

## Empirical Research

# 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

Nuria Ferres-Forga<sup>\*a</sup> , Justin Halberda<sup>b</sup>

[a] Center for Brain and Cognition, Universitat Pompeu Fabra, Barcelona, Spain. [b] Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD, USA.

## Abstract

We investigated whether training the Approximate Number System (ANS) would transfer to improved arithmetic performance in 7-8 year olds compared to a control group. All children participated in Pre- and Post-Training assessments of exact symbolic arithmetic (additions and subtractions) and approximate symbolic arithmetic abilities (a novel test). During 3 weeks of training (approximately 25 minutes per day, two days per week), we found that children in the ANS Training group had stable individual differences in ANS efficiency and increased in ANS efficiency, both within and across the training days. We also found that individual differences in ANS efficiency were related to symbolic arithmetic performance. Regarding arithmetic performance, both the ANS training group and the control group improved in all tests (exact and approximate arithmetics tests). Thus, the ANS training did not show a specific effect on arithmetic performance. However, considering the initial arithmetic level of children, we found that the trained children showed a higher improvement on the novel approximate arithmetic test compared to the control group, but only for those children with a low pre-training arithmetic score. Nevertheless, this difference within the low pre-training arithmetic score level was not observed in the exact arithmetic test. The limited benefits observed in these results suggest that this type of ANS discrimination training, through quantity comparison tasks, may not have an impact on symbolic arithmetics overall, although we cautiously propose that it could help with approximate arithmetic abilities for children at this age with below-average arithmetic performance.

**Keywords:** approximate number system, early mathematics, learning math in children, math ability, mathematics education, number sense, school mathematics

Journal of Numerical Cognition, 2020, Vol. 6(3), 275–303, <https://doi.org/10.5964/jnc.v6i3.277>

Received: 2019-12-19. Accepted: 2020-10-18. Published (VoR): 2020-12-03.

\*Corresponding author at: Universitat Pompeu Fabra, Ramon Trias Fargas, 25-27, 08005 Barcelona, Spain. E-mail: [nuriaferresforaga@gmail.com](mailto:nuriaferresforaga@gmail.com)



This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

When children begin their explicit mathematical training, typically their skills vary considerably (Haworth, Kovas, Petrill, & Plomin, 2007; Simzar, Domina, & Tran, 2016). The mathematical abilities required in the early school years range from simpler skills like intuitive estimation and comparison of numbers of items to more complex abilities involved in operations such as multiplication. Moreover, early mathematical skills are an important predictor of life outcomes such as health and salary (Parsons & Bynner, 2005; Peters, Hart, Tusler, & Fraenkel, 2014; Ritchie & Bates, 2013), and of later achievements in school and in college (Geary, Hoard, Nugent, & Bailey, 2013; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Papay, Murnane, & Willett, 2014; Watts, Duncan, Siegler, & Davis-Kean, 2014). Thus children who fall behind their peers are at a higher risk of being left behind throughout their schooling, resulting in lower academic achievement and continuing difficulties into adulthood

(Aubrey, Godfrey, & Dahl, 2006; Geary, 2013; Watts et al., 2014). Helping those students whose performance is weaker may be particularly challenging, because factors such as motivation (Simzar et al., 2016), math anxiety (Pletzer, Kronbichler, Nuerk, & Kerschbaum, 2015; Z. Wang et al., 2014), lack of a predisposition towards mathematics (Cerdeira et al., 2015), and teacher effects (Demaray & Elliot, 1998; Tournaki, 2003) come into play. Game-based training of basic math skills may be helpful, as it may be less threatening to those who have difficulty learning mathematics. Indeed, training math-related basic cognitive skills may increase students' confidence in their sense of number, and this might then transfer to their more school-based math abilities (Odic, Hock, & Halberda, 2014; Wang, Odic, Halberda, & Feigenson, 2016; but see Merkley, Matejko, & Ansari, 2017; Wang, Odic, Halberda, & Feigenson, 2017). The search for cognitively motivated, easily implemented, enjoyable and efficient training regimes is currently very active. Here, we implement a one-month intervention to test whether it can improve performance of arithmetic operations in children.

Classroom-based interventions and teaching have not always been found to be effective in helping students who struggle with math (e.g., Desimone & Long, 2010; Morgan, Farkas, & Maczuga, 2015). However, experimental results suggest that even very basic intuitive practice, for example in estimating and comparing the quantities of items in collections, might help them (e.g., Bonny & Lourenco, 2013; Bugden & Ansari, 2016). Previous research has found a relationship between individual differences in formal mathematics and the precision of a basic intuitive sense of number (for a review see Feigenson, Libertus, & Halberda, 2013). Although controversial (Butterworth, 2011; Dehaene, 2009; Elliott, Feigenson, Halberda, & Libertus, 2019; Fuhs, Nesbitt, & O'Rear, 2018; Gilmore et al., 2013; Le Corre & Carey, 2007; Li, Zhang, Wang, Ding, & Si, 2018; Libertus, Odic, Feigenson, & Halberda, 2016; Malone, Burgoyne, & Hulme, 2020; Mussolin, Nys, Content, & Leybaert, 2014; Noël & Rousselle, 2011; Núñez, 2017; Odic, Le Corre, & Halberda, 2015; Odic et al., 2016; Rips, Bloomfield, & Asmuth, 2008; Shusterman, Slusser, Halberda, & Odic, 2016; Soltész et al., 2010), this relationship can be found at multiple ages: in preschoolers (Bonny & Lourenco, 2013; Gray & Reeve, 2014; Libertus, Feigenson, & Halberda, 2011; Mazzocco, Feigenson, & Halberda, 2011b; Moore, VanMarle, & Geary, 2016; Shusterman et al., 2016; Starr, Libertus, & Brannon, 2013b), in primary school children (Gilmore, McCarthy, & Spelke, 2010; Nosworthy, Bugden, Archibald, Evans, & Ansari, 2013; Pinheiro-Chagas et al., 2014; Valle-Lisboa et al., 2016), in adolescents (Geary et al., 2013; Halberda, Mazzocco, & Feigenson, 2008) and in adults (Amalric & Dehaene, 2016; Castronovo & Göbel, 2012; Halberda, Ly, Wilmer, Naiman, & Germine, 2012).

The intuitive sense of number comes from a cognitive system often described as the Approximate Number System (ANS), comprising several components needed to understand and manipulate numerical quantities; among them, a rough ability to estimate magnitudes ("there are many/few marbles in that jar"), magnitude comparisons ("this jar has more marbles than that jar"), and the manipulation of numerical quantities through basic approximate arithmetic operations ("I can remove around half of the marbles from this jar to make it roughly equal to that jar"). Although not necessarily innate (Leibovich, Katzin, Harel, & Henik, 2017), the ANS seems to be present in young infants (Feigenson, Dehaene, & Spelke, 2004; Hyde, 2011; Hyde & Spelke, 2011), in people of all cultures (McCrink, Spelke, Dehaene, & Pica, 2013; Pica, Lemer, Izard, & Dehaene, 2004), and ages (Halberda et al., 2012), and even in other animals (e.g., Agrillo, Piffer, & Bisazza, 2011; Beran, Evans, & Harris, 2008; Dehaene, Dehaene-Lambertz, & Cohen, 1998; Jones & Brannon, 2012). While all people may have an ANS, the precision of the ANS varies across individuals throughout the lifespan (Halberda et al., 2012). A more precise ANS allows a person to make more confident distinctions between large close numbers, making more confident numerical distinctions. Interestingly, preschoolers, students across ages, and adults who have a more precise and accurate ANS also tend to do better on tests of formal, written, mathematics

(Amalric & Dehaene, 2016; Halberda et al., 2008; Y. He et al., 2016; Libertus et al., 2011). Importantly, this relationship holds even when general intelligence, verbal ability, and many other cognitive factors are controlled for (Halberda et al., 2008; Libertus, Feigenson, & Halberda, 2013).

While some studies suggest that executive functions (EF) might serve as the underlying link between better ANS performance and better school mathematics performance (Gilmore et al., 2013; Li et al., 2018), some recent work has found the opposite to be true (i.e., that ANS precision and number language or number knowledge survive as the only unique predictors of school math ability while EF are not unique (Malone et al., 2020, 2019; Purpura & Logan, 2015). This is an area that will benefit from additional exploration.

Also, early differences in ANS precision, but not other factors such as verbal or performance IQ, executive functions, language knowledge or visuospatial skills, predict later math performance at school (Mazzocco et al., 2011b). Because links between the ANS and formal mathematics abilities are not universally accepted, it remains important to test if ANS precision relates to arithmetic abilities.

The precision of the ANS may improve across life, especially during the school-age years, attaining its highest precision sometime around age 30 years (Halberda et al., 2012). These improvements justify the hope that interventions aimed at training the ANS might transfer to improvements in some arithmetic skills. Because these basic arithmetic skills can be built into computer tasks (Halberda et al., 2008; Hyde, Khanum, & Spelke, 2014; Park & Brannon, 2014), game-based ANS practice could help to train these early skills, with perhaps a more marked effect in students who are from particularly fragile social environments (Fuhs & McNeil, 2016; Valle-Lisboa et al., 2016; Wilson, Dehaene, Dubois, & Fayol, 2009) or possess low numeracy skills (Liang, Zhang, Long, Deng, & Liu, 2020; Praet & Desoete, 2014; Räsänen, Salminen, Wilson, Aunio, & Dehaene, 2009). Although several recent papers suggest that interventions can improve the ANS (e.g., Cochrane, Cui, Hubbard, & Green, 2019; Liang, Zhang, Long, Deng, & Liu, 2020; Van Herwegen, Costa, & Passolunghi, 2017; J. Wang et al., 2016), potentially transferring to school mathematics performance, these results remain controversial (Szűcs & Myers, 2017). The exact nature of the training and the strength of the results are still unknown. For example, in a brief (5-minute) intervention to improve ANS confidence, children showed immediate gains in symbolic arithmetic (Wang, Odic, Halberda, & Feigenson, 2016) although the causal link was questioned based on concerns about the study design (Merkley, Matejko, et al., 2017). That is, Merkley, Matejko, and Ansari (2017) argued that randomized control trials, with pre-testing of all participants, supported stronger inferences while the Wang et al. (2016) procedure relied on a treatment with post-test design (see also Wang, Odic, Halberda, & Feigenson, 2017). Thus, the possibility of ANS training transferring to symbolic math abilities remains an area of controversy. Even under the assumption that there is some transfer, it is also true that the persistence of these potential gains over an extended period remains unexplored. Other studies showed that training in ANS arithmetic (e.g., estimating the results of adding and subtracting clouds of dots) can lead to improvements in symbolic mathematics in children (Hyde et al., 2014; Obersteiner, Reiss, & Ufer, 2013; Park, Bermudez, Roberts, & Brannon, 2016), in adolescents (Knoll et al., 2016), and in adults (DeWind & Brannon, 2012; Park & Brannon, 2014), but training in simple magnitude comparisons of two collections (e.g., which jar has more marbles) did not produce comparable transfer (Lindskog & Winman, 2016; Park & Brannon, 2014, 2016; Pinheiro-Chagas et al., 2014).

In the present study, we aimed to add a contribution to the debate surrounding the possible link between the ANS (NB: whose very existence remains controversial) and symbolic arithmetic, using a robust pre-test,

intervention, post-test design with a well-matched control group. In the school system where we realized our study, the second-grade math program includes arithmetic (additions, subtractions, and some multiplications), basic geometry (spatial orientation, figures, measure of figures), measure of magnitudes (weight, volumes, distances, units of measurement), and basic statistics and probability, among other topics. While the specific program covered by other systems may vary in other countries, it is reasonable to expect that any school system at the target age will include arithmetic, as it gives children the basic steps necessary to succeed in other parts of mathematics. Thus, in our study we focused on arithmetic as the privileged area to test potential effects of ANS training.

In the current study, we focused on one grade level in a controlled environment, while carefully controlling dosing time and conditions of exposure to the ANS training game relative to its control. We studied second-graders, aged 7- to 8-years, for two reasons. First, previous studies have shown that many developmental changes occur around that age (Holloway & Ansari, 2009; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013; Siegler & Booth, 2004); and, second, at this grade students know enough arithmetics to be tested with a variety of arithmetic operations, including additions, subtractions, and even some multiplications. The arithmetic operations in second grade (additions, subtractions, and multiplications) represent different levels of difficulty for a child; for example additions are easier than subtractions (e.g., Knops, Dehaene, Berteletti, & Zorzi, 2014), and subtractions are easier than multiplications (Prado et al., 2011). In the present study, due to the approximative nature of the training, we introduced a novel arithmetic operations test to assess performance in approximate symbolic arithmetics, alongside a school-like test of exact symbolic arithmetics.

We asked if three-weeks of ANS training in approximate quantity comparisons could enhance arithmetic skills in children, and if improvements were related to their initial arithmetic abilities. This aspect of our work was motivated by the literature, as several studies suggest that children who are struggling the most may benefit from extended ANS game-based interventions (Langfus et al., 2019; Mazzocco, Feigenson, & Halberda, 2011a; Purpura & Logan, 2015; Valle-Lisboa et al., 2016). In order to understand the sources of children's differences in mathematical achievements, we must understand how the content taught in school relates to basic cognitive abilities. A crucial aspect of this endeavor – one with pivotal societal consequences – is to determine if some training can help struggling students improve. We believe that school-based training is a powerful way to progress in this direction.

In the present study, through our school-based intervention, we aimed to further test the controversial ANS construct and its possible link to symbolic arithmetic. Specifically, we evaluate whether individual differences in ANS efficiency are related to symbolic arithmetic performance; whether training the ANS for an extended period can improve children's ANS efficiency; whether such improvement will transfer to any arithmetic skills; and whether the initial arithmetic level of children before the training helps to determine the efficacy of such an intervention.

## Method

### Participants

Ninety-one second-graders (44 girls; average age = 7 years 10 months, range = 7 years 3 months to 9 years 3 months) were recruited from Hamelin International Laie School (<http://www.hamelininternacional->

laie.com/school/) in Barcelona, Catalunya. The school is a non-elite private school with no academic requirements or admissions tests for entering, and does not exclude children along the course of the school years based on academic performance. Tests scores for college entrance exams from the Hamelin International Laie School do not differ from other national schools. Participants came from families of middle to high socioeconomic status. According to teachers' reports, our sample did not include supremely gifted students or students who were greatly below the typical level. Thus, the sample is a fair representation of Catalan school students.

Participants came from four different classrooms (respectively, with  $n = 24, 24, 23$  and  $20$ ). We compared a control group, which relied on the standard school activities (Control,  $n = 44$ ) with an ANS Training group ( $n = 47$ ). Two teachers of mathematics served two classes each, randomly assigned to Control or ANS Training group so that teacher effects were controlled. Training activity occurred during the computer technology class. Throughout the study, computer and math classes were kept separate. Teachers and research staff never discussed relationships between computer activity and math classes, so that students were unaware that training was related with math performance.

## Materials

### Arithmetical Competence Assessment

We prepared three problem booklets for pencil-and-paper tests. Two of them contained straightforward arithmetic problems: one *additions* (Figure 1a) and one *subtractions* test booklet (Figure 1b), based on the math fluency subtest of the Woodcock-Johnson Tests of Achievement (Woodcock, McGrew, & Mather, 2001), with additional problems and a 6-minute testing time to better detect individual differences. For these, children had to write in the correct answer to the problems. The third booklet was a novel *operations* test, which we designed (Figure 1c) and presented a kind of problem that had not been taught before: each problem was presented along with its result, but the operation sign was omitted from the equation. Children had to mark whether the solution demanded an addition, a subtraction, or a multiplication sign. To correctly answer, students needed not assess whether the given answer was correct (though it always was). Rather, by estimating whether the solution resulted in a slightly larger, a much larger, or a smaller number, they could decide whether the correct operation sign was, respectively, an addition, a multiplication or a subtraction. Our aim was to introduce a novel test relating approximate arithmetic ability and ANS training not confounded with normal arithmetic activities at school during the period of training.

Motivated by findings in the literature (Bonny & Lourenco, 2013; Langfus et al., 2019; Mazzocco et al., 2011a), we wanted some way of characterizing children's initial arithmetic abilities. Because children's school grades at this age were not detailed enough to classify them according to their mathematical abilities, we used the additions and subtractions subtests before the training to identify three initial arithmetic performance groups (Below-Average, Average and Above-Average tertiles) and then use the novel operations subtest to measure pre- and post-training performance in approximate arithmetics.

Test difficulties were created under the supervision of the teachers. Two versions of each subtest were prepared, with different problems, so that pre- and post-training booklets were unique and counterbalanced across the sessions. Difficulty and problem order were randomized. The number of problems for each subtest was large, so that children could not complete all of them during a 6-minute allotted time: (210 additions

190 subtractions and 117 operations). In the Additions subtest, each addendum could reach at most 18, so that the highest sum was  $18 + 18$  and the lowest  $0 + 0$ . In the Subtractions subtest, both the minuend and subtrahend ranged between 0 and 18; the result of the subtractions was always positive. Additions and subtractions were presented in 10-sheet booklets, in a column operation algorithm form (Figure 1a and b); and the operations subtest was presented on 3 sheets (Figure 1c). The Operations test included problems of addition, subtraction and multiplication. For the additions in the Operations test, the maximum value for the first and second addenda was, respectively, 7 and 10 (i.e.,  $7 + 10 = 17$ ). For the subtractions in the Operation test, both the minuend and subtrahend ranged between 1 and 11, with the result always positive (i.e.,  $11 - 5 = 6$ ). For the multiplications, the timetables of 1 to 5 and 10 were used (i.e.,  $4 \times 5 = 20$ ), because children had been exposed only to them according to their teachers. The three types of operations and the problem difficulties were presented in random order. The booklet contained a roughly equivalent distribution of problem kinds (42 and 43 additions in the first and second version respectively; 38 and 39 subtractions; and 37 and 35 multiplications).

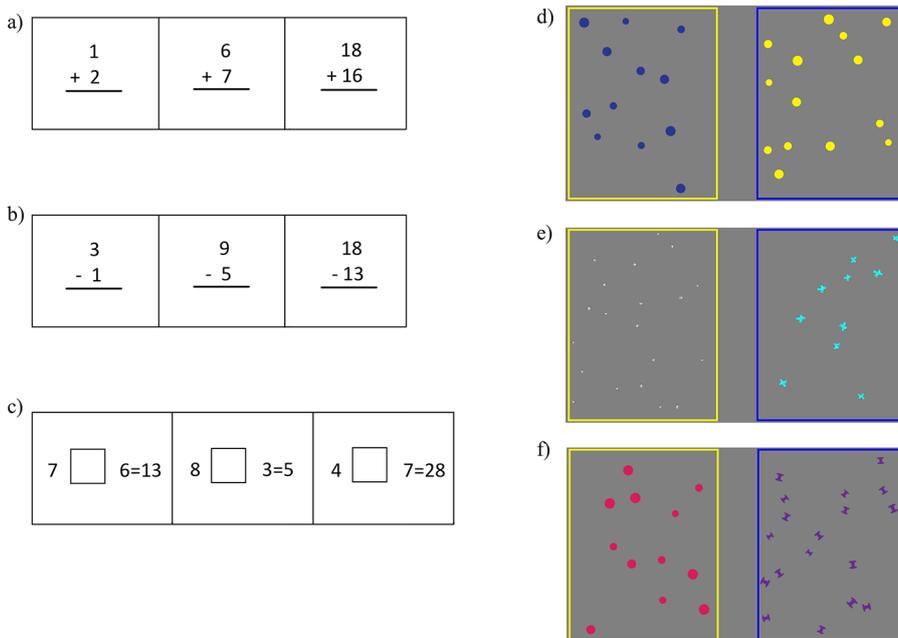


Figure 1. Examples of problems in additions subtest (a), subtractions subtest (b) and operations subtest (c). Examples of the three models of object size control for surface area: size confounded (d), size controlled (e) and stochastic size-control (f).

### Computer Activities

The computer classroom had 25 laptops (HP 620, Pentium (R) Dual Core 2.30GHz, 4GB DRAM, 64-bit; Windows 7 Home Premium), with individual headphones. Training and control groups ran different activities. Controls practiced two commercial programs, *Tux Paint*, and *Microsoft Word*. ANS Training children practiced a modified version of *Panamath* (Halberda et al., 2008). Written in Java SE6, the program generates trials displaying collections of items inside two rectangles appearing side by side. For example, twelve 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. The items were presented in seven different ratios (larger set/smaller set): 3, 2, 1.5, 1.25, 1.17, 1.14, 1.1. For example, in a 3-ratio trial children could see 21 blue vs.

7 yellow dots. Smaller ratios generate more difficult trials. On each trial, items were displayed for 1382 ms. To vary the relationship between surface area and number, there were 3 different models controlling object size. Forty-two percent of the trials were size-confounded. In them, items average size was equal for both sets, so that objects occupied cumulative surface areas congruent with number (Figure 1d). Forty-two percent of the trials were size-controlled, with the object average size smaller for the larger set, so that the ratio of the cumulative area occupied in each set was 1 (Figure 1e). In the remaining 16% of trials, object sizes were stochastically varied, so that children had no consistent size cue for number. In these, the average size of the objects varied randomly between being size-anticorrelated (where the numerically larger set had less total area on the screen), size-controlled, and size-confounded. We called these runs *stochastic size control* (Figure 1f). While there are a diversity of views about how to control for non-numerical cues in approximate number stimuli, our method here recapitulates one such approach that has appeared in several papers (Hellgren, Halberda, Forsman, Adén, & Libertus, 2013; Libertus et al., 2013; Libertus, Feigenson, Halberda, & Landau, 2014; Tosto et al., 2014). But, in fact, we expected that our children would rely on continuous cues to some extent and we did not desire to eliminate this possibility in our stimuli because this might be part and parcel of normal ANS functioning (Halberda, 2019).

## Procedure

The experiment had three phases: *Pre-training*, *Experimental/Control training*, and *Post-training*. The Pre-training and Post-training assessments were intended to measure the mathematical competence of the participants before and after training. These were conducted in children's classrooms, in the presence of their math teacher. From the child's perspective, they were independent of the computer activities. Children had to answer as many questions as possible in 6 minutes per test. The additions subtest was always ran first. Then children had to stop and wait until they were given the subtractions subtest. The final test was the operations subtest; as it contained a novel kind of problems teachers had to briefly explain what children were supposed to do beforehand. Teachers completed 3 examples on the board, in front of the class, one per each problem kind, after which children began the subtest. The Pre-training occurred two days before the first training session (initial assessment); the Post-training one day after training completion (final assessment).

The training phase was administered in six different sessions within a three-week period, at a pace of two sessions per week. This choice was motivated by previous studies (Langfus et al., 2019; Valle-Lisboa et al., 2016). Both the intervention and control groups were trained on the same days (Wednesdays and Fridays), for the same amount of time, and in the same computer classroom.

## Control

Control children practiced *Tux Paint* and *Microsoft Word*, learning standard activities available in these programs. No activity involved approximate number comparison. Computer classes started with teachers' instructions, after which children worked individually, in the same class as for the same period as the experimental group.

## Approximate Number System Training (ANS Training)

Children practiced the Panamath quantity discrimination game. During training, the computer teacher and the experimenter were always present. Children wore headphones and trained simultaneously. By observation, children appeared to like this game. The ANS Training presented children with a sequence of pictures, asking

them to make ordinal comparisons of the rapidly flashed collections appearing onscreen. On each trial, they saw two collections of items appearing onscreen side by side, and needed to rapidly estimate which side had more items (Figure 1 d-f), typing their answers on a keyboard (“f” and “j” keys for left or right side respectively). As each presentation lasted about 1.3 s, children could not count the items, but had to rapidly estimate which of the two sets was larger. Different collections of items were used on each run of trials, so that the game maintained children's interest. For example, one run could present blue vs. yellow dots; another could present cars vs. bears; yet another could display birds vs. dogs. There were 35 trials per run. Children completed approximately 24 runs over the course of three weeks. Feedback was provided after every response, with a high/low pitched beep indicating correct/incorrect answers.

When introducing ANS Training, children were told that they would play a game where they would see some objects – for example, blue and yellow dots – and would have to choose if there were more blue or yellow dots. They were informed that two different sounds would provide them with feedback. They were also told that the game would vary in difficulty, and that both speed and accuracy were important. Most children completed 24 runs. Six participants completed 21, one 22 and one 20. Each run was composed of 35 trials (or about 6-7 minutes). Each run started with the easiest ratio in the first five trials. Then, every five trials the game increased in difficulty, with the ratios becoming closer to 1, until the seven different ratios were presented in each run. This confidence-scaffolding procedure was implemented to increase ANS precision and confidence, based on previous studies with brief interventions (Odic, Hock, & Halberda, 2014; Wang et al., 2016). Children completed the ANS Training during their regular computer class time.

## Results

### ANS Training: Approximate Number Comparison

For our purpose, the primary measure for the effectiveness of the ANS training is post-pretest differences in symbolic math performance. However, before analyzing the effect of training, we assessed whether the training did engage the ANS. We looked for the main signature, ratio-dependent performance that results in a specific curve of percent correct as a function of ratio (Feigenson et al., 2004; Halberda et al., 2008; Libertus & Brannon, 2009; Starr, Libertus, & Brannon, 2013a), in which participants' accuracy at determining the bigger of two approximate numerosities decreases as the ratio between the numbers decreases. This ratio-dependence is predicted by Weber's law. Figure 2 presents the data from the ANS Training group, separated by the type of size control of the stimuli. The ratio-dependent performance curve appears for all three size-control trial types: as the numerical ratio between the two collections becomes easier, children's percentage of correct responses improves. Furthermore, notice that even though overall percent correct is somewhat lower, the curve of ratio-dependent performance appears also for the stochastic size-controlled and size-controlled trials. Children chose the numerically greater collection well above chance; this suggests that while they may sometimes have been led astray by the numerically smaller collection having a larger total surface area, they could still succeed on a majority of trials. Some portion of the difference in performance across the stimulus sets could be due to the differing EF demands of the stimuli (Gilmore et al., 2013; Li et al., 2018). And, we note that although performance exhibits the smooth curve of the ANS, other dimensions of the stimuli we did not control could contribute to number decisions (Abreu-Mendoza, Soto-Alba, & Arias-Trejo, 2013; Clayton, Gilmore, &

Inglis, 2015; L. He, Zhou, Zhou, He, & Chen, 2015; Hollingsworth, Simmons, Coates, & Cross, 1991; Izard & Dehaene, 2008; Leibovich & Henik, 2013; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004).

We note, however, that our main aim was not to identify which dimension is responsible for the ANS signature, but rather to ensure that the ANS is engaged regardless of the cues that empower it. The question of which perceptual features drive the ANS is an important one that continues to inspire debate, but it was not a focus of our present work.

Using the standard psychological model and fitting methods (e.g., Odic et al., 2014) we estimated the observed ANS precision for the whole ANS training group (i.e., Weber fraction or  $w$ ). We found the best fit model returned a  $w$  of 0.192 ( $SD = 0.09$ ). This value is not much different from what other studies with young participants have found (6-year-olds:  $w = 0.179$ , Halberda & Feigenson, 2008; Piazza & Izard, 2009). The curves in Figure 2 are generated by fitting a model of Weber's law to the mean performance of children in each ratio for each size control type. That is, each child contributes equally to the curves, the curves are fit to the combined data from all children, and the error bars are  $\pm SE$  for the group performance.

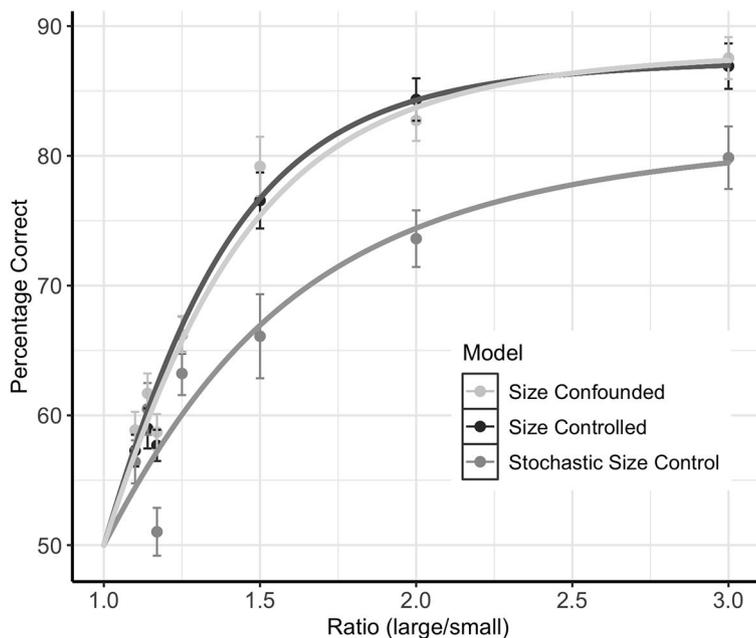


Figure 2. Accuracy on the ANS acuity task is a function of the ratio between numerical values. X-axis shows the seven different ratios presented during the training (1.1, 1.14, 1.17, 1.25, 1.5, 2, 3). The three ways of manipulating surface-area are shown, each with a corresponding Weber curve. The children's performance is consistent with their choice based on number.

It is interesting to assess to what extent, across the training sessions, children improved. However, the measure which best reveals such an improvement is not immediately obvious. Both response time (RT) and accuracy have been shown to be reliable measures of performance in this and similar tasks with children as young as four years of age. However, with few exceptions (e.g. Park & Brannon, 2014), the measures were not taken across an extended period of training (Bartelet, Vaessen, Blomert, & Ansari, 2014; Cantlon et al., 2009; Libertus et al., 2011, 2013; Mazzocco et al., 2011b; Murphy, Mazzocco, Hanich, & Early, 2007). During an extended training period, children can improve either by reducing response time or by increasing accuracy.

For this reason, *efficiency*, operationalized as the percentage correct divided by the RT, encompassing both improvements in RT (i.e., getting faster) and in percent correct (i.e., becoming more accurate), seems most appropriate (Bartelet et al., 2014; Libertus et al., 2013; Mazzocco et al., 2011b; Murphy et al., 2007; Park & Brannon, 2014). It increased across the ANS Training task ( $R^2 = .827, p < .001$ ; Figure 3). While not crucial to our purpose, we observe that while children greatly improved in their response times, they showed a tendency to be less accurate across training; thus, the improvements in ANS efficiency are primarily driven by improvements in RT.

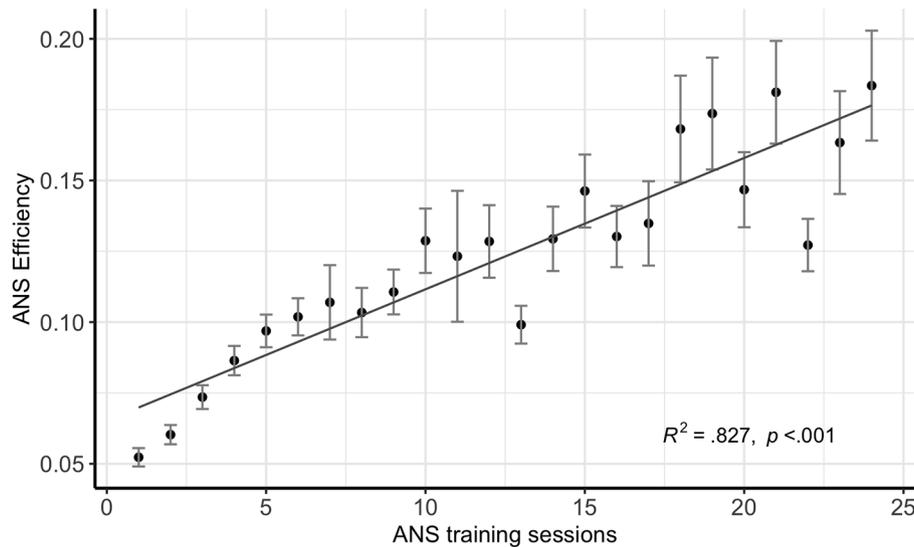


Figure 3. Efficiency (Percentage correct / RT) on the ANS comparison task as a function of progress in training (composed by 24 runs across three weeks;  $n = 47$ ). The line is a fit of the mean efficiency of the group in each session. Error bars present  $\pm SE$  for the group performance.

ANS training was also effective across the three individual daily runs, as revealed by a one-way repeated measures ANOVA with Run Order (first, second, third) as an independent factor and Efficiency as a dependent variable,  $F(2,92) = 36.67, p < .0001, \eta^2 = .072$ ; Figure 4.

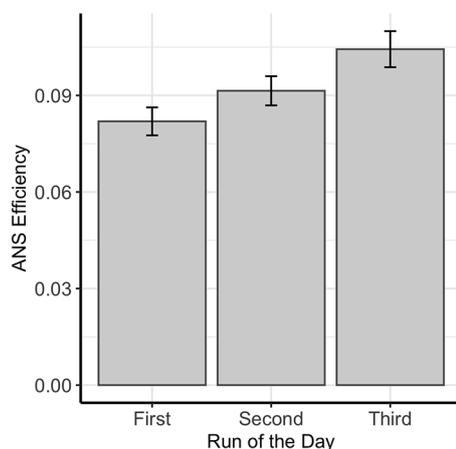


Figure 4. Efficiency (Percent Correct / RT) during the first, second and third run of each training day, showing improvement in efficiency across runs within each day.

Again, we stress that our main measure of interest is not improvement during training, but the effect of such training on school math; nevertheless, we can show that training affected children's ANS responses, at least as measured by efficiency.

Lastly, it remains an interest whether individual differences in ANS efficiency are stable over time. Here, we could test for this stability, and for improvements in ANS, by looking at regressions across testing days. Figure 5 shows a regression between ANS efficiency in the first day with each of the following testing days. The regressions were all significant (all  $p < .05$ ) and the overall regression was significant ( $R^2 = .2, p < .0001$ ). This demonstrates that individual differences in ANS efficiency were indeed stable over time.

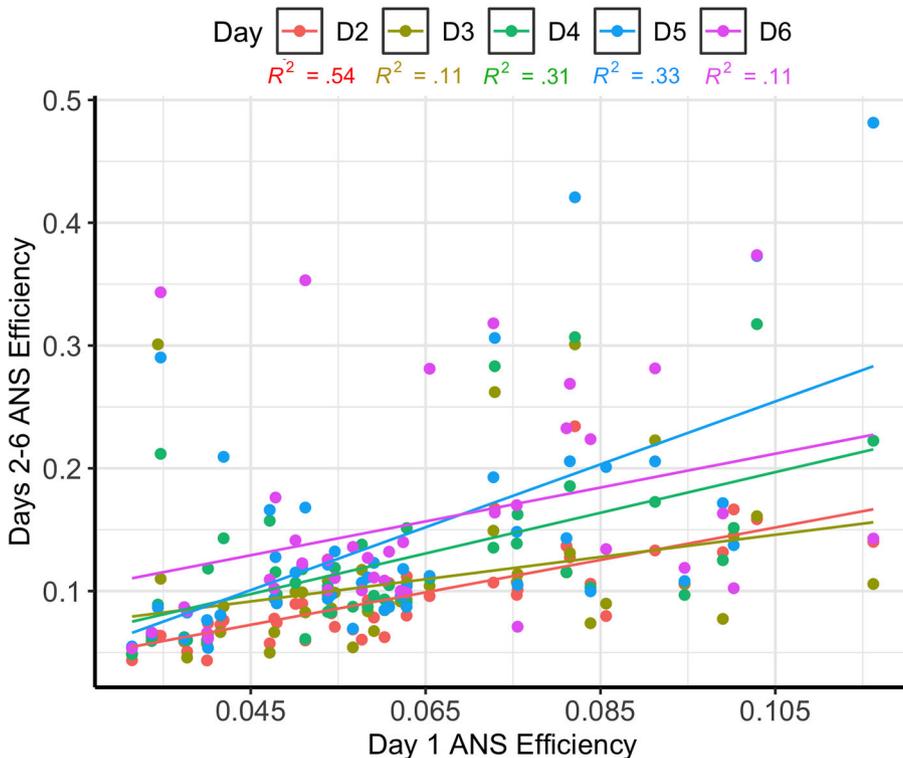
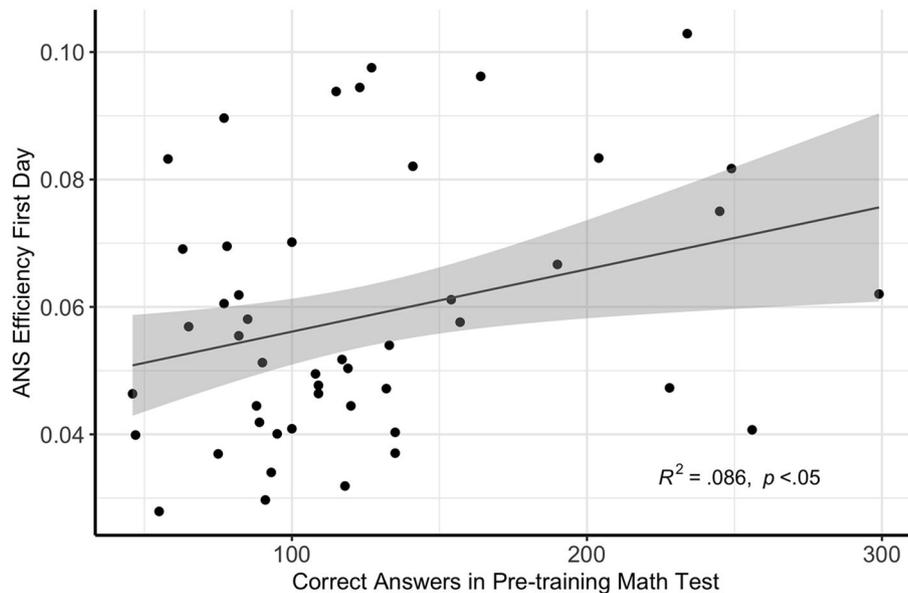


Figure 5. ANS Efficiency (Percent Correct / RT) across training days, showing stable individual differences in ANS Efficiency.

### ANS Efficiency and Pre-Training Math Performance

While still controversial, some previous research suggests that ANS performance is related to symbolic mathematics performance. Our Pre-Training assessment of Additions, Subtractions and Operations allows us to evaluate this claim for our participants. Indeed, the total number of correct answers in the Pre-Training math tests correlated with the ANS Efficiency before training,  $R^2 = .086, p < .05$ ; Figure 6. Such a correlation held also when analyzing the Pre-Training Symbolic Math Performance and the total ANS Efficiency across all training sessions. Collapsing across all 3-weeks of ANS-Training Efficiency, we again found that Symbolic Mathematics Performance in Pre-training correlated with ANS Efficiency ( $R^2 = .146, p < .01$ ). That is, the children with higher symbolic mathematics performance before training also showed a higher ANS efficiency across training. While previous demonstrations of this relationship between ANS acuity and math ability have relied on brief measures of ANS acuity (Booth & Siegler, 2006; Gilmore et al., 2010; Halberda et al., 2012, 2008; Y. He et al., 2016; Libertus et al., 2011; Libertus, Odic, & Halberda, 2012; Lyons & Beilock, 2011; Sasanguie, De Smedt,

Defever, & Reynvoet, 2012) here we extend these results to a much longer temporal interval, showing that this relationship holds even when ANS efficiency is measured across three weeks of training experience. These results suggest that the few studies that failed to find such a relation may be exceptions to the general trend (for discussion see Chen & Li, 2014; Fazio, Bailey, Thompson, & Siegler, 2014; Schneider et al., 2017).



*Figure 6.* Correlation between Approximate Number System performance before the training (measured as Efficiency during the first day of testing on the approximate number comparison task) and Symbolic Math Ability (measured as the number of correct answers in Additions, Subtractions and Operations in Pre-training). Each point represents a child ( $n = 47$ ).

## The Effect of ANS Training on Symbolic Math Performance

We now turn to our main interest: the effect of ANS training on symbolic math, or the difference between Pre- and Post-Training Symbolic Math Performance. A first rough assessment of such an effect can be obtained by collapsing into a single measure all Symbolic Math Tests (Additions, Subtractions, Operations) for all children. In a 2 X 2 ANOVA, with Training Condition (Control, ANS Training) and Phase (Pre-, Post-training) as independent variables, we found a main effect of Phase,  $F(1,89) = 111, p < .001, \eta^2 = .067$ , no effect of Training Condition,  $F(1,89) = 0.78, p = .38, \eta^2 = .008$ , and no interaction,  $F(1,89) = 1.76, p = .19, \eta^2 = .001$ ; [Figure 7](#). The data were also examined by estimating the Bayes factor, which suggested that the data were 3.33:1 in favor of the null hypothesis, or rather, 3.33 times more likely to occur under the model which only includes the main effect of Phase, rather than the model which includes the main effects of Phase and of Group, and their interaction. In sum, the Bayesian ANOVA reveals that the data provides support for the Phase main effect model, the null hypothesis, over any of the other models. Thus, when inspected globally, no difference was found between the Control and ANS Training children: they all improved in our Symbolic Math Tests during the period of the experiment. We notice that similar results were also found in a sample of Uruguayan seven-year-olds when considering all children as a group ([Langfus et al., 2019](#)). Importantly, this previous study ([Langfus et al., 2019](#)), along with others ([Bonny & Lourenco, 2013](#); [Mazzocco et al., 2011a](#)) have found stronger improvement for children who were initially low-performing in mathematics. These previous studies found these effects by splitting their samples into smaller groups based on initial math performance.

Motivated by this literature, and our own a priori interests, we planned to take this same approach here – splitting our sample into smaller groups based on pre-training arithmetic scores.

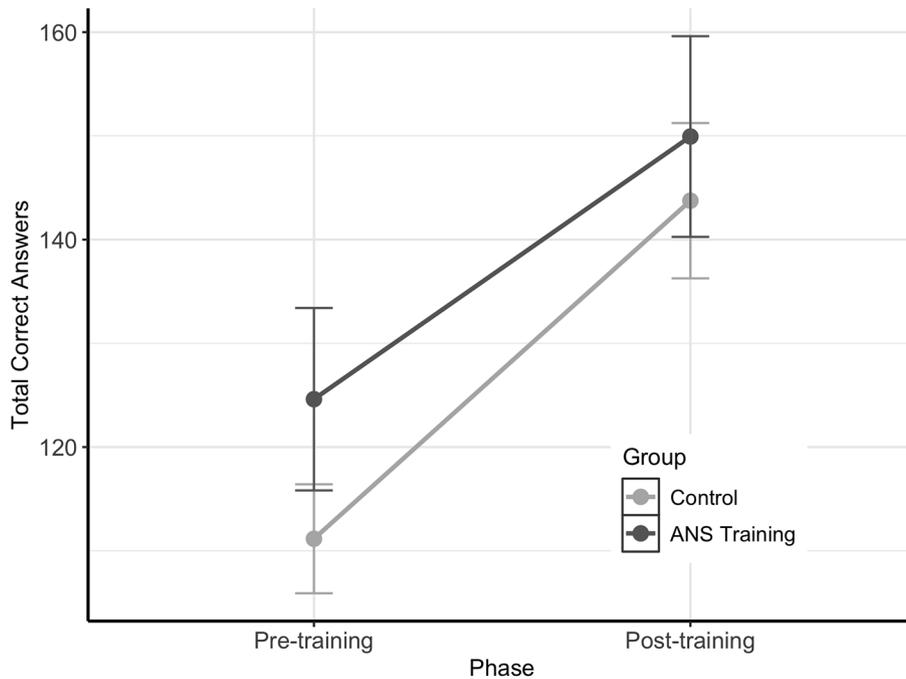


Figure 7. Total correct answers in Symbolic Math Assessment (Additions, Subtractions, and Operations) in Pre- and Post-training, for Control and ANS Training groups.

Considering the wide differences between children's mathematical abilities, and our interests in the potential effect of training on children with different starting points, we next investigated the relationship between initial math ability and improvement due to ANS training. To assess it, we inspected how the percentage of correct responses varied from pre- to post-test. We regressed number of correct answers on the Symbolic Math Test during pre-training against percentage growth in symbolic math performance from pre- to post-training. The regression showed that children who gave fewer correct answers on the pre-training Symbolic Math Test showed higher percentage gains in symbolic math ability from pre- to post-test ( $R^2 = .103, p < .01$ ; Figure 8).

This result suggested that an analysis at the level of tertiles (Below-Average, Average, Above-Average), which has been done in previous publications, may be informative. Converging evidence across several labs suggests that children who are lower-achieving in mathematics may show the strongest (or most easily detected) relationship between the ANS and school mathematics ability (Bonny & Lourenco, 2013; Langfus et al., 2019; Mazzocco et al., 2011a; Purpura & Logan, 2015; Valle-Lisboa et al., 2016). Therefore, inspired by this prior work, we next turned to investigating possible heterogeneity in the gain scores for children according to initial arithmetic performance. We split our groups into tertiles according to their results in the pre-training exact Arithmetic test, comprised of Additions and Subtractions but not the Operations subtest (Figure 9) – our novel measure which we planned to use as an outcome variable. We used the exact Arithmetic test because it includes the problems children are most familiar with and have practiced in school prior to our training.

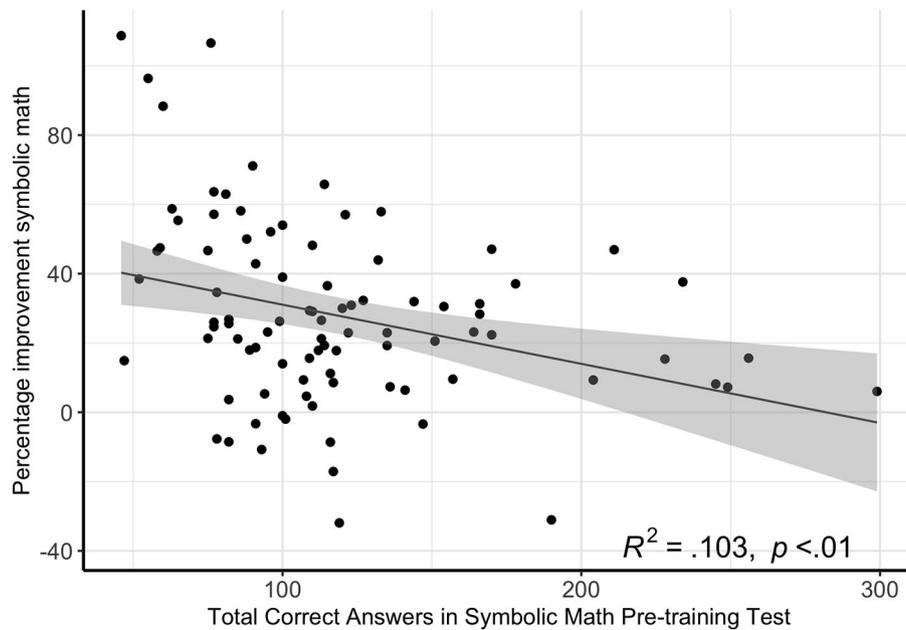


Figure 8. Percentage improvement in symbolic math for both Control and ANS Training children as a function of Total correct answers in the Symbolic Math Assessment in Pre-training.

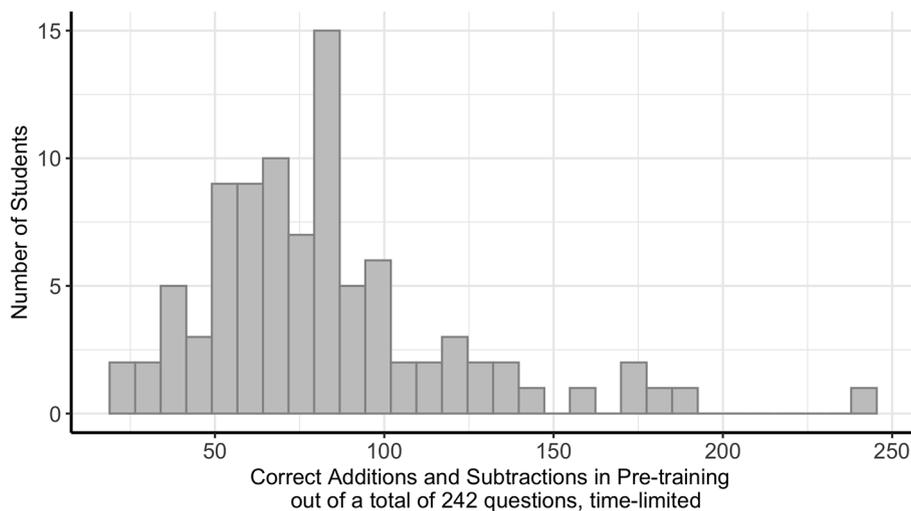


Figure 9. Histogram for correct answers in the exact Symbolic Arithmetic pre-test (Additions and Subtractions subtests).

The split into Below-Average, Average, and Above-Average tertiles regarding their initial arithmetic scores was carried out with a focus on creating roughly equal numbers of participants per group, and resulted in the following cutoffs: below 65 correct responses for the Below-Average tertile, between 66 and 86 for the Average tertile, and above 86 correct responses for the Above-Average tertile (Figure 9;  $n_{\text{Below-Control}} = 15$ ,  $n_{\text{Below-ANS}} = 15$ ,  $n_{\text{Average-Control}} = 16$ ,  $n_{\text{Average-ANS}} = 16$ ,  $n_{\text{Above-Control}} = 13$ ,  $n_{\text{Above-ANS}} = 16$ ). Notice that these tertiles are only indicating performance relative to children's own classmates on our exact Arithmetic test before the training; they may not capture a more universal notion of "low-achieving" children (Desimone & Long, 2010), nor "gifted" children (Swanson, 2006). However, given the typical breadth in the composition of our sample, and the fact that our speeded arithmetic tasks are based on the math fluency subtest of the Woodcock-Johnson Tests of

Achievement (Woodcock et al., 2001), our split is likely to capture differences valid beyond our limited sample (see also Langfus et al., 2019; Valle-Lisboa et al., 2016).

Considering that the grouping variable was measured in the pretest, we expected no differences in the number of correct answers for ANS Training and Control children in the Symbolic Math Pre-training Test. A 2 x 3 ANOVA, with Training Condition (ANS Training, Control) and Symbolic Math Level (Below-Average, Average, and Above-Average tertiles) as independent variables, showed that, as expected Symbolic Math Level differed,  $F(2,85) = 77.3, p < .001, \eta^2 = .645$ , but Training Condition did not,  $F(1,85) = 3.0, p = .087, \eta^2 = .034$ . However, an interaction between Training Condition and Symbolic Math Level,  $F(2,85) = 5.53, p < .01, \eta^2 = .115$ , appeared. Bonferroni *post hoc* tests revealed that the interaction was driven by the Above-Average tertile where the Control group obtained lower scores than the ANS Training group ( $p < .05$ ; Figure 10). While this difference appears to be a random result of grouping, it is worth bearing in mind that, because of a ceiling effect, it may mask potential ANS training benefits in the Above-Average tertile.

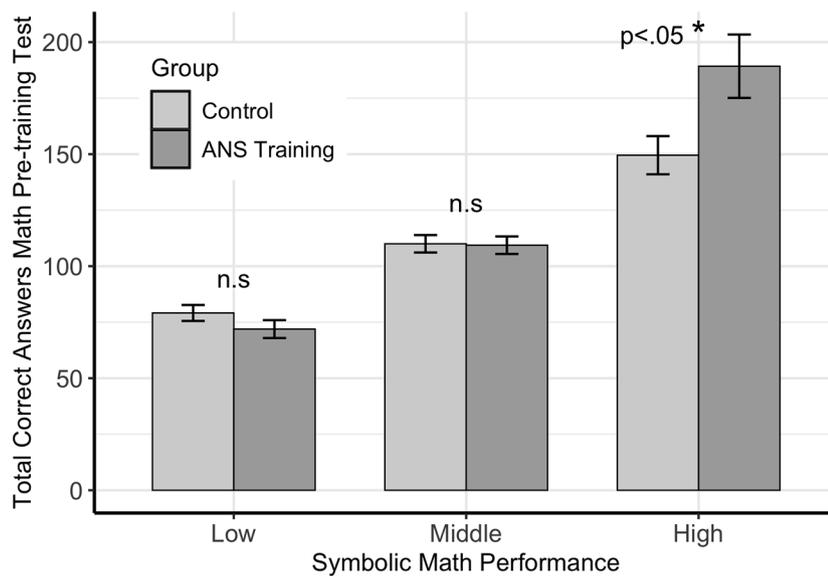


Figure 10. Total correct answers in the Symbolic Math Pre-Training test (Additions, Subtractions and Operations) for each level (Below-Average, Average, Above-Average) and for each condition (Control and ANS Training). For the Above-Average tertile there was a significant difference between Control and ANS Training groups.

We relied on the Operations subtest as our outcome variable, as it assesses how children deal with aspects of mathematical reasoning that they have not been trained to solve, and it was not used as a grouping variable for grouping our children. At the same time, we found it comparable with more traditional arithmetic tests, because it highly correlates with them both before and after training (Additions and Operations: at pre-test  $r = .68, p < .001$ , at post-test  $r = .68, p < .001$ ; Subtractions and Operations: at pre-test  $r = .64, p < .001$ , at post-test  $r = .69, p < .001$ ). In order to compare results across groups, we combined Pre- and Post-training scores into percentage change scores (correct responses in Post-training - correct responses in Pre-training)/correct responses in Pre-training. A two-way between-participants ANOVA with Symbolic Math Level (Below-Average, Average, Above-Average) and Training Condition (ANS Training, Control) as factors and percent change on the Operations subtest as the dependent variable revealed no effect of Symbolic Math Level,  $F(2,85) = 0.8, p = .45, \eta^2 = .019$ , or Training Condition,  $F(1,85) = 0.17, p = .68, \eta^2 = .002$ . However, their interaction was significant,

$F(2,85) = 8.49, p < .001, \eta^2 = .166$ ; Figure 11. Bonferroni *post-hoc* tests revealed that the change was higher for the ANS Training group than for the Control group in the Below-Average tertile of children ( $p < .05$ ), but not for the Average tertile ( $p = .41$ ). The Above-Average tertile of children showed an effect in the opposite direction, perhaps due to the ceiling effect in the ANS Training group: the Control group scored higher than the ANS Training group ( $p < .01$ ); however, as we noticed, in this tertile the ANS group was abnormally high, and better than the Control group already at Pre-Training in all tests, including the Operations test,  $t(27) = 2.09, p < .05$ . This makes it hard to interpret the results, probably contaminated by ceiling effects at baseline.

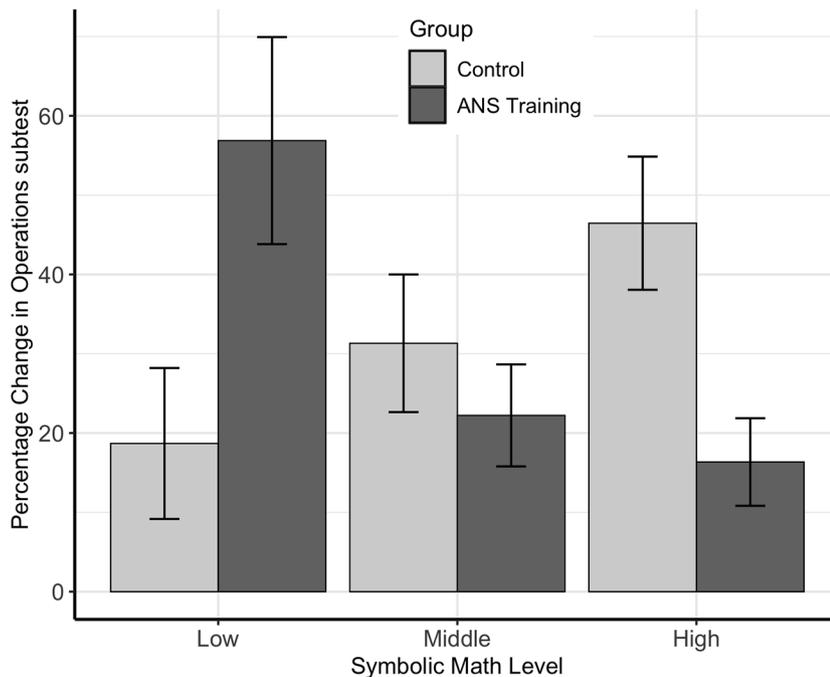


Figure 11. Percentage change in the Operations subtest by Symbolic Math Level (Below-Average, Average, Above-Average) and Training Condition (Control, ANS Training).

Considering the size of the sample (91 participants: Control  $n = 44$  and ANS Training  $n = 47$ ), we wanted to maintain all participants in order to not lose statistical power. However, with the concern that the observed interaction in the improvement in the Operations test could be driven by the initial differences before the training, specifically in the above-average tertile (Figure 10), and considering that precisely this tertile was the only one not well matched regarding the number of participants between both groups ( $n_{\text{Above-Control}} = 13$ ,  $n_{\text{Above-ANS}} = 16$ ), we proceeded to analyze all participants whose initial arithmetic scores were within 2.5 *SDs* from the mean, consequently excluding three participants that coincidentally belonged to the Above-Average tertile from the ANS Training group. This equates the number of subjects in Control ( $n = 44$ ) and ANS Training ( $n = 44$ ) groups. Thus, with the sample of now 88 participants, we tested again if there were differences in the number of correct answers for the Symbolic Math Pre-training Test. The  $2 \times 3$  ANOVA, with Training Condition (ANS Training, Control) and Symbolic Math Level (Below-Average, Average, and Above-Average tertiles) as independent variables and number of correct answers for the Symbolic Math Pre-training Test as the dependent variable, showed that Symbolic Math Level differed,  $F(2,82) = 85.5, p < .001, \eta^2 = .67$ , but Training Condition did not,  $F(1,82) = 0.77, p = .38, \eta^2 = .009$ , and there was no interaction,  $F(2,82) = 2.64, p = .08, \eta^2 = .06$ . The three levels for both groups now had an equal number of participants and non-significantly

different initial scores across Control and Training groups. Thus, we proceeded to analyze the improvement in Operations subtest. The two-way between-participants ANOVA with Symbolic Math Level (Below-Average, Average, Above-Average) and Training Condition (ANS Training, Control) as factors and percent change on the Operations subtest as the dependent variable revealed no effect of Symbolic Math Level,  $F(2,82) = 0.78$ ,  $p = .46$ ,  $\eta^2 = .019$ , or Training Condition,  $F(1,82) = 0.15$ ,  $p = .69$ ,  $\eta^2 = .002$ . However, again, their interaction was significant,  $F(2,82) = 7.88$ ,  $p < .001$ ,  $\eta^2 = .16$ . Bonferroni *post-hoc* tests once more revealed that the change was higher for the ANS Training group than for the Control group in the Below-Average tertile of children ( $p < .05$ ), but not for the Average tertile ( $p = .41$ ). The Above-Average tertile of children showed again an effect in the opposite direction: the Control group scored higher than the ANS Training group ( $p < .01$ ). In this case, both groups in the Above-Average tertile had the same number of correct answers on average in Operations post-test ( $M_{PreControl} = 40$ ,  $M_{PreANS} = 50$ ,  $M_{PostControl} = 57$ ,  $M_{PostANS} = 57$ ). Therefore, it may be reasonable to consider that the limited time (six minutes) given to the children to do the test possibly caused a ceiling effect of the number of correct problems to be solved. This prevents a detection of potentially greater improvements in the Above-Average tertile in the ANS Training group.

Although our outcome variable was the percentage of change in the Operations subtest, we also analyzed this improvement in the exact Arithmetic test (Addition and Subtraction subtests). Considering that, for the whole sample, the number of correct answers in Operation subtest was highly correlated with the number of correct answers obtained in the Addition and Subtraction subtests pre and post training, it may have been expected for the exact Arithmetic test to show similar results in the percentage of change of the below-average tertile. However, in a 2 X 2 ANOVA, with Symbolic Math Level (Below-Average, Average, Above-Average) and Training Condition (ANS Training, Control) as factors and percent change on the exact Arithmetic test as the dependent variable, showed that Symbolic Math Level differed,  $F(2,82) = 10.6$ ,  $p < .001$ ,  $\eta^2 = .20$ , but Training Condition did not,  $F(1,82) = 2.1$ ,  $p = .15$ ,  $\eta^2 = .02$ , and no interaction was found,  $F(2,82) = 1.1$ ,  $p = .34$ ,  $\eta^2 = .02$ . Thus, no difference was found between the Control and the ANS Training groups regarding the exact Arithmetic test, they all improved on exact additions and subtractions during the period of the experiment. Notice that this is specifically relevant for the Below-Average level, in which there is a significant difference of improvement between the ANS Training group and the control group but only in the Operations subtest (approximate symbolic arithmetics) and not in the exact Arithmetic test.

With a concern that our one significant improvement based on training in the below-average tertile might be a spurious result, we also checked if children showed improvement in all three problem types of the Operations subtest. Focusing on the Below-Average level and the Operations subtest, where the training seemed to be effective, we assessed if any specific equation type (additions, subtractions or multiplications) drove the observed difference between groups. A 2 X 3 mixed ANOVA, with Training Condition (ANS Training, Control) as a between-participant factor and Operation subtest equation type (addition, subtraction, multiplication) as a within-participant factor revealed a main effect of Training Condition, with ANS Training children showing greater percentage change than Control children,  $F(1,84) = 5.64$ ,  $p < .05$ ,  $\eta^2 = .06$ , but no effect of problem type,  $F(2,84) = 0.08$ ,  $p = .9$ ,  $\eta^2 = .002$ , and no interaction,  $F(2,84) = 0.27$ ,  $p = .76$ ,  $\eta^2 = .006$ . Thus, children who performed Below-Average in the pre-training arithmetic scores and had ANS training, improved in all types of equations of the Operations subtest, without any of the equation types solely driving the difference between groups. This suggests some stability of this effect, but of course, more research is needed.

In light of the obtained results, we focus our discussion on the Below-Average tertile regarding their initial arithmetic level, where training may be effective, although only in approximate symbolic arithmetics (Operations subtest) and not in exact symbolic arithmetics (exact Arithmetic test: Additions and Subtractions subtests).

## Discussion

A considerable amount of research has highlighted the importance of the ANS to math ability, although this role is still controversial and many questions need to be answered before any causal link can be claimed with high confidence. In the present study, we try to make some contributions to this ongoing discussion. Here we found severe restrictions on how and to whom training the ANS resulted in improvements of arithmetic skills. We extensively trained children on an ANS confidence-scaffolding quantity discrimination task (increasing trial difficulty during sessions). We found that training on speeded, non-symbolic, quantity comparison engages the approximate number system, as shown by the size- ratio signature of the responses, even for size-controlled and stochastic size-controlled trials. Furthermore, ANS training increased ANS efficiency both over the course of the 3-week intervention and within daily sessions. These results confirm and extend the finding that ANS can be partially trained (Knoll et al., 2016; Obersteiner et al., 2013; Park et al., 2016; Park & Brannon, 2014).

In relation to other ongoing debates, we found that initial symbolic math ability related to ANS efficiency (Figure 6), that individual differences in ANS efficiency were measurable, and that these individual differences were stable throughout training (Figure 5).

Concerning another controversy, we think it is very likely that our children were relying on some continuous extent cues (Leibovich & Henik, 2013; Leibovich et al., 2017; Merkley, Scerif, & Ansari, 2017; Szűcs & Myers, 2017), however, the resulting representations might still be numerical (Halberda, 2019).

Our main interest, though, was to assess to what extent ANS training would transfer to symbolic mathematics, specifically, arithmetics. Here, our results showed limited benefits. Overall, both Control and ANS Training children improved from Pre- to Post-training. As children's knowledge progresses over the course of a one-month class activity, this overall improvement was expected. However, we also expected to detect some training-induced improvements in the ANS training group compared to the Control group and, overall, these were not observed. Taking the intervention group as a whole, the ANS training did not help children improve their results in the exact Arithmetics test (Additions and Subtractions tests) nor in the novel Operations test. This is in fact the null hypothesis of the present study, that is, that the training did not have an overall effect. The one positive result of note was that, within the children who performed below-average in the pre-training exact Arithmetic test, training the ANS did seem to show a transfer to approximate symbolic arithmetics. This appears to be consistent with patterns seen in other samples of this age (Langfus et al., 2019; Valle-Lisboa et al., 2016). Present results suggest that those who belonged in the ANS training group and had below-average pre-training scores showed a higher improvement in the novel Operations test, which assesses approximate symbolic arithmetics. This improvement, however, was not seen in the exact Arithmetic test (Additions and Subtractions tests).

We would like to stress that the results obtained are positive solely regarding a marginal aspect of arithmetics (symbolic approximate arithmetics), and for a limited number of children (below average in the Arithmetic pre-test). Furthermore, considering the low number of participants per tertile, a word of caution is necessary

when interpreting these results. They should be taken as a suggestion since further research with larger sample sizes in each tertile is needed to support this claim.

This being said, and as an attempt at giving a possible explanation for the obtained results, we cautiously present the following considerations. Different ANS training regimens may train different components of the ANS, such as the approximate sense of cardinality, or of arithmetic transformations, or of quantity comparison, or of a sense of ratios. Other training studies with adults (Park & Brannon, 2013, 2014) and preschoolers (Park et al., 2016) improved exact symbolic arithmetic by training approximate arithmetic transformations. In those studies, the cognitive component shared by their ANS training and the symbolic tests was the arithmetic manipulation of numerical quantities. We tentatively suggest that the lack of this cognitive component in our ANS training could be an explaining factor regarding why we obtained null results in the exact symbolic arithmetic. In the present study, we instead focused on training the ANS through approximate quantity comparison, a component also present in our measure, the Operations test. This test requires a previous understanding of how basic arithmetic operations change quantities. It has the sensitivity to assesses approximate symbolic arithmetics since no exact calculation is required; a comparison of both sides of the equal sign ( $3 \square 5 = 15$ ) allows for a correct answer. Simply stated, “more” for addition, “less” for subtraction, and “much more” for multiplication. For this reason, this novel test may have detected improvements in approximate symbolic arithmetics resulting from practice with the approximate nature of our ANS training procedure. The novel Operations test as a measure of success could, therefore, constitute a strength of our study, allowing us to better detect a positive transfer, if any, from ANS comparison training to approximate symbolic arithmetics.

There are further explanations to be considered. One potential tension in our results is that we found that Pre-training Symbolic Math Test performance is significantly related to ANS performance (Figure 6) but that ANS training did not lead to greater gains in Symbolic Maths performance (e.g., Figure 7). While not unexpected – because the relationship in Figure 6 has developed over the lifetime of the child while the non-result in Figure 7 relies on just 1-month of training – this difference in results leaves open several possibilities: e.g., perhaps 1-month of training isn’t enough; or training the ANS through approximate quantity comparison is not the most beneficial type of ANS training; or the relationship in Figure 6 relies on EF while the non-relationship in Figure 7 does not; or that the ANS is important to approximate symbolic arithmetic but not to exact arithmetic per se; etc. In a way, the diverse results in the present study mirror the diversity of interpretations that are current and appropriate in literature, since we are far from understanding the mechanisms that underlie the relation between ANS and symbolic mathematics (Szkudlarek & Brannon, 2017).

We cautiously suggest that ANS quantity comparison training could help children who are struggling with arithmetics to improve some very basic aspects, namely, symbolic approximation. Given that we found some improvements only in approximate symbolic arithmetic, it remains to be studied if these could later positively impact the comprehension and execution of exact symbolic arithmetic. While approximate symbolic arithmetic is not typically taught in schools, exact symbolic arithmetic is, and it is a target outcome of math education.

In the present study, results suggest that the ANS exists, that individual differences in ANS efficiency exist and are stable, and that ANS training can improve ANS efficiency. However, we found that training the ANS, by approximate quantity comparisons, leads to very limited transfer in 7-8 year olds – only in children with low pre-training arithmetic scores and only in approximate symbolic arithmetics.

## Funding

The research was supported by a McDonnell Scholar Award to J.H.

## Competing Interests

The authors have declared that no competing interests exist.

## Acknowledgments

Special thanks to Sònia Sas, general director of Hamelin Laie International School at the time of testing, for giving access to the participants and the school facilities essential to conduct this research. We thank the teachers Mar Moreno, Teresa López Guerrero and Catherine Teti for their collaboration with the experimenter during the completion of the research. We thank Luis Morís Fernández for advice in data analysis, and Ariadna Batalla Ferrés and Luca L. Bonatti for useful advice on the manuscript.

## 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:***

- A-test-additions.pdf
- B-test-additions.pdf
- A-test-subtractions.pdf
- B-test-subtractions.pdf
- A-test-operators.pdf
- B-test-operators.pdf

***Instructions to the teachers to guarantee the correct procedure during the tests. Version in Catalan (original), and version in English (translation):***

- Instruccions\_professores\_Catala.pdf
- Teachers\_instructions\_English.pdf

***Results of the math tests from the 91 children:***

- Results\_math\_tests.csv

### Index of Supplementary Materials

Ferres-Forga, N., & Halberda, J. (2020). *Supplementary materials to "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"* [Tests, instructions, and results]. PsychOpen. <https://doi.org/10.23668/psycharchives.4351>

## References

- Abreu-Mendoza, R. A., Soto-Alba, E. E., & Arias-Trejo, N. (2013). Area vs. density: Influence of visual variables and cardinality knowledge in early number comparison. *Frontiers in Psychology, 4*, Article 805. <https://doi.org/10.3389/fpsyg.2013.00805>
- Agrillo, C., Piffer, L., & Bisazza, A. (2011). Number versus continuous quantity in numerosity judgments by fish. *Cognition, 119*(2), 281-287. <https://doi.org/10.1016/j.cognition.2010.10.022>
- 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>
- Aubrey, C., Godfrey, R., & Dahl, S. (2006). Early mathematics development and later achievement: Further evidence. *Mathematics Education Research Journal, 18*(1), 27-46. <https://doi.org/10.1007/BF03217428>
- Bartelet, D., Vaessen, A., Blomert, L., & Ansari, D. (2014). What basic number processing measures in kindergarten explain unique variability in first-grade arithmetic proficiency? *Journal of Experimental Child Psychology, 117*(1), 12-28. <https://doi.org/10.1016/j.jecp.2013.08.010>
- Beran, M. J., Evans, T. A., & Harris, E. H. (2008). Perception of food amounts by chimpanzees based on the number, size, contour length and visibility of items. *Animal Behaviour, 75*(5), 1793-1802. <https://doi.org/10.1016/j.anbehav.2007.10.035>
- 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>
- Bugden, S., & Ansari, D. (2016). Probing the nature of deficits in the "Approximate number system" in children with persistent Developmental Dyscalculia. *Developmental Science, 19*(5), 817-833. <https://doi.org/10.1111/desc.12324>
- Butterworth, B. (2011). Foundational numerical capacities and the origins of dyscalculia. In S. Dehaene & E. M. Brannon (Eds.), *Space, time and number in the brain: Searching for the foundations of mathematical thought* (pp. 249-265). <https://doi.org/10.1016/B978-0-12-385948-8.00016-5>
- Cantlon, J. F., Libertus, M. E., Pinel, P., Dehaene, S., Brannon, E. M., & Pelphrey, K. A. (2009). The neural development of an abstract concept of number. *Journal of Cognitive Neuroscience, 21*(11), 2217-2229. <https://doi.org/10.1162/jocn.2008.21159>
- 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>
- Clayton, S., Gilmore, C., & Inglis, M. (2015). Dot comparison stimuli are not all alike: The effect of different visual controls on ANS measurement. *Acta Psychologica, 161*, 177-184. <https://doi.org/10.1016/j.actpsy.2015.09.007>
- Cochrane, A., Cui, L., Hubbard, E. M., & Green, C. S. (2019). "Approximate number system" training: A perceptual learning approach. *Attention, Perception & Psychophysics, 81*(3), 621-636. <https://doi.org/10.3758/s13414-018-01636-w>
- Dehaene, S. (2009). Origins of mathematical intuitions: The case of arithmetic. *Annals of the New York Academy of Sciences, 1156*, 232-259. <https://doi.org/10.1111/j.1749-6632.2009.04469.x>
- 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](https://doi.org/10.1016/S0166-2236(98)01263-6)
- 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>
- Desimone, L., & Long, D. (2010). Teacher effects and the achievement gap: Do teacher and teaching quality influence the achievement gap between black and white and high-and low-SES Students in the Early Grades? *Teachers College Record, 112*, 3024-3073. Retrieved from [http://www.gse.upenn.edu/pdf/desimone/ECLS\\_TCR.pdf](http://www.gse.upenn.edu/pdf/desimone/ECLS_TCR.pdf)
- 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>
- Elliott, L., Feigenson, L., Halberda, J., & Libertus, M. E. (2019). Bidirectional, longitudinal associations between math ability and approximate number system precision in childhood. *Journal of Cognition and Development, 20*(1), 56-74. <https://doi.org/10.1080/15248372.2018.1551218>
- 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>
- Fuhs, M. W., & McNeil, N. M. (2016). The role of non-numerical stimulus features in approximate number system training in preschoolers from low-income homes. *Journal of Cognition and Development, 17*, 737-764. <https://doi.org/10.1080/15248372.2015.1105228>
- Fuhs, M. W., Nesbitt, K. T., & O'Rear, C. D. (2018). Approximate number system task performance: Associations with domain-general and domain-specific cognitive skills in young children. *Journal of Numerical Cognition, 4*(3), 590-612. <https://doi.org/10.5964/jnc.v4i3.141>
- Geary, D. C. (2013). Early foundations for mathematics learning and their relations to learning disabilities. *Current Directions in Psychological Science, 22*(1), 23-27. <https://doi.org/10.1177/0963721412469398>

- Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' functional numeracy is predicted by their school entry number system knowledge. *PLoS One*, *8*(1), Article e54651. <https://doi.org/10.1371/journal.pone.0054651>
- Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., . . . Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS One*, *8*(6), Article e67374. <https://doi.org/10.1371/journal.pone.0067374>
- Gilmore, C., 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>
- Gray, S. A., & Reeve, R. A. (2014). Preschoolers' dot enumeration abilities are markers of their arithmetic competence. *PLoS One*, *9*(4), Article 94428. <https://doi.org/10.1371/journal.pone.0094428>
- Halberda, J. (2019). Perceptual input is not conceptual content. *Trends in Cognitive Sciences*, *23*(8), 636-638. <https://doi.org/https://doi.org/10.1016/j.tics.2019.05.007https://doi.org/10.1016/j.tics.2019.05.007>
- 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., Mazocco, 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>
- Haworth, C. M. A., Kovas, Y., Petrill, S. A., & Plomin, R. (2007). Developmental origins of low mathematics performance and normal variation in twins from 7 to 9 years. *Twin Research and Human Genetics*, *10*(1), 106-117. <https://doi.org/10.1375/twin.10.1.106>
- He, L., Zhou, K., Zhou, T., He, S., & Chen, L. (2015). Topology-defined units in numerosity perception. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(41), E5647-E5655. <https://doi.org/10.1073/pnas.1512408112>
- 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>
- Hellgren, K., Halberda, J., Forsman, L., Adén, U., & Libertus, M. E. (2013). Compromised approximate number system acuity in extremely preterm school-aged children. *Developmental Medicine and Child Neurology*, *55*(12), 1109-1114. <https://doi.org/10.1111/dmcn.12206>
- 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. (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. (2011). Two systems of non-symbolic numerical cognition. *Frontiers in Human Neuroscience*, *5*, Article 150. <https://doi.org/10.3389/fnhum.2011.00150>
- 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>
- Hyde, D. C., & Spelke, E. S. (2011). Neural signatures of number processing in human infants: Evidence for two core systems underlying numerical cognition. *Developmental Science*, *14*(2), 360-371. <https://doi.org/10.1111/j.1467-7687.2010.00987.x>
- 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>
- Jones, S. M., & Brannon, E. M. (2012). Prosimian primates show ratio dependence in spontaneous quantity discriminations. *Frontiers in Psychology*, *3*, Article 550. <https://doi.org/10.3389/fpsyg.2012.00550>
- Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, *45*(3), 850-867. <https://doi.org/10.1037/a0014939>
- Knoll, L. J., Fuhrmann, D., Sakhardande, A. L., Stamp, F., Speekenbrink, M., & Blakemore, S.-J. (2016). A window of opportunity for cognitive training in adolescence. *Psychological Science*, *27*(12), 1620-1631. <https://doi.org/10.1177/0956797616671327>
- Knops, A., Dehaene, S., Berteletti, I., & Zorzi, M. (2014). Can approximate mental calculation account for operational momentum in addition and subtraction? *Quarterly Journal of Experimental Psychology*, *67*(8), 1541-1556. <https://doi.org/10.1080/17470218.2014.890234>
- 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). <https://doi.org/10.1016/B978-0-12-815952-1.00002-5>
- 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., & Henik, A. (2013). Magnitude processing in non-symbolic stimuli. *Frontiers in Psychology*, *4*, Article 375. <https://doi.org/10.3389/fpsyg.2013.00375>
- 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>
- Li, H., Zhang, M., Wang, X., Ding, X., & Si, J. (2018). The central executive mediates the relationship between children's approximate number system acuity and arithmetic strategy utilization in computational estimation. *Frontiers in Psychology*, *9*, Article 943. <https://doi.org/10.3389/fpsyg.2018.00943>

- Liang, Y., Zhang, L., Long, Y., Deng, Q., & Liu, Y. (2020). Promoting effects of RtI-based mathematical play training on number sense growth among low-SES preschool children. *Early Education and Development, 31*(3), 335-353. <https://doi.org/10.1080/10409289.2019.1664261>
- 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. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science, 14*, 1292-1300. <https://doi.org/10.1111/j.1467-7687.2011.01080.x>
- Libertus, M. E., Feigenson, L., & Halberda, J. (2013). 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., & Landau, B. (2014). Understanding the mapping between numerical approximation and number words: Evidence from Williams syndrome and typical development. *Developmental Science, 17*(6), 905-919. <https://doi.org/10.1111/desc.12154>
- 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>
- Libertus, M. E., Odic, D., & Halberda, J. (2012). Intuitive sense of number correlates with math scores on college-entrance examination. *Acta Psychologica, 141*(3), 373-379. <https://doi.org/10.1016/j.actpsy.2012.09.009>
- Lindskog, M., & Winman, A. (2016). No evidence of learning in non-symbolic numerical tasks – A comment on Park and Brannon (2014). *Cognition, 150*, 243-247. <https://doi.org/10.1016/j.cognition.2016.01.005>
- 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>
- Malone, S. A., Burgoyne, K., & Hulme, C. (2020). Number knowledge and the approximate number system are two critical foundations for early arithmetic development. *Journal of Educational Psychology, 112*(6), 1167-1182. <https://doi.org/10.1037/edu0000426>
- Malone, S. A., Pritchard, V. E., Heron-Delaney, M., Burgoyne, K., Lervåg, A., & Hulme, C. (2019). The relationship between numerosity discrimination and arithmetic skill reflects the approximate number system and cannot be explained by inhibitory control. *Journal of Experimental Child Psychology, 184*, 220-231. <https://doi.org/10.1016/j.jecp.2019.02.009>
- 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>

- Merkley, R., Matejko, A. A., & Ansari, D. (2017). Strong causal claims require strong evidence: A commentary on Wang and colleagues. *Journal of Experimental Child Psychology*, *153*, 163-167. <https://doi.org/10.1016/j.jecp.2016.07.008>
- Merkley, R., Scerif, G., & Ansari, D. (2017). What is the precise role of cognitive control in the development of a sense of number? *Behavioral and Brain Sciences*, *40*, Article e179. <https://doi.org/10.1017/S0140525X1600217X>
- Moore, A. M., VanMarle, K., & Geary, D. C. (2016). Kindergartners' fluent processing of symbolic numerical magnitude is predicted by their cardinal knowledge and implicit understanding of arithmetic 2 years earlier. *Journal of Experimental Child Psychology*, *150*, 31-47. <https://doi.org/10.1016/j.jecp.2016.05.003>
- Morgan, P. L., Farkas, G., & Maczuga, S. (2015). Which instructional practices most help first-grade students with and without mathematics difficulties? *Educational Evaluation and Policy Analysis*, *37*(2), 184-205. <https://doi.org/10.3102/0162373714536608>
- Murphy, M. M., Mazzocco, M. M. M., Hanich, L. B., & Early, M. C. (2007). Cognitive characteristics of children with Mathematics Learning Disability (MLD) vary as a function of the cutoff criterion used to define MLD. *Journal of Learning Disabilities*, *40*(5), 458-478. <https://doi.org/10.1177/00222194070400050901>
- Mussolin, C., Nys, J., Content, A., & Leybaert, J. (2014). Symbolic number abilities predict later approximate number system acuity in preschool children. *PLoS One*, *9*(3), Article e91839. <https://doi.org/10.1371/journal.pone.0091839>
- 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>
- Nosworthy, N., Bugden, S., Archibald, L., Evans, B., & Ansari, D. (2013). A two-minute paper-and-pencil test of symbolic and nonsymbolic numerical magnitude processing explains variability in primary school children's arithmetic competence. *PLoS One*, *8*(7), Article e67918. <https://doi.org/10.1371/journal.pone.0067918>
- Núñez, R. E. (2017, June). Is there really an evolved capacity for number? *Trends in Cognitive Sciences*, *21*(6), 409-424. <https://doi.org/10.1016/j.tics.2017.03.005>
- Obersteiner, A., Reiss, K., & Ufer, S. (2013). How training on exact or approximate mental representations of number can enhance first-grade students' basic number processing and arithmetic skills. *Learning and Instruction*, *23*(1), 125-135. <https://doi.org/10.1016/j.learninstruc.2012.08.004>
- 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., Lisboa, J. V., 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>
- Papay, J. P., Murnane, R. J., & Willett, J. B. (2014). High-school exit examinations and the schooling decisions of teenagers: Evidence from regression-discontinuity approaches. *Journal of Research on Educational Effectiveness*, *7*(1), 1-27. <https://doi.org/10.1080/19345747.2013.819398>

- 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>
- Park, J., & Brannon, E. M. (2016). How to interpret cognitive training studies: A reply to Lindskog & Winman. *Cognition*, *150*, 247-251. <https://doi.org/10.1016/j.cognition.2016.02.012>
- Parsons, S., & Bynner, J. (2005). *Does numeracy matter more?* London, United Kingdom: National Research and Development Centre for Adult Literacy and Numeracy. Retrieved from [http://dera.ioe.ac.uk/22317/1/doc\\_2979.pdf](http://dera.ioe.ac.uk/22317/1/doc_2979.pdf)
- Peters, E., Hart, P. S., Tusler, M., & Fraenkel, L. (2014). Numbers matter to informed patient choices: A randomized design across age and numeracy levels. *Medical Decision Making*, *34*(4), 430-442. <https://doi.org/10.1177/0272989X13511705>
- 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>
- 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>
- 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>
- Prado, J., Mutreja, R., Zhang, H., Mehta, R., Desroches, A. S., Minas, J. E., & Booth, J. R. (2011). Distinct representations of subtraction and multiplication in the neural systems for numerosity and language. *Human Brain Mapping*, *32*(11), 1932-1947. <https://doi.org/10.1002/hbm.21159>
- Praet, M., & Desoete, A. (2014). Enhancing young children's arithmetic skills through non-intensive, computerised kindergarten interventions: A randomised controlled study. *Teaching and Teacher Education*, *39*, 56-65. <https://doi.org/10.1016/j.tate.2013.12.003>
- 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>
- Rips, L. J., Bloomfield, A., & Asmuth, J. (2008). From numerical concepts to concepts of number. *Behavioral and Brain Sciences*, 31(6), 623-687. <https://doi.org/10.1017/S0140525X08005566>
- Ritchie, S. J., & Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychological Science*, 24(7), 1301-1308. <https://doi.org/10.1177/0956797612466268>
- 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>
- Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number-space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology*, 114(3), 418-431. <https://doi.org/10.1016/j.jecp.2012.10.012>
- 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>
- Shusterman, A., Slusser, E., 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>
- 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>
- 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>
- Soltész, F., Szűcs, D., & Szűcs, L. (2010). Relationships between magnitude representation, counting and memory in 4- to 7-year-old children: A developmental study. *Behavioral and Brain Functions*, 6(1), Article 13. <https://doi.org/10.1186/1744-9081-6-13>
- 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>
- Swanson, J. D. (2006). Breaking through assumptions about low-income, minority gifted students. *Gifted Child Quarterly*, 50(1), 11-25. <https://doi.org/10.1177/001698620605000103>
- Szkudlarek, E., & Brannon, E. M. (2017). Does the approximate number system serve as a foundation for symbolic mathematics? *Language Learning and Development*, 13(2), 171-190. <https://doi.org/10.1080/15475441.2016.1263573>

- Szűcs, D., & Myers, T. (2017). A critical analysis of design, facts, bias and inference in the approximate number system training literature: A systematic review. *Trends in Neuroscience and Education*, 6, 187-203. <https://doi.org/10.1016/j.tine.2016.11.002>
- Tosto, M. G., Petrill, S. A., Halberda, J., Trzaskowski, M., Tikhomirova, T. N., Bogdanova, O. Y., . . . Kovas, Y. (2014). Why do we differ in number sense? Evidence from a genetically sensitive investigation. *Intelligence*, 43(100), 35-46. <https://doi.org/10.1016/j.intell.2013.12.007>
- 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 Herwegen, J., Costa, H. M., & Passolunghi, M. C. (2017). Improving approximate number sense abilities in preschoolers: PLUS games. *School Psychology Quarterly*, 32(4), 497-508. <https://doi.org/10.1037/spq0000191>
- 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>
- Wang, J., Odic, D., Halberda, J., & Feigenson, L. (2017). Better together: Multiple lines of evidence for a link between approximate and exact number representations: A reply to Merkley, Matejko, and Ansari. *Journal of Experimental Child Psychology*, 153, 168-172. <https://doi.org/10.1016/j.jecp.2016.09.005>
- Wang, Z., Hart, S. A., Kovas, Y., Lukowski, S., Soden, B., Thompson, L. A., . . . Petrill, S. A. (2014). Who is afraid of math? Two sources of genetic variance for mathematical anxiety. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 55(9), 1056-1064. <https://doi.org/10.1111/jcpp.12224>
- Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher*, 43(7), 352-360. <https://doi.org/10.3102/0013189X14553660>
- 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>
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement: (WJ-III)*. Rolling Meadows, IL, USA: Riverside Publishing.