Received: 7 February 2022 | Revised: 17 November 2022 | Accepted: 27 January 2023

DOI: 10.1111/desc.13386

Developmental Science WILEY

RESEARCH ARTICLE

Number sense at 12 months predicts 4-year-olds' maths skills

Gisella Decarli[1] | Donatella Zingaro[2] | Luca Surian[1] | Manuela Piazza[2]

[1]Department of Psychology and Cognitive Science, University of Trento, Trento, Italy

[2]Center for Mind/Brain Sciences, University of Trento, Italy

Correspondence
Manuela Piazza, Center for Mind/Brain Sciences, University of Trento, Italy.
Email: manuela.piazza@unitn.it;

Gisella Decarli, Department of Psychology and Cognitive Science, University of Trento, Italy,
Email: decarli.gisella@gmail.com

Abstract

Preverbal infants spontaneously represent the number of objects in collections. Is this 'sense of number' (also referred to as Approximate Number System, ANS) part of the cognitive foundations of mathematical skills? Multiple studies reported a correlation between the ANS and mathematical achievement in children. However, some have suggested that such correlation might be mediated by general-purpose inhibitory skills. We addressed the question using a longitudinal approach: we tested the ANS of 60 12 months old infants and, when they were 4 years old (final $N = 40$), their symbolic math achievement as well as general intelligence and inhibitory skills. Results showed that the ANS at 12 months is a specific predictor of later maths skills independent from general intelligence or inhibitory skills. The correlation between ANS and maths persists when both abilities are measured at four years. These results confirm that the ANS has an early, specific and longstanding relation with mathematical abilities in childhood.

KEYWORDS

approximate number system, inhibition, longitudinal study, maths achievement, maths skills, number sense

Research Highlights

- In the literature there is a lively debate about the correlation between the ANS and maths skills.
- We longitudinally tested a sample of 60 preverbal infants at 12 months and rested them at 4 years (final sample of 40 infants).
- The ANS tested at 12 months predicted later symbolic mathematical skills at 4 years, even when controlling for inhibition, general intelligence and perceptual skills.
- The ANS tested at 4 years remained linked with symbolic maths skills, confirming this early and longstanding relation in childhood.

1 INTRODUCTION

Representing the approximate number of objects in the environment is a core and phylogenetically ancient ability, referred to as 'Approximate Number System', or ANS, which humans display soon after birth and share with many non-human species as it is highly adaptive for survival (Dehaene, 2001; Nieder, 2021). Thanks to this sense of number, preverbal infants spontaneously compare and mentally combine sets in the form of approximate proto-arithmetical operations well before the process of enculturation allows them to use symbols to represent, compare and combine exact numbers.

It has been proposed that in humans the ANS acts as a foundational building block for the acquisition of symbolic numerical and mathematical skills, providing the young learners with a sort of domain-specific 'start-up kit' (Butterworth, 2018; Dehaene & Cohen, 2007; Piazza, 2010; see Chen & Li, 2014 and Schneider et al., 2017 for meta-analyses assessing the link between ANS and math). According to this idea young children, during the first stages of learning number words and mental arithmetic, resort to their pre-existing intuitions of the laws governing quantities and their transformations and use them to make sense of the novel cultural acquisitions. Proving this proposal true has important theoretical significance but also potential societal implications: it would imply that the integrity of the ANS in young children could provide an objective early behavioural marker to guide educational intervention and detect early risks for dyscalculia (a specific learning disability in the maths area) thus targeting early remediation.

However, this proposal remains rather controversial. On one side the empirical data supporting this hypothesis, mainly based on the report of specific correlations between the ANS and formal maths, are not always consistent (e.g., see Halberda et al., 2008; Libertus et al., 2013; Price et al., 2012; as examples of positive findings, and e.g., Sasanguie et al., 2013 for failure to observe it; see also De Smedt et al., 2013 and Szűcs & Myers, 2017 for reviews that show the weakness of ANS role in accounting for math achievement). On the other hand, an alternative account posits that the performance in tasks assessing the ANS (numerosity comparison, where participants are asked to compare two sets of objects and indicate the more numerous) mostly reflects the ability to inhibit the images' visual features. Indeed, the characteristics of the stimuli (such as size, convex hull, density etc.) are inevitably confounded with numerosity and covary with the different quantities that are represented. In line with this, performances in number comparison tasks has been found to be affected by the level of congruency between the continuous visual features and number (Clayton & Gilmore, 2015; Gilmore et al., 2013). Importantly, children with Developmental Dyscalculia (DD) seem to differ from controls especially, if not only, on incongruent trials (Bugden & Ansari, 2016; Piazza et al., 2018).

On the bases of these observations some authors proposed the hypothesis that inhibitory control could be the mediator of the correlation found between ANS acuity and mathematical scores (Gilmore et al., 2013). According to this hypothesis the ability to exert inhibitory control would be crucial both for the dots comparison task (for suppressing visual irrelevant information in favour of number information), and for the formal maths tasks (for performing complex calculation). Thus, the correlation observed between ANS and symbolic math task would be fully dependent upon the common role of inhibition. However, there is some evidence against this hypothesis (e.g., Castaldi et al., 2018; Malone et al., 2019). Keller and Libertus (2015), for example, found that the correlation between children's accuracy in the non-symbolic comparison task and their mathematical scores in a standardized test persisted also when controlling for inhibitory control capacity. Therefore, it is still an open question whether mathematical acquisition is supported by the ANS, by more domain-general inhibitory skills, or both.

The only truly conclusive piece of evidence attesting the foundational role of the ANS in formal maths rests on the observations that the ANS acuity in young pre-verbal individuals (where the effects of enculturation can be readily excluded) is a specific and reliable longitudinal predictor of their later symbolic mathematical achievement. To date, however, there is only one published study that used this approach (Starr et al., 2013). In this study, 6 months old infants' ANS acuity was found to be a specific longitudinal predictor of the performance in early symbolic numerical abilities at 3.5 years of age.

Given the large theoretical and potentially societal impact of this issue, and given the current replication crisis in psychological science, it is extremely important to verify (and potentially extend) the validity of this initial observation. Moreover, given the current debate on whether inhibitory skills are the main drivers of the correlation between ANS and maths achievement and that the aforementioned longitudinal study did not tackle the potential role of inhibitory skills, a new study is needed to incorporate this important issue.

Here we report a longitudinal study where we tested 12-month-old pre-numerate infants on an implicit number sense task and a perceptual face recognition task. This control allows us to assess the specificity of the possible correlation between numerosity perception and math, that is not guided by more general perceptual abilities. Notably and differently from Starr et al. (in this study, half infants were assessed with a control color detection task and the other half with a control size detection task), we used the same control task for all participants, thus allowing participants' assessment in the same non-numerical perceptual capacity. Infants were then re-evaluated at 4 years of age on a set of non-symbolic and symbolic formal math skills as well as general processing skills including general intelligence and inhibitory abilities. The main aim of this study was to verify whether, using participants from a different linguistic and cultural background, of a different age range and with a different set of stimuli and tasks we could replicate the important initial observation of a positive and specific longitudinal correlation between early ANS and later formal math skills, to test the specificity of this link and to verify the stability of this correlation when controlling for general-purpose inhibitory skills.

2 MATERIALS AND METHODS

2.1 Participants

A group of sixty Caucasian infants was tested in a first session (hereafter referred to as T1) (M = 12 months and 4 days; SD = 24 months

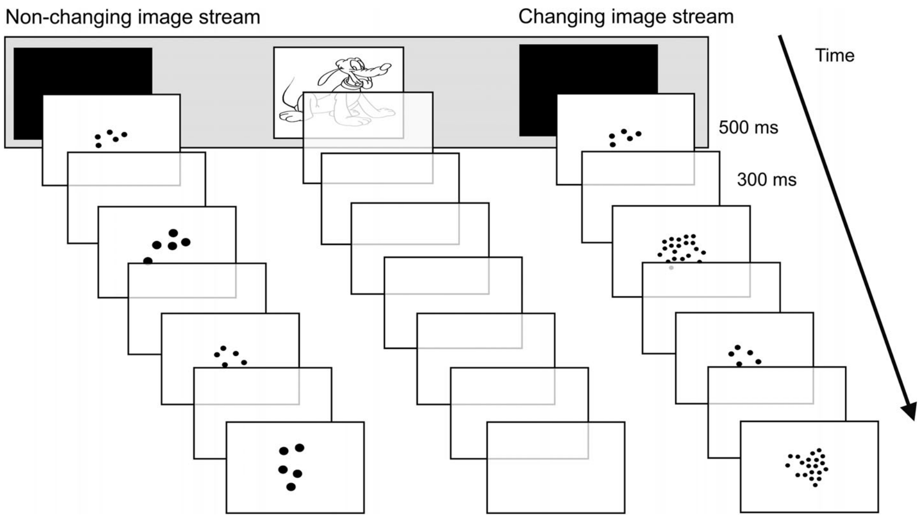

FIGURE 1 Schema of number change detection paradigm used at T1. One of the two lateral streams (here the right one) showed images that change in numerosity, while the other stream (here the left one) showed the same numerosity over time.

and 4 days; age range: 11 months and 17 days–14 months; 27 girls), which took place between winter 2016 and summer 2017. The size of the sample was determined by reference to Starr et al. (whose final sample was of 48 infants) considering the drop-out that characterises longitudinal studies with young children. We took this study as a reference point since it assessed for the first time the possible predictive role of ANS in math and since it was the first attempt to disentangle the directionality of this relation. Data at T1 from 10 infants were excluded due to fussiness ($N = 9$) or mother interference ($N = 1$), and 10 additional families were unable to come back to the Lab 3 years later for the second testing session of the longitudinal design (hereafter referred to as T2). T2 tests took place in summer 2020 and a final sample of 40 children was included in the analyses ($M = 51$ months and 3 days; $SD = 2$ months and 5 days; age range 47 months and 6 days to 58 months and 9 days; 19 girls). We are aware that the sample size could have been larger, but there are some elements to consider. First, we set a long-term study, passing 3 years between T1 and T2 which increased the number of dropouts, normally occurring in this kind of studies and requiring a high commitment by the participants. Second, the T2 of the study took place during the pandemic period, leading to even more problems for the participants' recruitment.

For both experimental sessions parents gave written informed consent. The study was approved by the Ethical Committee of the University of Trento.

2.2 | Procedure at T1

At T1, infants were tested with the visual change detection paradigm (Libertus & Brannon, 2010). They were shown four trials with numerical stimuli and four trials of face stimuli, the latter serving as a control.

The structure and timing of the two tasks were identical. Infants sat on a parent lap in front of a desk (approximately 100 cm) with three 17-inch monitors placed next to each other (screen resolution: 1280 × 1024) and partially covered by a rectangular black cardboard that only revealed the screens (see Figure 1). A small webcam was placed on the central monitor. The webcam was attached to the experimenter's computer located behind the desk where the experiment took place, and hidden by a thick curtain. Parents were blind to the scope of the study and instructed not to look/point at the screens. Infants observed four trials for the numerical change detection and four for the face change detection. At the beginning of each trial an attention-getter appeared in the central monitor and the experimenter started the trial as soon as the infant looked at it. The central screen turned black and remained so for the entire duration of the trial, while the two lateral screens started displaying streams of visual stimuli concurrently. One of the two screens displayed the same content over time (hereafter referred to as 'the non-changing stream'), while the other presented two different contents in alternating order ('the changing stream'). The side of the changing stream alternated across trials and the order was counterbalanced across participants. A trial consisted in a sequence of 20 stimuli (each presented for 500 ms and followed by 300 ms of blank), and lasted 2.6 min. Infants' fixations were recorded online by an expert observer. A second observer coded offline infants' fixations. The average inter-observer reliability was high ($r = 0.95$).

2.2.1 | Numerical change detection task

In this task the stimuli were sets of dots, black on a white background. In the non-changing stream, the sets always contained five dots while in the changing stream they alternated between two numerosities. There were four trials, presented in the same order: in the first two trials, the changing stream alternated images of 5 and 20 dots (1:4 ratio) and in the last two images were of 5 and 10 dots (1:2 ratio). For each ratio stimuli were generated randomly assigning perceptual variables to each image such that, on average, the size of the dots in the changing sets was held constant across numerosities on half the images and the total occupied area of the dots was held constant on the other half. In the non-changing sets, these parameters varied simultaneously such that, across images, they covered all values assigned to the paired non-changing sets (these two controls were intermingled across trials; stimuli constructed using the same logic as in Dehaene et al., 2005; Piazza et al., 2013). For each infant, we calculated a preference score (as in Libertus & Brannon, 2010) by subtracting the average percentage of looking time to the non-changing image stream from the percent looking time to the changing image stream. We performed t-tests comparing the preference scores with zero for each level of difficulty. In the 1:4 ratio infants preferred to look at the changing image stream compared to the non-changing (M = 0.17, SD = 0.22) and this preference was significantly higher than zero ($t(38) = 4.82$; $p < 0.001$). However, we did not observe a significant preference for the changing stream in the 1:2 ratio (M = 0.04, SD = 0.21; $t(36) = 1.24$; $p = 0.22$). This indicates that the stimuli in this condition were too difficult to allow infants to detect a numerical difference across the streams. This seems at odds with previous studies where 1:2 ratios were found to be readily discriminable in 6 months old infants (Libertus & Brannon, 2010). One explanation for this discrepancy could be that we used a different method to generate the stimuli. Indeed, the images in the non-changing image stream could vary in size or total occupied area within the same trial, making the sequence potentially more difficult to elaborate. Alternatively, it is possible that infants were less attentive in the 1:2 condition. Indeed, the trials of the 1:2 condition were always presented after the trials of the 1:4 condition. Considering these results, to compute further analyses, we selected only the 1:4 ratio for the number detection.

2.2.2 | Non-numerical change detection task

This task was identical to the numerical change detection except that the stimuli were images of child faces (male or female, alternating across trials, with neutral expressions; see Supplementary Materials). The pictures were grey-scaled, equalised in luminance and cropped with an oval sized 12.5 cm × 17.1 cm. While the non-changing stream displayed the same identical face, the changing stream displayed two alternating faces of the same gender but whose key elements (eyes, nose and mouth) differed either in their shape ('featural change') or in their relative position ('second-order change'). Thus, also for this task, we presented two levels of difficulty (see Mondloch et al., 2002; Quinn et al., 2013). For the second-order change, starting from each original face, we generated a new one where the eyes were moved 1.2 cm further apart and the mouth was moved 1.2 cm further down (an approximate 2% change in spatial separation; see Mondloch et al., 2002). Infants were presented with two trials of featural change and two trials of second-order change. We calculated for each infant a preference score as used in the numerical change detection task and we found similar results: in the first order change (featural change) trials, infants looked longer at the changing ($M = 0.1$, $SD = 0.2$; $t(37) = 2.99$; $p = 0.005$) compared to the non-changing stream, whereas in the second order (relative position change) trials they did not show a significant preference for the changing stream (M = 0.005, SD = 0.18; $t(38) = 0.17$; $p = 0.87$). Here the feature change trials were always presented before the second order trials, leaving open the possibility that the failure to observe a preference for the second order condition was due to drop of attention. To compute further analysis, we thus selected the first order condition for the face detection.

2.3 | Procedure at T2

At T2, children completed seven different tasks (4 maths-related, and 3 non-maths related, see Table 1). Data collection at T2 occurred during the COVID-19 pandemic and was performed following the 'Operational protocol to fight and contain the spread of the Sars-Cov-2 in the workplace' at University of Trento. Task order was randomized across children. When completing all the tasks children received a little present as a compensation for their participation and a certificate of participation.

2.3.1 | Maths-related tasks

Numerosity comparison

This computerised task was modelled following Piazza et al. (2010). The child sat on a chair 50 cm distant from the laptop (screen resolution 1368 × 720). Stimuli were pairs of arrays of black dots displayed in two white discs on either side of a central white fixation point. Children were asked to choose, by pointing, and as quickly as possible, to the array that contained more dots. The child's response was recorded by the experimenter who pressed the corresponding left or right response-key on the computer keyboard. The images remained on the screen until the experimenter entered the child's response. The pair of images that were displayed always had one set (n1) which contained a constant number of items (either 16 or 32). The other set (n2) could include 12, 13, 14, 15, 17, 18, 19, 20 dots when n1 was 16 and twice these values when n1 was 32. Perceptual variables were randomly assigned to each stimulus pair such that, on average, the size of the dots in the n2 array was held constant on half the trials and the total occupied area of the dots in the n2 array was held constant on the other half; in the n1 arrays, these parameters varied simultaneously such that, across trials, they covered all values assigned to the different n2 arrays (see Piazza et al., 2013). The experiment began with 4 training trials, where feedback was given to help children understand the task, followed by 64 trials, divided in 4 blocks. As an index of the ANS acuity for each child we recovered the internal Weber fraction w by fitting their individual psychometric curve for the 16 and 32 references with a single sigmoid function of the log n1/n2 ratio (see Dehaene, 2007).

Spontaneous focus on numerosity (SFON)

To test the extent to which children spontaneously focus on numerosity we used a modified version of the SFON task (SFON; Hannula &

TABLE 1 Schema of the tasks used for assessing participants at T1 (12 months) and at T2 (4 years).

12 months			4.5 years		
Construct	*Test*	*Measure*	*Construct*	*Test*	*Measure*
ANS acuity	Numerical change detection task	Numerical preference score	ANS acuity	Non-symbolic number comparison task	Weber fraction (w)
			Symbolic Math ability	TEDI-MATH	Standardized math score
			Counting knowledge	Give-A-Number task	Number knowledge
			Spontaneous focus on number	SFON task	Accuracy
Perceptual discrimination ability	Non-Numerical change detection task (face)	Non-Numerical preference score	General Intelligence	WPPSI-IV	IQ score
			Inhibitory skills	Statue (NEPSY-II)	Inhibition score
			Face processing skills	CFMT-C	Face recognition score

Lehtinen, 2005), which consists of a figure copy task. We presented children with three sheets each displaying an animal (a dinosaur, a ladybird and a zebra). The figures were randomly presented to children. Each sheet displayed a figure on the left and the same figure but with some missing parts on the right. Participants were told to complete the figure on the right so that the two looked exactly the same. The dinosaur had two missing spines, the zebra three missing stripes and the ladybirds four missing dots. Accuracy was calculated by dividing each child's final score for the maximum score of the task (max = 3; one point was given for each figure).

Give a number (GaN)

This task was modelled following Le Corre and Carey (2007). A box containing 10 identical small plastic dog toys was placed on the table in front of the child. On the first trial the child was asked to take two dogs out of the box and put them on the table. In the case the child's success, the experimenter asked for a larger number (i.e., the following trial requested was $X + 1$). If the child failed, $X - 1$ toys were requested on the subsequent trial. The trials proceeded in the order 2,3,4,5,6,7 and 8 until the child answered correctly at least twice for N and failed at least twice for $N + 1$, or until the child successfully provided eight toys. The largest number correctly processed determined his/her knower-level (from 2 to 8).

Symbolic maths processing (Tedi-Math)

Children's formal mathematical ability was assessed through six subtests taken from the TEDI-MATH (Grégoire et al., 2004, Italian adaptation), a standardised battery that evaluates different aspects of early maths achievement. Because we wanted to get a pure measure of the knowledge and use of symbolic numbers we excluded those subtests that tackle non-symbolic numerical skills (e.g. approximate number comparison). The selected sub-tests tested the following knowledge: knowledge of the verbal number sequence (verbally reciting numbers; test 1.d and 1.f), of the counting principles (counting visual items; test 2.c and 2.b2), of the semantic of number words (comparing oral number words; test 3.b3), and of basic arithmetic (adding and subtracting visual items; test 5.a) (more details on these tests can be found in the Table S1). The final score for each child was calculated by dividing the sum of the child's scores on the subtests by the sum of the maximum scores on the subtests.

2.3.2 | Non-maths related tasks

General Intelligence

To assess general intelligence we presented children with one verbal (Receptive Vocabulary) and one nonverbal (Cube Design) subtest of the standardised WPPSI-IV test (Wechsler, 2012), and we calculated a composite IQ score for each child.

Inhibitory skills

To assess general inhibitory skills we used the subtest 'Statue' from the 'Attention and Executive Functions' section of the standardised test NEPSY-II (Korkman et al., 2007). Children were asked to maintain a specific position without moving the body, eye, or mouth for 75 s and ignore the experimenter's sound distractors occurring at predetermined time intervals. According to the level of adherence to the instructions, their inhibitory skills were quantified based on the conversion of the errors in a norm-referenced scaled score, with the higher score indicating a better inhibition behaviour (see Supplementary Materials).

The rationale behind this test's choice was to assess inhibitory skills with a standardized measure that has been validated on a big sample of children and that presents a high reliability. Moreover, this task is language-free, allowing us to measure inhibitory skills and excluding the possible interference of the child's linguistic level of knowledge. Therefore, with this task, we aimed to select a pure measure of inhibitory capacity.

Face recognition

To assess face processing skills we used the standardised Cambridge Face Memory Test for Children (CFMT-C; Croydon et al., 2014) that is essentially a face recognition memory test, and computed accuracy for each child.

3 | RESULTS

3.1 | Results at T2

Analyzing the pattern of results at T2 we found that the ANS acuity in comparing large numerosities (indexed by the internal Weber fraction

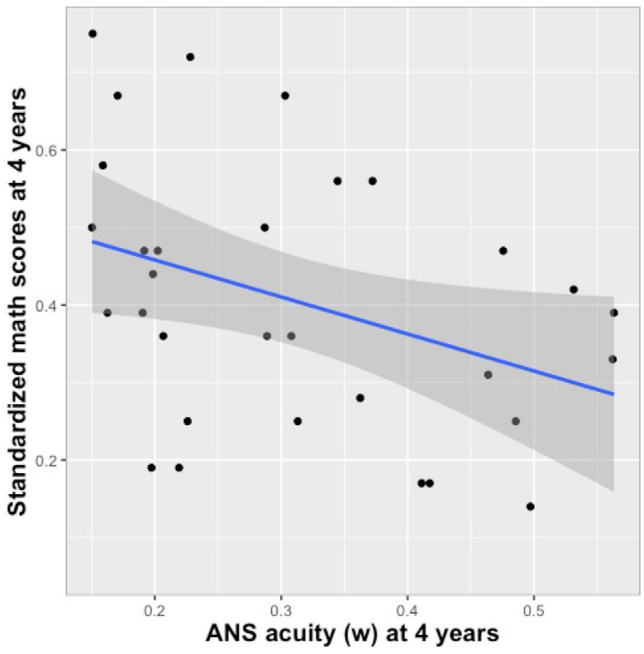

FIGURE 2 Correlation between ANS acuity, indexed by the internal Weber fraction *w*, and accuracy score in the standardised maths test Tedi-Math. Black dots represent each participant and grey bands represent the confidence thresholds given from the linear modelling statistics.

w) at 4 years of age correlates with performance in the symbolic mathematical test (Tedi-Math test) at the same age (r_s = –.42, p = 0.018; see Figure 2): children with high ANS acuity performed high in the Tedi-Math test. This link held both when controlled for IQ (r_s = –0.45, p = 0.012) and for inhibitory skills (r_s = –0.37, p = 0.04). Performance in the Tedi-Math test also correlated with performance in both the GaN and the SFON tests (r_s = 0.57, $p < 0.001$; r_s = 0.36, p = 0.022, respectively). Both links held even when controlling for IQ (r_s = 0.53, $p < 0.001$; r_s = 0.34, p = 0.035, respectively), and the GaN (but not the SFON) also when controlling for inhibitory skills (partial correlation between the GaN and formal maths, controlled for inhibitory skills: r_s = 0.51, p = 0.001).

The Tedi-Math also correlated with general intelligence (r_p = 0.33, p = 0.039), but not with inhibitory skills (r_s = 0.26, p = 0.11) nor with face recognition skills (r_p = 0.11, p = 0.51). In order to more directly identify which indexes best predicted performance in the symbolic maths processing (Tedi-Math) we performed a model comparison analysis. We considered all models resulting from all possible combinations of the predictors and ranked them on the basis of the Bayesian Information Criterion (BIC; Raftery, 1995; see Table 2 for the beta's values; see also Table S2). As predictors we included non-symbolic numerical skills (ANS acuity and SFON skills), as well as non-numerical skills, both general (inhibition, general IQ) and domain specific (face processing skills). We did not include the GaN task as a predictor as it taps on the very same knowledge of numerical symbols that are also probed in the Tedi-Math. The model that best predicted symbolic maths processing was the one that included the ANS acuity alone (BIC = –15.23). We also performed the same analyses including the two T1 indexes (number and face preference score, respectively) and we found again that the best model was the one with the ANS acuity alone (BIC = –15.22). Interestingly, however, the second best model predicting Tedi-Math's score included the infants' number preference scores at T1 (BIC = –14.1; see also Supplementary Materials).

TABLE 2 Predictors used for the model comparison analysis and the corresponding beta values.

Predictor	Beta
ANS acuity (T2)	–0.38
SFON (T2)	0.4
General Intelligence (T2)	0.33
Inhibitory skill (T2)	0.19
Face perception (T2)	0.11
Numerical Preference Score (T1)	0.36
Face Preference Score (T1)	–0.27

3.2 | Results of longitudinal analyses (T1-T2)

Compared to Starr et al. (2013), in this study we tested slightly older children, grown in a different socio-cultural environment, and used a different set of tasks to assess the various skills. Despite these differences, we closely replicated the main longitudinal results of Starr et al. (2013). First, we found that the inter-individual variability of children's sensitivity to numerosity is stable in time even when the methods for estimating it was very different: number acuity measured by an implicit change detection paradigm at 12 months significantly correlated with number acuity measured by an explicit number comparison task at 4 years of age (r_s = –0.45, p = 0.013; see Figure 3a). We also found this link persisted even after controlling for IQ (r_s = –0.45, p = 0.013). Additionally, we also show that this link was not accounted for by inhibitory skills (r_s = –0.42, p = 0.024). Moreover, as in Starr et al., we confirm that this link was specific in that the measure of non-numerical perceptual abilities (face change detection) at T1 was not a reliable predictor of number comparison at T2 (r_p = –0.27, p = 0.105, see Figure 4).

Second, and even more crucially, we replicated the significant longitudinal correlation between the ANS acuity measured at 12 months and the performance in the standardised symbolic maths test at 4 years (r_p = 0.36, p = 0.022; Figure 3b). Our data show that this correlation also held when controlling for IQ (r_p = 0.33, p = 0.04). Here we also tested inhibitory skills, as some suggested that the link between maths ability and the ability to compare numerosities could be accounted for by general-purposes inhibitory skills (e.g., Gilmore et al., 2013; Szűcs et al., 2013). Partial correlations controlling for inhibitory skills indicated that the ANS at 12 months remained a reliable predictor of later symbolic maths skills (r_s = 0.36, p = 0.03). Notably, the ANS acuity at 12 months and inhibitory skills at 4 years did not correlate (r_s = 0.07, p = 0.702).

Third, in order to test the specificity of the longitudinal correlation between non-symbolic number processing at T1 and symbolic maths

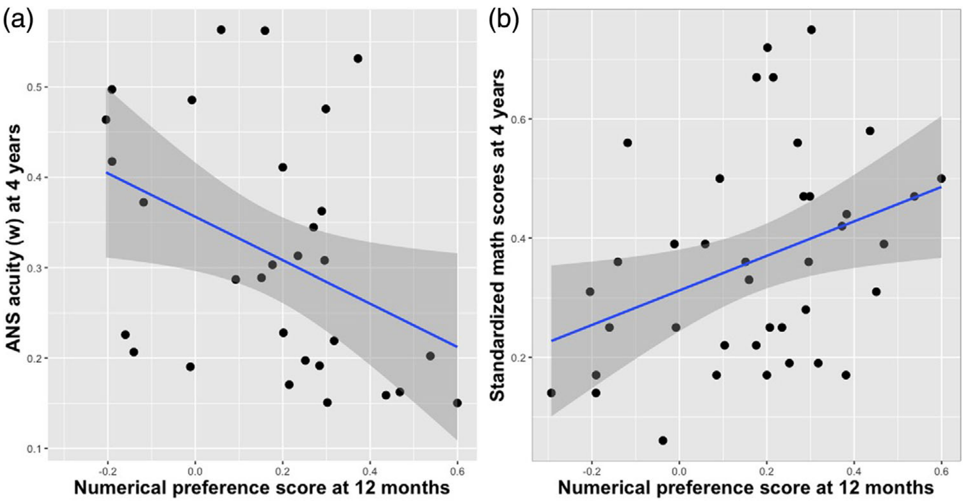

FIGURE 3 (a) Correlation between ANS acuity, indexed by the Weber fraction, and numerical preference scores in the change detection paradigm at 12 months. (b) Correlation between accuracy scores in the Tedi-Math and numerical preference scores in the change detection paradigm at 12 months. Black dots represent each participant and grey bands represent the confidence thresholds given from the linear modelling statistics.

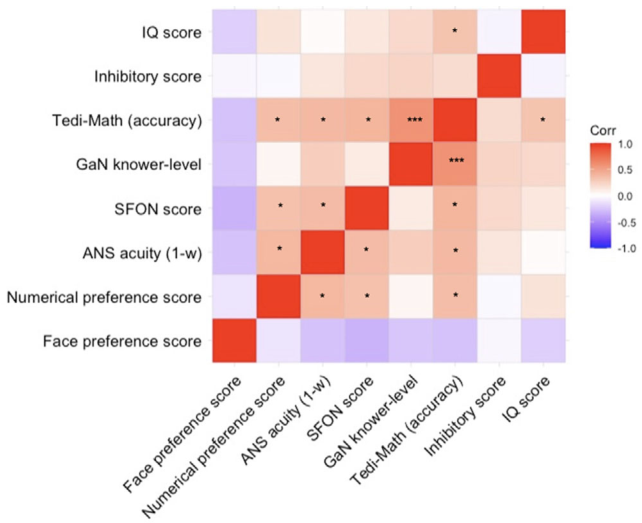

FIGURE 4 Matrix of correlations for the performances in the tasks used at T1 and T2. The gradient indicates the strength and the direction of the correlations. The stars indicate significant correlations ($* = p < 0.05$; $*** = p < 0.001$). Note that the w fraction has been inverted (i.e., subtracting the true values from 1) to show the positive correlation with the mathematical index.

skills at T2 we verified that performance in face processing, our perceptual control task at 12 months, did not predict Tedi-Math scores. As expected, we did not find an association between these two skills ($r_p = -0.27$, $p = 0.105$; $r_p = -0.19$, $p = 0.235$). In addition, we found that the link between early ANS at T1 and math at T2 persists after controlling for face processing at T1 ($r_p = 0.34$, $p = 0.04$). Moreover, we have compared the magnitude of the two correlations (i.e., ANS acuity at T1 and math at T2 vs. face perception at T1 and math at T2). The R package 'cocor' (see Diedenhofen & Musch, 2015) was used in the analysis to allow the comparison of two overlapping correlations based on dependent groups. In particular, we implemented Hittner's method (Hittner et al., 2003; using a modified index of Fisher's Z transformation) and Zou's method (Zou, 2007; using confidence intervals for assessing the difference). The results of both analyses led to the convergent conclusion that the difference between the two correlations ANS at T1-math and face perception-math is significant, and the null hypothesis about the equality of the two correlations can thus be rejected (Zou: 95% C.I. for T1 ANS acuity/math—Face perception/math: 0.15 1.01; Hittner: $z = 2.51$, $p = 0.012$).

Finally, we did not find a significant correlation between either ANS at T1 and GaN ($r_s = 0.07$, $p = 0.69$), nor ANS at T2 and GaN ($r_s = -0.22$, $p = 0.24$). One possible explanation is that there is not enough variability between subjects in our GaN task: almost half of the children tested at T2 (N = 17) were at ceiling, reaching the highest knower-level (7 or 8). This result is in line with the findings obtained by Sarneka and Carey (2008), who demonstrated that many children at 4 years of age are already able to manage the counting rules and can be considered cardinal principle-knowers. Therefore, as in our case, at 3–4 year of age children seem to possess the capacity to manage the cardinal principle and implement it to give the right response to the GaN task.

4 | DISCUSSION

Understanding what are the main cognitive skills that ground mathematical acquisitions is of great theoretical and societal interest. One main-stream hypothesis holds that a key skill is the preverbal ability to represent and compare the number of items in a set, an ability that is thought to be supported by a mechanism named the Approximate Number System (ANS). To date, however, only one study has directly addressed the link between ANS and maths measuring the ANS in preverbal infants and maths skills three years later, thus excluding the potential effects of education on the first measures of the ANS (Starr et al., 2013).

Here, we examine this link in another longitudinal study, assessing pre-numerate infants (at 12 months) and re-testing them 3 years later (at 4 years). Specifically, at T1 infants were tested with two versions of the change detection paradigm, aiming at assessing numerical acuity on one side and face perception, used as a control, on the other. At T2, the same children were tested with maths-related tasks assessing symbolic maths achievement (with a standardised battery for early maths achievement), the ability to understand the counting principles (GaN task), the ability to spontaneously attend numerosity (SFON task) and the ability to compare the number of objects in visual sets (ANS), as well as domain-general abilities (inhibitory skills and general intelligence). In order to test for the specificity of the ANS to maths correlation we also tested a non-numerical skill at T2 (face processing), which we expected that would not be predicted by early ANS measures.

Analysing the patterns of correlations between tasks at T2, we found that children with higher ANS acuity were also those who performed better in the standardised early maths achievement test, even when controlling for general intelligence or inhibitory skills. This evidence is in line with previous results and provides further support to the claim that the inborn ability to represent quantities is tightly linked to early mathematical skills, and that this link is unlikely mediated by domain-general abilities (Keller & Libertus, 2015).

This finding becomes even more convincing when we look at the longitudinal correlations, where we observed that the ANS acuity measured at 12 months is a specific and reliable predictor of maths acquisition at 4 years. This is the first time that, despite testing older participants from a different linguistic and cultural background and using different tests, the important initial findings of Starr et al. (2013) are replicated, confirming that ANS is one key system grounding early symbolic numerical acquisitions. We hypothesise that a good sense of numerosity plays a pivotal role in learning early number semantics: it fosters the process of magnitude-to-symbol mapping, leading children to attribute meaning to number words and to map them into their pre-existing representations of coarse magnitudes. This is also in line with previous reports of a deficit in the ANS in children with developmental dyscalculia (e.g., Decarli et al., 2020; Decarli et al., 2023; Piazza et al., 2010) and of a benefit in arithmetic tasks after an ANS training (e.g., Park & Brannon, 2013 but see also Szkudlarek et al., 2021 for failure to observe it). Importantly, we also tested for the specificity of the longitudinal ANS-to-maths link by including, both at T1 and at T2, control perceptual tests of perceptual non-numerical skills. We chose face processing as a control domain because from both neuroimaging and behavioural studies we know that might develop independently from number processing (e.g., Chinello et al., 2013). Accordingly, here we found a clear dissociation: neither the face change detection skills at T1 predicted symbolic maths skills at T2, nor the numerical change-detection tests at T1 predicted face perception skills at T2.

Crucially, we found that the ANS to symbolic maths correlation persisted also when controlling not only for general intelligence (as in Starr et al., 2013), but also for domain-general inhibitory skills. Recent theories have proposed that inhibitory skills can account for the link between ANS and maths (see for example Gilmore et al., 2013). These authors proposed that dot comparison tasks require a high level of inhibition capacities (to inhibit responses to non-numerical aspects of the sets, such as size or density) and that this capacity (and not numerical representations per se) can explain most of the inter-individual variance in symbolic mathematical scores. Our data do not go in this direction: we did not find a correlation between maths abilities and inhibitory skills and the correlations that we found between the ANS and symbolic maths persisted also once controlled for inhibition.

Our results suggest that, in the first steps of development, the role played by the ANS is predominant compared to that potentially played by the inhibitory system. The ability to inhibit a response might become more relevant when the requests of the maths task are higher, potentially for performing more complex arithmetical operations. Alternatively, it might be important to revise the definition of the precise kinds of inhibitory skills that may play a role in linking the ANS to maths achievement. Indeed, the inhibition literature suggests that we could distinguish between inhibitory response and interference control (Nigg, 2000; see also Malone et al., 2019). The former refers to the control of an impulsive response in favour of another, while the latter refers to the ability to intentionally concentrate on some stimuli while inhibiting the attentional focus to others. Our inhibitory task is primarily based on inhibitory response rather than on interference control and this leaves open the possibility that another type of inhibition is related to early mathematical abilities (see Traverso et al., 2019 for positive training effects of interference on maths).

Among the non-standardized tests that we used at T2 are the Give a Number and the Spontaneous Focus on Numerosity tests. The former one assesses the understanding of the cardinality principle (e.g., Sarneka & Carey, 2008), and coherently we found that it heavily correlates with standardised tests of symbolic number processing, persisting even once controlling for inhibitory skills. The latter assesses the ability of children to spontaneously attend to the feature 'number' of a complex scene. We found that also SFON was correlated with symbolic maths skills, but this link did not persist when controlling for inhibition. In this task children are asked to integrate a figure with missing parts starting from a model. Due to the nature of the task, children might be required to inhibit the tendency to draw some extra/minus parts at the animals. Therefore, the correlation between SFON and maths can be partially explained by the role played by the inhibitory ability there.

In summary, our data support the claim that ANS acuity plays a causal role in the early mathematical achievement. Indeed, we demonstrated a link before the acquisition of formal math and before the learning of language (and in turn of early counting). However, we remain cautious about this interpretation as there may be other variables mediating this relationship.

Nevertheless, this study represents an important step in the definition of the cognitive foundations of early mathematical acquisition, and may indicate the importance of early educational or remediation interventions to foster early maths acquisitions. Considering the recent debate about the role of inhibitory skills in maths, we did not provide evidence in favour of it. However, our study leaves open some questions that should be addressed in future studies that should explore the possible predictive and distinct role of different types of inhibitory functions.

AUTHOR CONTRIBUTIONS

M.P. and G.D. developed the study concept. G.D., D.Z. and M.P. contributed to the study design. G.D. and D.Z. collected the data. G.D. and D.Z. performed the data analysis and interpretation under the supervision of M.P. G.D., M.P. and D.Z. drafted the manuscript. L.S. provided critical revisions. All authors approved the final version of the manuscript for submission.

ACKNOWLEDGMENTS

We warmly thank the many families who participated in this study. We also thank Massimo Vescovi, Chiara Zanonato, Leonardo Venturoso and Francesca Anderle for helping in various stages of data collection. This research was supported by the Center for Mind/Brain Sciences (M.P. and D.Z.) and the Department of Psychology and Cognitive Science (G.D. and L.S.) of the University of Trento (Italy).

Open Access Funding provided by Universita degli Studi di Trento within the CRUI-CARE Agreement.

CONFLICT OF INTERESTS STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data available on request from the corresponding authors.

ETHICS STATEMENT

This study was reviewed and approved by the Ethics Committee of the University of Trento ('Lo sviluppo del senso del numero durante l'infanzia', 2015–026). Written informed consent to participate in this study was provided by the participants' legal guardian. The study meets the requirements of the Declaration of Helsinki.

ORCID

Gisella Decarli https://orcid.org/0000-0002-6784-4457
Donatella Zingaro https://orcid.org/0000-0002-3658-6045

REFERENCES

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. (2018). *Dyscalculia: From science to education*. Routledge.

Castaldi, E., Mirassou, A., Dehaene, S., Piazza, M., & Eger, E. (2018). Asymmetrical interference between number and item size perception provides evidence for a domain specific impairment in dyscalculia. *PLOS ONE*, *13*(12), e0209256. https://doi.org/10.1371/journal.pone.0209256

Clayton, S., & Gilmore, C. (2015). Inhibition in dot comparison tasks. *ZDM*, *47*(5), 759–770. https://doi.org/10.1007/s11858-014-0655-2

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

Chinello, A., Cattani, V., Bonfiglioli, C., Dehaene, S., & Piazza, M. (2013). Objects, numbers, fingers, space: Clustering of ventral and dorsal functions in young children and adults. *Developmental Science*, *16*(3), 377–393. https://doi.org/10.1111/desc.12028

Croydon, A., Pimperton, H., Ewing, L., Duchaine, B. C., & Pellicano, E. (2014). The Cambridge Face Memory Test for Children (CFMT-C): A new tool for measuring face recognition skills in childhood. *Neuropsychologia*, *62*, 60–67. https://doi.org/10.1016/j.neuropsychologia.2014.07.008

Decarli, G., Paris, E., Tencati, C., Nardelli, C., Vescovi, M., Surian, L., & Piazza, M. (2020). Impaired large numerosity estimation and intact subitizing in developmental dyscalculia. *PloS One*, *15*(12), e0244578. https://doi.org/10.1371/journal.pone.0244578

Decarli, G., Sella, F., Lanfranchi, S., Gerotto, G., Gerola, S., Cossu, G., & Zorzi, M. (2023). Severe developmental dyscalculia is characterized by core deficits in both symbolic and nonsymbolic number sense. *Psychological Science*, *34*(1), 8–21. https://doi.org/10.1177/09567976221097947

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

Dehaene, S. (2007). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P. Haggard (Ed.) *Attention and Performance XII: Sensorimotor Foundations of Higher Cognition*, Oxford University Press, (pp. 527–574).

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, *56*(2), 384–398. https://doi.org/10.1016/j.neuron.2007.10.004

Dehaene, S., Izard, V., & Piazza, M. (2005). Control over non-numerical parameters in numerosity experiments. Unpublished manuscript.

De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*(2), 48–55. https://doi.org/10.1016/j.tine.2013.06.001

Diedenhofen, B., & Musch, J. (2015). Cocor: A comprehensive solution for the statistical comparison of correlations. *PLOS ONE*, *10*(4), e0121945. https://doi.org/10.1371/journal.pone.0121945

Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., Simms, V., & Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PloS One*, *8*(6), e67374. https://doi.org/10.1371/journal.pone.0067374

Grégoire, J., Noël, M., & Van Nieuwenhoven, C. (2004). *Tedi-Math*. Harcourt.

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

Hannula, M. M., & Lehtinen, E. (2005). Spontaneous focusing on numerosity and mathematical skills of young children. *Learning and Instruction*, *15*(3), 237–256. https://doi.org/10.1016/j.learninstruc.2005.04.005

Hittner, J. B., May, K., & Silver, N. C. (2003). A Monte Carlo evaluation of tests for comparing dependent correlations. *The Journal of General Psychology*, *130*(2), 149–168. https://doi.org/10.1080/00221300309601282

Keller, L., & Libertus, M. (2015). Inhibitory control may not explain the link between approximation and math abilities in kindergarteners from middle class families. *Frontiers in Psychology*, *6*, 685. https://doi.org/10.3389/fpsyg.2015.00685

Korkman, M., Kirk, U., & Kemp, S. (2007). *Nepsy-II*. Pearson.

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

Libertus, M. E., & Brannon, E. M. (2010). Stable individual differences in number discrimination in infancy. *Developmental Science*, *13*(6), 900–906. https://doi.org/10.1111/j.1467-7687.2009.00948.x

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

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

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, *31*(5), 553–566. https://doi.org/10.1068/p3339

Nieder, A. (2021). The evolutionary history of brains for numbers. *Trends in Cognitive Sciences*, *25*(7), 608–621. https://doi.org/10.1016/j.tics.2021.03.012

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*, *126*(2), 220–246. https://doi.org/10.1037/0033-2909.126.2.220

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

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, *14*(12), 542–551. https://doi.org/10.1016/j.tics.2010.09.008

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., Dehaene, S., & Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*(1), 33–41. https://doi.org/10.1016/j.cognition.2010.03.012

Piazza, M., Pica, P., Izard, V., Spelke, E. S., & Dehaene, S. (2013). Education enhances the acuity of the nonverbal approximate number system. *Psychological Science*, *24*(6), 1037–1043. https://doi.org/10.1177/0956797612464057

Piazza, M., De Feo, V., Panzeri, S., & Dehaene, S. (2018). Learning to focus on number. *Cognition*, *181*, 35–45. https://doi.org/10.1016/j.cognition.2018.07.011

Price, G. R., Palmer, D., Battista, C., & Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: Reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychologica*, *140*(1), 50–57. https://doi.org/10.1016/j.actpsy.2012.02.008

Quinn, P. C., Tanaka, J. W., Lee, K., Pascalis, O., & Slater, A. M. (2013). Are faces special to infants? An investigation of configural and featural processing for the upper and lower regions of houses in 3- to 7-month-olds. *Visual Cognition*, *21*(1), 23–37. https://doi.org/10.1080/13506285.2013.764370

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163. https://doi.org/10.2307/271063

Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition*, *108*(3), 662–674. https://doi.org/10.1016/j.cognition.2008.05.007

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), e12372. https://doi.org/10.1111/desc.12372

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

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

Szűcs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex*, *49*(10), 2674–2688. https://doi.org/10.1016/j.cortex.2013.06.007

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

Traverso, L., Viterbori, P., & Usai, M. C. (2019). Effectiveness of an Executive Function Training in Italian Preschool Educational Services and Far Transfer Effects to Pre-academic Skills. *Frontiers in Psychology*, *10*, 2053. https://www.frontiersin.org/article/10.3389/fpsyg.2019.02053

Wechsler, D. (2012). Wechsler preschool and primary scale of intelligence—fourth edition. The Psychological Corporation San Antonio, TX.

Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods*, *12*(4), 399. https://doi.org/10.1037/1082-989X.12.4.399

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Decarli, G., Zingaro, D., Surian, L., & Piazza, M. (2023). Number sense at 12 months predicts 4-year-olds' maths skills. *Developmental Science*, e13386. https://doi.org/10.1111/desc.13386