the Symbolic Tests in the Pre- and Post-Training Phases, for the Numerical Estimation and Quantity Discrimination Training Conditions
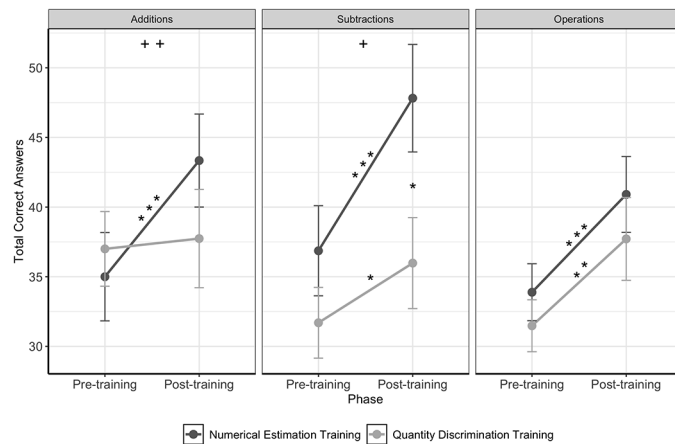
Note. Asterisks indicate significant post hoc tests. Plus signs indicate significant interactions, with number of symbols indicating the *p*-values of the tests.

$^{*}p < .05.$ $^{**}p < .01.$ $^{***}p < .001.$ $^{+}p < .05.$ $^{++}p < .01.$

Discussion

Many studies have revealed a connection between the ANS and school math performance (Chen & Li, 2014; Fazio, Bailey, Thompson, & Siegler, 2014; Schneider et al., 2017 for reviews). Other studies have shown the trainability of ANS abilities (Au et al., 2018; DeWind & Brannon, 2012; Ferres-Forga & Halberda, 2020; Hyde et al., 2014). However, transfer to improvements in arithmetic has been less consistent and depends on the ANS ability trained. When quantity discrimination was trained, the transfer to arithmetic abilities has been scarce (Ferres-Forga & Halberda, 2020) or none-at-all (Park & Brannon, 2014). When approximate arithmetic was trained, results have been inconsistent (Bugden et al., 2021; Park et al., 2016; Park & Brannon, 2013; Szkudlarek et al., 2021). Finally, training based on quantity estimation, which the current study focuses on, has been less studied. When it has been, it was mostly limited to children with learning difficulties or low socioeconomic status (Räsänen, Salminen, Wilson, Aunio, & Dehaene, 2009; Sella et al., 2016; Wilson, Dehaene, Dubois, & Fayol, 2009; Wilson, Revkin, Cohen, Cohen, & Dehaene, 2006).

The mapping task between non-symbolic and symbolic number representations focuses on the estimation of quantities (Brankaer et al., 2014; Castronovo & Göbel, 2012; Mundy & Gilmore, 2009), an ability that relies on the ANS. Indeed, in a mapping task, participants first make an estimation of the quantity of items and then they form a linguistic mapping to number words or Arabic digits (Carey, 2009; Odic & Starr, 2018). The accuracy of this mapping varies; it is related to children's math abilities (Booth & Siegler, 2006; Brankaer et al., 2014; Holloway & Ansari, 2009; Libertus et al., 2016; Marinova et al., 2021; Mazzocco et al., 2011a; Mundy & Gilmore, 2009), and can even be a predictor of children's arithmetic learning (Booth & Siegler, 2008; Göbel et al., 2014; Habermann et al., 2020; Kolkman et al., 2013; Lyons et al., 2014; Marinova et al., 2021; Sasanguie et al., 2012).

In the present study, we trained the mapping from estimated quantities to Arabic digits in 7-year-olds during a 3-week period, with the aim of transferring improvements to arithmetic. We introduced a novel training regime, the Numerical Estimation Training (NET), implemented in a computer intervention ("Digits" game). For our control group, we replicated the Quantity Discrimination Training regime (QDT) from Ferres-Forga and Halberda (2020), implemented with a version of the "Panamath" game. Both training tasks rely on ANS abilities – respectively, quantity estimation and quantity discrimination – but while the QDT does not involve any symbolic representations, the NET links the estimation of quantities to the symbolic language of mathematics. We considered the QDT a better control than a business as usual control ("non-numerical standard school activities"). This is because it is also a numerical cognition

training regime and its differences with the NET are minor compared to "non-numerical standard school activities". Due to the different ANS abilities trained by the NET and the QDT, the regimes differed in the inclusion of a calibration procedure with passive trials that also compensated for the partial feedback in the NET tasks (see *NET* and *QDT* in *Computer Activities*). Apart from these differences, both training regimes were equivalent in terms of procedure (*Training Phase*), employed computer programs with many similarities (*Computer Activities*), and were identically inserted in the normal school routine. The QDT regime and the pre and post math tests were replicated from Ferres-Forga and Halberda (2020) under strict conditions (i.e., same school and period of the school year), but with new children, hence math knowledge-level of children was very comparable if not identical across studies. This allowed us to also evaluate the current study on the basis of the results of Ferres-Forga and Halberda's previous training study. That previous study established that the QDT does provide some benefits in approximate arithmetic compared to a "non-numerical standard school activities" control group. Choosing the QDT regime for our control group was particularly challenging for the NET regime because it could potentially reduce detectable effects of the NET over the QTD. This makes our positive results all the more telling.

Despite the challenging control training of the QDT, we found that the NET regime transferred stronger improvements to arithmetic performance. Surely, some of the improvements seen in both groups could be due to testing repetition effects or to the standard school activity during the period of the training studies, but these factors affected both groups identically.

In both training regimes, we detected the signatures of ANS engagement, and in both regimes children's efficiency increased, confirming ANS trainability.

We measured the transfer of the training to arithmetic performance with the Pre-training and Post-training assessments, composed of the Additions Test, the Subtractions Test and the Operations Test. Importantly, the arithmetic knowledge necessary to perform these tests was neither explained nor practiced in any of the training conditions, because we wanted to assess the transfer of the trained aspects of numerical cognition to the performance in formal school arithmetic. Also, it is worth stressing that in the Pre-test phase, both groups started from a very comparable level, because there were no differences in any test at the Pre-training phase. Thus, between-group differences are not likely to have affected the results.

The data suggest that the NET regime improved children's math abilities more than the QDT regime, at least in the two tests that focus on exact arithmetic: the Additions and the Subtractions Tests. In the Additions Test, the NET group improved the number of correct answers after training while the QDT group did not. This suggests that the NET regime transferred to an improvement in children's abilities to solve exact additions quickly and correctly.

In the Subtractions Test, both groups significantly increased in their number of correct answers within 6 minutes, although improvement was much higher for the NET group ($p < .0001$) than for the QDT group ($p = .027$). We cannot attribute the QDT group's lower improvement to their training, because children had been attending regular math classes during the period of the intervention and math practice may have had an effect. However, because in Post-training the NET group scored significantly higher than the QDT group, the transfer to improved calculation of subtractions may be considered a real effect of the NET regime. This interpretation is supported by the aforementioned study (Ferres-Forga & Halberda, 2020), where the intervention group (which was identical to our current QDT protocol), did not improve in exact arithmetic (neither in addition nor in subtraction) above a "non-numerical standard school activities" control group. In view of this, we suggest that the NET regime may have transferred an improvement into children's abilities to correctly solve exact subtractions.

In the Operations test we detected a similarly high pre-post test improvement in both groups. This test is sensitive to children's approximate arithmetic level because it does not require explicit exact calculations but it does require an understanding of how basic arithmetic operations change quantities. In Ferres-Forga and Halberda (2020), the QDT regime improved performance in the Operations Test compared to a "non-numerical standard school activities" control, but only in children with a low pre-training math level. In the present study, we found a general improvement in both groups regardless of children's initial arithmetic abilities. As such, the present study cannot claim on its own that the trainings accounted for the improvements in the Operations Test (approximate arithmetic). The improvements could potentially be due to testing repetition effects, since the Operations Test was novel for the children and was not practiced at school during the intervention. For these reasons, it is only by taking into consideration Ferres-Forga

and Halberda's results, that we very cautiously suggest that both training regimes may have transferred benefits to approximate arithmetic. More research is needed to claim these benefits.

What could have generated the NET advantage in exact arithmetic improvements? The results suggest that training the mapping from estimated quantities to Arabic digits with the NET has transferred to an improvement in children's arithmetic abilities. Recruiting the ANS to estimate quantities and mapping these estimations to Arabic digits, along with the calibration provided by the passive trials (Izard & Dehaene, 2008), may have led children to better comprehend the meanings of digits. Such an explanation is consistent with findings that this mapping is related to children's mathematical performance (Booth & Siegler, 2006; Brankaer et al., 2014; Holloway & Ansari, 2009; Libertus et al., 2016; Marinova et al., 2021; Mazzocco et al., 2011a; Mundy & Gilmore, 2009) and can predict how well they will learn arithmetic skills (Booth & Siegler, 2008; Göbel et al., 2014; Habermann et al., 2020; Kolkman et al., 2013; Lyons et al., 2014; Marinova et al., 2021; Sasanguie et al., 2012). Further research is needed in order to consolidate this explanation and to explore if other forms of training with symbols (e.g. without quantities involved), could transfer to benefits in arithmetic.

A solid comprehension of the meaning of the numerical symbols used in arithmetic is important for the understanding of the arithmetic calculations themselves. We submit that the educational system may be overestimating 7-8 year olds' comprehension of this basic aspect of mathematical language. We propose that an appropriate training of mapping from estimated quantities to digits, such as the NET, could be a potent way to foster an overall improvement in children's arithmetic abilities.

Supplementary Materials

The Supplementary Materials contain the following items (for access see Index of Supplementary Materials below):

- Math Tests: Two versions of each test were prepared, with different problems, so that pre- and post-training booklets were unique and counterbalanced across the sessions
- Instructions to the teachers to guarantee the correct procedure during tests. The version in Catalan (original), and the version in English (translation)
- Results of the math tests from the 90 children
- Access to training materials (computer games)

Index of Supplementary Materials

Ferres-Forga, N., Halberda, J., Batalla-Ferres, A., & Bonatti, L. L. (2022). *Supplementary materials to "Improving mathematics performance in 7-year-old children: Training the mapping from estimated quantities to Arabic digits"* [Tests, instructions, results, and access to training materials]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.5633

PsychOpen GOLD

References

Amalric, M., & Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians. *Proceedings of the National Academy of Sciences, 113*(18), 4909–4917. https://doi.org/10.1073/pnas.1603205113

Au, J., Jaeggi, S. M., & Buschkuehl, M. (2018). Effects of non-symbolic arithmetic training on symbolic arithmetic and the approximate number system. *Acta Psychologica, 185*, 1–12. https://doi.org/10.1016/j.actpsy.2018.01.005

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition, 86*, 201–221. https://doi.org/10.1016/S0010-0277(02)00178-6

Bonny, J. W., & Lourenco, S. F. (2013). The approximate number system and its relation to early math achievement: Evidence from the preschool years. *Journal of Experimental Child Psychology, 114*(3), 375–388. https://doi.org/10.1016/j.jecp.2012.09.015

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*(1), 189–201. https://doi.org/10.1037/0012-1649.41.6.189

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development, 79*(4), 1016–1031. https://doi.org/10.1111/j.1467-8624.2008.01173.x

Brankaer, C., Ghesquière, P., & De Smedt, B. (2014). Children's mapping between non-symbolic and symbolic numerical magnitudes and its association with timed and untimed tests of mathematics achievement. *PLoS One, 9*(4), Article e93565. https://doi.org/10.1371/journal.pone.0093565

Brannon, E. M. (2006). The representation of numerical magnitude. *Current Opinion in Neurobiology, 16*(2), 222–229. https://doi.org/10.1016/j.conb.2006.03.002

Bugden, S., Szkudlarek, E., & Brannon, E. M. (2021). Approximate arithmetic training does not improve symbolic math in third and fourth grade children. *Trends in Neuroscience and Education, 22*, Article 100149. https://doi.org/10.1016/j.tine.2021.100149

Bulthé, J., De Smedt, B., & Op de Beeck, H. P. (2014). Format-dependent representations of symbolic and non-symbolic numbers in the human cortex as revealed by multi-voxel pattern analyses. *NeuroImage, 87*, 311–322. https://doi.org/10.1016/j.neuroimage.2013.10.049

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Sciences, 14*(12), 534–541. https://doi.org/10.1016/j.tics.2010.09.007

Butterworth, B., & Walsh, V. (2011). Neural basis of mathematical cognition. *Current Biology, 21*(16), R618–R621. https://doi.org/10.1016/j.cub.2011.07.005

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences, 13*(2), 83–91. https://doi.org/10.1016/j.tics.2008.11.007

Carey, S. (2004). Bootstrapping and the origin of concepts. *Daedalus, 133*(1), 59–68. https://doi.org/10.1162/001152604772746701

Carey, S. (2009). *The origin of concepts*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195367638.001.0001

Castronovo, J., & Göbel, S. M. (2012). Impact of high mathematics education on the number sense. *PLoS One, 7*(4), Article e33832. https://doi.org/10.1371/journal.pone.0033832

Cerda, G., Pérez, C., Navarro, J. I., Aguilar, M., Casas, J. A., & Aragón, E. (2015). Explanatory model of emotional-cognitive variables in school mathematics performance: A longitudinal study in primary school. *Frontiers in Psychology, 6*, Article 1363. https://doi.org/10.3389/fpsyg.2015.01363

Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica, 148*, 163–172. https://doi.org/10.1016/j.actpsy.2014.01.016

Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review, 8*(4), 698–707. https://doi.org/10.3758/BF03196206

Coubart, A., Izard, V., Spelke, E. S., Marie, J., & Streri, A. (2014). Dissociation between small and large numerosities in newborn infants. *Developmental Science, 17*(1), 11–22. https://doi.org/10.1111/desc.12108

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*(1–2), 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. Oxford University Press.

Dehaene, S. (2001). Précis of the number sense. *Mind & Language, 16*(1), 16–36. https://doi.org/10.1111/1468-0017.00154

Dehaene, S. (1993). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Attention & Performance XXII. Sensorimotor foundations of higher cognition* (pp. 527–574). Harvard Universitiy Press. https://doi.org/10.1093/acprof:oso/9780199231447.003.0024

Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(2), 314–326. https://doi.org/10.1037/0278-7393.21.2.314

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences, 21*(8), 355–361. https://doi.org/10.1016/S0166-2236(98)01263-6

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance, 16*(3), 626–641. https://doi.org/10.1037/0096-1523.16.3.626

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*(3–6), 487–506. https://doi.org/10.1080/02643290244000239

Demaray, M. K., & Elliot, S. N. (1998). Teachers' judgments of students' academic functioning: A comparison of actual and predicted performances. *School Psychology Quarterly, 13*(1), 8–24. https://doi.org/10.1037/h0088969

De Smedt, B., & Gilmore, C. (2011). Defective number module or impaired access? Numerical magnitude processing in first graders with mathematical difficulties. *Journal of Experimental Child Psychology, 108*(2), 278–292. https://doi.org/10.1016/j.jecp.2010.09.003

DeWind, N. K., & Brannon, E. M. (2012). Malleability of the approximate number system: Effects of feedback and training. *Frontiers in Human Neuroscience, 6*, Article 68. https://doi.org/10.3389/fnhum.2012.00068

Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology, 123*(1), 53–72. https://doi.org/10.1016/j.jecp.2014.01.013

Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Development Perspectives, 7*(2), 74–79. https://doi.org/10.1111/cdep.12019

Ferres-Forga, N., & Halberda, J. (2020). Approximate number system discrimination training for 7-8 year olds improves approximate, but not exact, arithmetics, and only in children with low pre-training arithmetic scores. *Journal of Numerical Cognition, 6*(3), 275–303. https://doi.org/10.5964/jnc.v6i3.277

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44*(1–2), 43–74. https://doi.org/10.1016/0010-0277(92)90050-R

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature, 447*(7144), 589–591. https://doi.org/10.1038/nature05850

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2010). Non-symbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. *Cognition, 115*(3), 394–406. https://doi.org/10.1016/j.cognition.2010.02.002

Göbel, S. M., Watson, S. E., Lervåg, A., & Hulme, C. (2014). Children's arithmetic development: It is number knowledge, not the approximate number sense, that counts. *Psychological Science, 25*(3), 789–798. https://doi.org/10.1177/0956797613516471

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science, 306*(5695), 496–499. https://doi.org/10.1126/science.1094492

Habermann, S., Donlan, C., Göbel, S. M., & Hulme, C. (2020). The critical role of Arabic numeral knowledge as a longitudinal predictor of arithmetic development. *Journal of Experimental Child Psychology, 193*, Article 104794. https://doi.org/10.1016/j.jecp.2019.104794

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "Number Sense": The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology, 44*(5), 1457–1465. https://doi.org/10.1037/a0012682

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences of the United States of America, 109*(28), 11116–11120. https://doi.org/10.1073/pnas.1200196109

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*(7213), 665–668. https://doi.org/10.1038/nature07246

He, Y., Zhou, X., Shi, D., Song, H., Zhang, H., & Shi, J. (2016). New evidence on causal relationship between Approximate Number System (ANS) acuity and arithmetic ability in elementary-school students: A longitudinal cross-lagged analysis. *Frontiers in Psychology, 7*, Article 1052. https://doi.org/10.3389/fpsyg.2016.01052

Hinrichs, J. V., Yurko, D. S., & Hu, J.-m. (1981). Two-digit number comparison: Use of place information. *Journal of Experimental Psychology: Human Perception and Performance, 7*(4), 890–901. https://doi.org/10.1037/0096-1523.7.4.890

Hollingsworth, W. H., Simmons, J. P., Coates, T. R., & Cross, H. A. (1991). Perceived numerosity as a function of array number, speed of array development, and density of array items. *Bulletin of the Psychonomic Society, 29*(5), 448–450. https://doi.org/10.3758/BF03333967

Holloway, I. D., & Ansari, D. (2008). Domain-specific and domain-general changes in children's development of number comparison. *Developmental Science, 11*(5), 644–649. https://doi.org/10.1111/j.1467-7687.2008.00712.x

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology, 103*(1), 17–29. https://doi.org/10.1016/j.jecp.2008.04.001

Hyde, D. C., Khanum, S., & Spelke, E. S. (2014). Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. *Cognition, 131*(1), 92–107. https://doi.org/10.1016/j.cognition.2013.12.007

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition, 106*(3), 1221–1247. https://doi.org/10.1016/j.cognition.2007.06.004

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences of the United States of America, 106*(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106

Kemp, S. (1984). Estimating the sizes of sports crowds. *Perceptual and Motor Skills, 59*(3), 723–729. https://doi.org/10.2466/pms.1984.59.3.723

Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. M. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning and Instruction, 25*, 95–103. https://doi.org/10.1016/j.learninstruc.2012.12.001

Krueger, L. E. (1982). Single judgments of numerosity. *Perception & Psychophysics, 31*(2), 175–182. https://doi.org/10.3758/BF03206218

Krueger, L. E. (1989). Reconciling Fechner and Stevens: Toward a unified psychophysical law. *Behavioral and Brain Sciences, 12*(2), 251–267. https://doi.org/10.1017/S0140525X0004855X

Langfus, J., Maiche, A., De León, D., Fitipalde, D., Mailhos, Á., & Halberda, J. (2019). Chapter 2 – The Effects of SES, grade-repeating, and IQ in a game-based approximate math intervention. In D. C. Geary, D. B. Berch, & K. M. Koepke (Eds.), *Cognitive Foundations for Improving Mathematical Learning* (Vol. 5, pp. 37–67). Academic Press. https://doi.org/10.1016/B978-0-12-815952-1.00002-5

Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development, 78*(6), 1723–1743. https://doi.org/10.1111/j.1467-8624.2007.01087.x

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition, 105*(2), 395–438. https://doi.org/10.1016/j.cognition.2006.10.005

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2017). From "sense of number" to "sense of magnitude": The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences, 40*, Article e164. https://doi.org/10.1017/S0140525X16000960

Lemer, C., Dehaene, S., Spelke, E. S., & Cohen, L. (2003). Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia, 41*(14), 1942–1958. https://doi.org/10.1016/S0028-3932(03)00123-4

Libertus, M. E., & Brannon, E. M. (2009). Behavioral and neural basis of number sense in infancy. *Current Directions in Psychological Science, 18*(6), 346–351. https://doi.org/10.1111/j.1467-8721.2009.01665.x

Libertus, M. E., Feigenson, L., & Halberda, J. (2013a). Is approximate number precision a stable predictor of math ability? *Learning and Individual Differences, 25*, 126–133. https://doi.org/10.1016/j.lindif.2013.02.001

Libertus, M. E., Feigenson, L., & Halberda, J. (2013b). Numerical approximation abilities correlate with and predict informal but not formal mathematics abilities. *Journal of Experimental Child Psychology, 116*(4), 829–838. https://doi.org/10.1016/j.jecp.2013.08.003

Libertus, M. E., Odic, D., Feigenson, L., & Halberda, J. (2016). The precision of mapping between number words and the approximate number system predicts children's formal math abilities. *Journal of Experimental Child Psychology, 150*, 207–226. https://doi.org/10.1016/j.jecp.2016.06.003

Lussier, C. A., & Cantlon, J. F. (2017). Developmental bias for number words in the intraparietal sulcus. *Developmental Science, 20*(3), Article e12385. https://doi.org/10.1111/desc.12385

Lyons, I. M., Ansari, D., & Beilock, S. L. (2012). Symbolic estrangement: Evidence against a strong association between numerical symbols and the quantities they represent. *Journal of Experimental Psychology: General, 141*(4), 635–641. https://doi.org/10.1037/a0027248

Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition, 121*(2), 256–261. https://doi.org/10.1016/j.cognition.2011.07.009

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in Grades 1-6. *Developmental Science, 17*(5), 714–726. https://doi.org/10.1111/desc.12152

Marinova, M., Reynvoet, B., & Sasanguie, D. (2021). Mapping between number notations in kindergarten and the role of home numeracy. *Cognitive Development, 57*, Article 101002. https://doi.org/10.1016/j.cogdev.2020.101002

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011a). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development, 82*(4), 1224–1237. https://doi.org/10.1111/j.1467-8624.2011.01608.x

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011b). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PLoS One, 6*(9), Article e23749. https://doi.org/10.1371/journal.pone.0023749

McCrink, K., Spelke, E. S., Dehaene, S., & Pica, P. (2013). Non-symbolic halving in an Amazonian indigene group. *Developmental Science, 16*(3), 451–462. https://doi.org/10.1111/desc.12037

Moll, K., Snowling, M. J., Göbel, S. M., & Hulme, C. (2015). Early language and executive skills predict variations in number and arithmetic skills in children at family-risk of dyslexia and typically developing controls. *Learning and Instruction, 38*, 53–62. https://doi.org/10.1016/j.learninstruc.2015.03.004

Moore, A. M., & Ashcraft, M. H. (2015). Children's mathematical performance: Five cognitive tasks across five grades. *Journal of Experimental Child Psychology, 135*, 1–24. https://doi.org/10.1016/j.jecp.2015.02.003

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature, 215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Mundy, E., & Gilmore, C. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology, 103*(4), 490–502. https://doi.org/10.1016/j.jecp.2009.02.003

Noël, M.-P., & Rousselle, L. (2011). Developmental changes in the profiles of dyscalculia: An explanation based on a double exact-and-approximate number representation model. *Frontiers in Human Neuroscience, 5*, Article 165. https://doi.org/10.3389/fnhum.2011.00165

Núñez-Peña, M. I., & Suárez-Pellicioni, M. (2014). Less precise representation of numerical magnitude in high math-anxious individuals: An ERP study of the size and distance effects. *Biological Psychology, 103*, 176–183. https://doi.org/10.1016/j.biopsycho.2014.09.004

Odic, D., Hock, H., & Halberda, J. (2014). Hysteresis affects approximate number discrimination in young children. *Journal of Experimental Psychology: General, 143*(1), 255–265. https://doi.org/10.1037/a0030825

Odic, D., Le Corre, M., & Halberda, J. (2015). Children's mappings between number words and the approximate number system. *Cognition, 138*, 102–121. https://doi.org/10.1016/j.cognition.2015.01.008

Odic, D., & Starr, A. (2018). An introduction to the Approximate Number System. *Child Development Perspectives, 12*(4), 223–229. https://doi.org/10.1111/cdep.12288

Odic, D., Valle-Lisboa, J., Eisinger, R., Olivera, M. G., Maiche, A., & Halberda, J. (2016). Approximate number and approximate time discrimination each correlate with school math abilities in young children. *Acta Psychologica, 163*, 17–26. https://doi.org/10.1016/j.actpsy.2015.10.010

Park, J., Bermudez, V., Roberts, R. C., & Brannon, E. M. (2016). Non-symbolic approximate arithmetic training improves math performance in preschoolers. *Journal of Experimental Child Psychology, 152*, 278–293. https://doi.org/10.1016/j.jecp.2016.07.011

Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychological Science, 24*(10), 2013–2019. https://doi.org/10.1177/0956797613482944

Park, J., & Brannon, E. M. (2014). Improving arithmetic performance with number sense training: An investigation of underlying mechanism. *Cognition, 133*(1), 188–200. https://doi.org/10.1016/j.cognition.2014.06.011

Piazza, M., & Eger, E. (2016). Neural foundations and functional specificity of number representations. *Neuropsychologia, 83*, 257–273. https://doi.org/10.1016/j.neuropsychologia.2015.09.025

Piazza, M., & Izard, V. (2009). How humans count: Numerosity and the parietal cortex. *The Neuroscientist, 15*(3), 261–273. https://doi.org/10.1177/1073858409333073

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron, 44*(3), 547–555. https://doi.org/10.1016/j.neuron.2004.10.014

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron, 53*(2), 293–305. https://doi.org/10.1016/j.neuron.2006.11.022

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306*(5695), 499–503. https://doi.org/10.1126/science.1102085

Pinel, P., Dehaene, S., Rivière, D., & LeBihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *NeuroImage, 14*(5), 1013–1026. https://doi.org/10.1006/nimg.2001.0913

Pinhas, M., Donohue, S. E., Woldorff, M. G., & Brannon, E. M. (2014). Electrophysiological evidence for the involvement of the approximate number system in preschoolers' processing of spoken number words. *Journal of Cognitive Neuroscience, 26*(9), 1891–1904. https://doi.org/10.1162/jocn_a_00631

Pinheiro-Chagas, P., Wood, G., Knops, A., Krinzinger, H., Lonnemann, J., Starling-Alves, I., . . . Haase, V. G. (2014). In how many ways is the approximate number system associated with exact calculation? *PLoS One, 9*(11), Article e111155. https://doi.org/10.1371/journal.pone.0111155

Pletzer, B., Kronbichler, M., Nuerk, H. C., & Kerschbaum, H. H. (2015). Mathematics anxiety reduces default mode network deactivation in response to numerical tasks. *Frontiers in Human Neuroscience, 9*, Article 202. https://doi.org/10.3389/fnhum.2015.00202

Purpura, D. J., & Logan, J. A. R. (2015). The nonlinear relations of the approximate number system and mathematical language to early mathematics development. *Developmental Psychology, 51*(12), 1717–1724. https://doi.org/10.1037/dev0000055

Räsänen, P., Salminen, J., Wilson, A. J., Aunio, P., & Dehaene, S. (2009). Computer-assisted intervention for children with low numeracy skills. *Cognitive Development, 24*(4), 450–472. https://doi.org/10.1016/j.cogdev.2009.09.003

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science, 19*(6), 607–614. https://doi.org/10.1111/j.1467-9280.2008.02130.x

Rousselle, L., & Noël, M. P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude processing. *Cognition, 102*(3), 361–395. https://doi.org/10.1016/j.cognition.2006.01.005

Russo, J., MacDonald, A., & Russo, T. (2021). The influence of making predictions on the accuracy of numerosity estimates in elementary-aged children. *International Journal of Science and Mathematics Education.* Advance online publication. https://doi.org/10.1007/s10763-021-10156-3

Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology, 30*(2), 344–357. https://doi.org/10.1111/j.2044-835X.2011.02048.x

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science, 20*(3), Article e12372. https://doi.org/10.1111/desc.12372

Sella, F., Tressoldi, P., Lucangeli, D., & Zorzi, M. (2016). Training numerical skills with the adaptive videogame "The Number Race": A randomized controlled trial on preschoolers. *Trends in Neuroscience and Education, 5*(1), 20–29. https://doi.org/10.1016/j.tine.2016.02.002

Shum, J., Hermes, D., Foster, B. L., Dastjerdi, M., Rangarajan, V., Winawer, J., Miller, K. J., & Parvizi, J. (2013). A brain area for visual numerals. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33*(16), 6709–6715. https://doi.org/10.1523/JNEUROSCI.4558-12.2013

Shusterman, A., Slusser, E. B., Halberda, J., & Odic, D. (2016). Acquisition of the cardinal principle coincides with improvement in approximate number system acuity in preschoolers. *PLoS One, 11*(4), Article e0153072. https://doi.org/10.1371/journal.pone.0153072

Siegel, A. W., Goldsmith, L. T., & Madson, C. R. (1982). Skill in estimation problems of extent and numerosity. *Journal for Research in Mathematics Education, 13*(3), 211–232. https://doi.org/10.2307/748557

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*(2), 428–444. https://doi.org/10.1111/j.1467-8624.2004.00684.x

Siegler, R. S., & Booth, J. L. (2005). Development of numerical estimation: A review. In J. I. D. Campbell (Ed.), *The handbook of mathematical cognition* (pp. 197–212). https://doi.org/10.4324/9780203998045-20

Simzar, R. M., Domina, T., & Tran, C. (2016). Eighth grade algebra course placement and student motivation for mathematics. *AERA Open, 2*(1), 1–26. https://doi.org/10.1177/2332858415625227

Slusser, E., Ditta, A., & Sarnecka, B. W. (2013). Connecting numbers to discrete quantification: A step in the child's construction of integer concepts. *Cognition, 129*(1), 31–41. https://doi.org/10.1016/j.cognition.2013.05.011

Slusser, E., & Sarnecka, B. W. (2011). Find the picture of eight turtles: A link between children's counting and their knowledge of number word semantics. *Journal of Experimental Child Psychology, 110*(1), 38–51. https://doi.org/10.1016/j.jecp.2011.03.006

Starr, A., Libertus, M. E., & Brannon, E. M. (2013a). Infants show ratio-dependent number discrimination regardless of set size. *Infancy, 18*(6), 927–941. https://doi.org/10.1111/infa.12008

Starr, A., Libertus, M. E., & Brannon, E. M. (2013b). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences of the United States of America, 110*(45), 18116–18120. https://doi.org/10.1073/pnas.1302751110

Stoet, G., Bailey, D. H., Moore, A. M., & Geary, D. C. (2016). Countries with higher levels of gender equality show larger national sex differences in mathematics anxiety and relatively lower parental mathematics valuation for girls. *PLoS One, 11*, Article e0153857. https://doi.org/10.1371/journal.pone.0153857

Sullivan, J., & Barner, D. (2014). The development of structural analogy in number-line estimation. *Journal of Experimental Child Psychology, 128*, 171–189. https://doi.org/10.1016/j.jecp.2014.07.004

Szkudlarek, E., & Brannon, E. M. (2018). Approximate arithmetic training improves informal math performance in low achieving preschoolers. *Frontiers in Psychology, 9*, Article 606. https://doi.org/10.3389/fpsyg.2018.00606

Szkudlarek, E., Park, J., & Brannon, E. M. (2021). Failure to replicate the benefit of approximate arithmetic training for symbolic arithmetic fluency in adults. *Cognition, 207*, Article 104521. https://doi.org/10.1016/j.cognition.2020.104521

Temple, E., & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences of the United States of America, 95*(13), 7836–7841. https://doi.org/10.1073/pnas.95.13.7836

Thien, L. M., & Ong, M. Y. (2015). Malaysian and Singaporean students' affective characteristics and mathematics performance: Evidence from PISA 2012. *SpringerPlus, 4*(1), Article 563. https://doi.org/10.1186/s40064-015-1358-z

Tobia, V., Bonifacci, P., & Marzocchi, G. M. (2021). Symbolic versus non-symbolic training for improving early numeracy in preschoolers at risk of developing difficulties in mathematics. *Research in Developmental Disabilities, 111*, Article 103893. https://doi.org/10.1016/j.ridd.2021.103893

Tournaki, N. (2003). Effect of student characteristics on teachers' predictions of student success. *The Journal of Educational Research, 96*(5), 310–319. https://doi.org/10.1080/00220670309597643

Valle-Lisboa, J., Cabana, Á., Eisinger, R., Mailhos, Á., Luzardo, M., Halberda, J., & Maiche, A. (2016). Cognitive abilities that mediate SES's effect on elementary mathematics learning: The Uruguayan tablet-based intervention. *Prospects, 46*(2), 301–315. https://doi.org/10.1007/s11125-017-9392-y

van Oeffelen, M. P., & Vos, P. G. (1982). A probabilistic model for the discrimination of visual number. *Perception & Psychophysics, 32*(2), 163–170. https://doi.org/10.3758/BF03204275

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child Development, 85*(3), 1062–1076. https://doi.org/10.1111/cdev.12165

Wagner, J. B., & Johnson, S. C. (2011). An association between understanding cardinality and analog magnitude representations in preschoolers. *Cognition, 119*(1), 10–22. https://doi.org/10.1016/j.cognition.2010.11.014

Wang, J., Odic, D., Halberda, J., & Feigenson, L. (2016). Changing the precision of preschoolers' approximate number system representations changes their symbolic math performance. *Journal of Experimental Child Psychology, 147*, 82–99. https://doi.org/10.1016/j.jecp.2016.03.002

Wilson, A. J., Dehaene, S., Dubois, O., & Fayol, M. (2009). Effects of an adaptive game intervention on accessing number sense in low-socioeconomic-status kindergarten children. *Mind, Brain and Education, 3*(4), 224–234. https://doi.org/10.1111/j.1751-228X.2009.01075.x

Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of "The Number Race", an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions, 2*(1), Article 20. https://doi.org/10.1186/1744-9081-2-20

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36*(2), 155–193. https://doi.org/10.1016/0010-0277(90)90003-3

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology, 24*(2), 220–251. https://doi.org/10.1016/0010-0285(92)90008-P

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*(1), B1–B11. https://doi.org/10.1016/S0010-0277(99)00066-9

Zbrodoff, N. J., & Logan, G. D. (2005). *What everyone finds: The problem-size effect.* (J. D. Campbell, Ed.). Psychology Press.

PsychOpen GOLD

MCLS

Journal of Numerical Cognition (JNC) is an official journal of the Mathematical Cognition and Learning Society (MCLS).

ⓛⓟ leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.

PsychOpen ^GOLD