Received: 3 February 2016 | Accepted: 26 September 2016
DOI: 10.1111/desc.12533

PAPER

WILEY Developmental Science

Children's intuitive sense of number develops independently of their perception of area, density, length, and time

Darko Odic

Department of Psychology, University of British Columbia, Vancouver, Canada

Correspondence
Darko Odic, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, British Columbia V6T 1Z4, Canada.
Email: darko.odic@psych.ubc.ca

Abstract

Young children can quickly and intuitively represent the number of objects in a visual scene through the Approximate Number System (ANS). The precision of the ANS – indexed as the most difficult ratio of two numbers that children can reliably discriminate – is well known to improve with development: whereas infants require relatively large ratios to discriminate number, children can discriminate finer and finer changes in number between toddlerhood and early adulthood. Which factors drive the developmental improvements in ANS precision? Here, we investigate the influence of four non-numeric dimensions – area, density, line length, and time – on ANS development, exploring the degree to which the ANS develops independently from these other dimensions, from inhibitory control, and from domain-general factors such as attention and working memory that are shared between these tasks. A sample of 185 children between the ages of 2 and 12 years completed five discrimination tasks: approximate number, area, density, length, and time. We report three main findings. First, logistic growth models applied to both accuracy and Weber fractions (w; an index of ANS precision) across age reveal distinct developmental trajectories across the five dimensions: while area and length develop by adolescence, time and density do not develop fully until early adulthood, with ANS precision developing at an intermediate rate. Second, we find that ANS precision develops independently of the other four dimensions, which in turn develop independently of the ANS. Third, we find that ANS precision also develops independently from individual differences in inhibitory control (indexed as the difference in accuracy and w between Congruent and Incongruent ANS trials). Together, these results are the first to provide evidence for domain-specific improvements in ANS precision, and place children's maturing perception of number, space, and time into a broader developmental context.

RESEARCH HIGHLIGHTS

- The Approximate Number System (ANS) provides children with intuitive but imprecise representations of number.
- Here, we test which factors drive the improvement of ANS precision with age by comparing the developmental trajectories of the ANS with those of area, density, line length, and time representations between toddlerhood and adulthood.
- This is the first study to report concurrent Weber fractions for these five dimensions across a broad age; in addition, growth modelling and partial correlation analyses revealed that ANS precision develops independently of area, density, length, and time, and from children's ability to inhibit non-numeric dimensions during the ANS task.
- These results place children's developing sense of number, space, and time in a broader and richer developmental context.

1 | INTRODUCTION

Thinking about number and quantity is at the heart of everything we do: we select the shortest line at the grocery store; choose the least dense part of the auditorium to sit in; estimate how much wine in our glass is enough. Our ability to reason about number, space, and time is foundational for other cognitive abilities, and individual differences in these representations predict musical performance (Grondin & Killeen, 2009), sports performance (Witt, Linkenauger, Bakdash, & Proffitt, 2008), and everyday activities such as reasoning about money (Marques & Dehaene, 2004). Thus, understanding the ontology and the development of number, space, and time representations is of interest in many subfields of psychology and neuroscience, including developmental, cognitive, comparative and computational psychology.

Although children's early emerging representations of space and time have long been the focus of study, research has recently shown that most human and non-human animals also have an intuitive sense of number, often termed the Approximate Number System (ANS). For example, within hours of birth, newborns expect the number of visual objects to match perceptually to the number of sounds that they hear (Izard, Sann, Spelke, & Streri, 2009). Later in development, babies that have habituated to a particular number of objects (e.g. 32 dots) can subsequently detect numerically large changes in the display (e.g. a change to 16 dots; Feigenson, Dehaene, & Spelke, 2004; Jordan & Brannon, 2006; Xu & Spelke, 2000). The ANS has similarly been found in many non-human animals, including rhesus macaques (Brannon & Terrace, 1998; Cantlon & Brannon, 2006; Nieder & Miller, 2004), rats (Meck & Church, 1983), pigeons (Brannon, Wusthoff, Gallistel, & Gibbon, 2001; Emmerton, 1998), and even guppies (Piffer, Agrillo, & Hyde, 2011). Converging evidence from cognitive, developmental, computational, comparative and neurophysiological psychology suggests that the ANS is localized to a particular brain region – the intraparietal sulcus (IPS; Emerson & Cantlon, 2015; Nieder, 2005; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; Roitman, Brannon, & Platt, 2007) – and that it is universally shared across different cultures, including those that do not have words for numbers or mathematical concepts (Frank, Everett, Fedorenko, & Gibson, 2008; Gordon, 2004; Pica, Lemer, Izard, & Dehaene, 2004).

The cost of having such an intuitive number system, however, is its imprecision. Our ability numerically to discriminate two sets of objects depends on their ratio: discriminating a large ratio, such as 30 vs. 10 dots (a ratio of 3.0) is easy even for newborns (Xu & Spelke, 2000), while discriminating a small ratio, such as 15 vs. 14 dots (a ratio of 1.07) is challenging even for most adults (Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Libertus, Odic, & Halberda, 2013). Individual differences in the ANS are typically indexed through each person's Weber fraction (w), a behavioral index of the noise in the underlying neural tuning curves that represent number (Halberda & Odic, 2014; Nieder, 2005; Piazza et al., 2004; Pica et al., 2004).

Work by Halberda and colleagues (2012) demonstrates that – despite the universality of the ANS – there are large individual differences in the precision of the system at virtually every age. ANS precision also undergoes significant developmental improvement with age. Thus, while a typical 9-month-old infant has a w of about 0.5 (i.e. can reliably discriminate ratios of about 1.5), a typical 19-year-old college student has w values of about 0.15 (Cordes & Brannon, 2008; Halberda et al., 2012; Libertus et al., 2012; Xu & Spelke, 2000). ANS development begins from very early in infancy and does not peak until late adolescence or early adulthood, subsequently declining in old age (Halberda et al., 2012).

Why does the ANS show these large developmental changes? In other words, what are the main factors that drive the development of ANS precision? The existing literature broadly suggests three possibilities.

The first possibility is that the ANS, as a dedicated system for representing number, may develop as a result of domain-specific effects, including the maturation of specific brain circuits that implement it, or alternatively as a result of children's increased familiarity with and expertise at using the system. In other words, the ANS may develop for its own reasons and along its own developmental trajectory. As an analogy to this explanation, consider the maturation of children's visual acuity: while all typically developing children are born with the ability to see, the development of very specialized muscles and circuits, including the orbital muscles, the fovea, and dedicated circuits between the eyes and the visual cortex, will lead to a developmental peak in early toddlerhood (Mayer & Dobson, 1982; Yuodelis & Hendrickson, 1986). Because these muscles and circuits are dedicated to vision, one can state that – despite the fact that many other brain and body regions are developing in parallel with vision – the development of visual acuity is causally independent from the development of, for example, audition or motor control. Domain-specific development of the ANS is supported by the finding that its precision is especially improved by education in mathematics (Piazza, Pica, Izard, Spelke, & Dehaene, 2013; see also Lindskog, Winman, & Juslin, 2014), and that an infant's ANS precision at 6 months predicts their precision in preschool (Starr, Libertus, & Brannon, 2013).

The second possibility is that the development of ANS precision has nothing to do with the ANS itself, and that, instead, it may be capturing children's ability to process and discriminate non-numeric dimensions, including object density, surface area, contour extent, etc. (Cantrell & Smith, 2013; Dakin, Tibber, Greenwood, Kingdom, & Morgan, 2011; Defever, Reynvoet, & Gebuis, 2013; Gebuis & Reynvoet, 2012; Szucs, Nobes, Devine, Gabriel, & Gebuis, 2013). This theory is motivated by the observation that, on a typical task measuring ANS precision, participants are asked to identify the numerically larger set of dots (e.g. the number of yellow vs. blue dots in Figure 1). However, such a display necessarily contains information about a host of non-numeric dimensions, including the size and density of the dots. Children may perform above chance, then, not because they have an intuitive number system, but instead by discriminating the myriad of correlated non-numeric dimensions. Hence, development in number discrimination tasks may, in fact, simply reflect an improvement in these non-numeric dimensions.

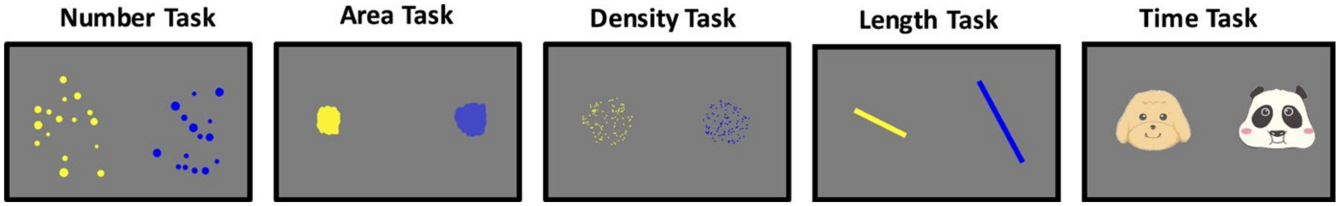

FIGURE 1 Example stimuli from the five discrimination tasks. In the Time task, the two characters are animated and hold their breath for a certain amount of time

The idea that ANS precision may be contaminated by non-numeric dimensions is supported by arguments for a generalized magnitude system, the hypothesized unified sense of magnitude that underpins all reasoning about number, space, and time (Bueti & Walsh, 2009; Lourenco & Longo, 2010, 2011; Walsh, 2003). The existence of the generalized magnitude system is supported by several key findings. First, many non-numeric dimensions, especially area and density, influence number discrimination performance (e.g. participants frequently select the denser display as the more numerous; Dakin et al., 2011; Durgin, 1995; Gebuis & Reynvoet, 2012; Szucs et al., 2013). Second, there are a number of cross-over and association effects between number, space, and time, from infancy onwards (de Hevia, Izard, Coubart, Spelke, & Streri, 2014; Lourenco & Longo, 2010), including effects showing that the perception of large numbers biases our attention towards the right side of space (Fischer, Castel, Dodd, & Pratt, 2003; Wood, Willmes, Nuerk, & Fischer, 2008) and towards longer durations (Walsh, 2003). Third, research has repeatedly shown that number, space, and time representations are all instantiated in the IPS (Bueti & Walsh, 2009; Pinel, Piazza, Le Bihan, & Dehaene, 2004), that administration of electrical stimulation to this region affects number, space, and time perception (Cappelletti et al., 2013), and that single-unit recordings reveal an overlap in the neurons that code for number and line length (Tudusciuc & Nieder, 2007). Under some formulations of this theory, the generalized magnitude system differentiates into several sub-systems as children learn about and interact with the world (Cantrell & Smith, 2013; Lourenco & Longo, 2010, 2011), while others argue that the generalized magnitude system persists deep into adulthood, providing a common currency by which we represent and reason about magnitude (Bueti & Walsh, 2009; Fabbri, Cancellieri, & Natale, 2012; Lu, Mo, & Hodges, 2011; Xuan, Zhang, He, & Chen, 2007). Hence, ANS development could instead be the development of a single, unified, generalized magnitude system, and thus the development of the ANS should be tightly coupled to the development of space and time representations.

The third possibility is that the primary factor driving ANS development is the maturation of various domain-general abilities, including attention, working-memory, inhibitory control, or more general parietal lobe maturation. For example, Xenidou-Dervou, De Smedt, van der Schoot, and van Lieshout (2013) have shown that ANS discrimination tasks put a load on working memory, while Ratcliff, Love, Thompson, and Opfer (2012) have shown that various decision-making factors, including response biases, play into children's and adult's performance. One especially likely factor that may drive ANS development is children's improving inhibitory control: Gilmore and colleagues (2013) have shown that ANS precision at least partially captures how well children perform on trials in which they have to inhibit various non-numeric dimensions, such as size, in order to answer the number question correctly (i.e. Incongruent trials, where the more numerous dots are smaller than the less numerous dots), as opposed to trials where various non-numeric dimensions correlate with number (i.e. Congruent trials, where the more numerous dots are also bigger). ANS development, then, may simply be capturing the development of children's ability to precisely attend to number as opposed to other dimensions (although see DeWind, Adams, Platt, & Brannon, 2015; Keller & Libertus, 2015; Starr & Brannon, 2015b).

Overall, theories accounting for the development of ANS precision can be split into three (non-mutually exclusive) possibilities: (1) domain-specific ANS development, (2) the development of several other non-numeric dimensions, including area, density, line length and time, and (3) the maturation of domain-general abilities such as inhibitory control.

The existing data on ANS development and its relationship to non-numeric dimensions unfortunately cannot adjudicate among these three possibilities. First, few studies have examined the development of the ANS in conjunction with other dimensions. Odic, Libertus, Feigenson, and Halberda (2013; see also Odic, Pietroski, Hunter, Lidz, & Halberda, 2013) showed that ANS precision develops independently from surface-area precision, but only studied a narrow age-range of 3- to 6-year-old children. Work by Starr and Brannon (2015a) has investigated the relationship between number, brightness and line length, but focused primarily on correlations between the precision of these dimensions, showing differences in the correlations between preschoolers and adults. Correlations between dimensions, however, can stem from several factors, including shared domain-general skills such as working memory (see Odic et al., 2016 and Van Opstal & Verguts, 2013), making it difficult to make any strong claims about the co-development of these dimensions. Other studies have similarly focused on the development of number and one or two other dimensions, but usually over a narrow age range that does not capture the entire developmental trajectory for these dimensions (e.g. Abreu-Mendoza & Arias-Trejo, 2015; Brannon, Lutz, & Cordes, 2006; Dormal & Pesenti, 2012; Droit-Volet & Wearden, 2001). Here, we report data on a broad age-range of participants, namely 2- to 12-year-old children and college-aged adults, and across five discrimination tasks: number, area, density, line length and time. Unlike previous work, we can quantify the developmental trajectory for each dimension, and examine whether number develops independently from any or all of these dimensions.

The second major challenge in understanding ANS development has been how to control for all possible domain-general factors, including attention, working memory, inhibitory control, etc., that may be responsible for the relationship between the ANS and non-numeric dimensions, as well as for any developmental effects. Here, we take a novel approach to this problem: rather than attempting to measure each domain-general ability individually, we instead rely on the fact that, whatever the range of possible domain-general factors that affect ANS development may be, they must be shared between the ANS and the four non-numeric dimensions. Droit-Volet, Clément, and Fayol, (2008), for example, have demonstrated that time discrimination tasks put a large load on children's working memory; hence, by controlling for time discrimination performance when examining ANS development, we can also largely control for working memory differences in ANS precision. Because attention, decision-making factors, and even general parietal lobe development should all be shared across number, area, density, length, and time, controlling for these factors should also largely control for the various domain-general abilities that could be driving ANS development (see Odic et al., 2016 for evidence that this approach is appropriate for number and time representations when concurrently measuring working memory).

The goals of our experiment are, thus, twofold: besides being the first to report developmental data on a range of quantity discrimination tasks across a broad age range, we also expect to find different patterns of results depending on which of the three theories reviewed above is the best explanation for ANS development. If the development of the ANS is accounted for by the development of a generalized magnitude system, we should find that all age-related variability in ANS precision is accounted for by the age-related differences in area, density, length and time precision. In addition, because these dimensions have been shown to be represented in the IPS, controlling for these dimensions should also largely control for more domain-general development of the parietal lobe and the IPS itself. Furthermore, by examining the development of children's performance on Congruent versus Incongruent ANS trials, we can assess the role of inhibitory control development on the ANS (Gilmore et al., 2013). Importantly, if we found that the ANS develops along distinct developmental trajectories and independently from area, length, time, density and inhibitory control, we would have evidence for an important role of domain-specific ANS development.

2 | EXPERIMENT

2.1 | Participants

A total of 185 children between the ages of 2 and 12 participated in the study (93 boys and 92 girls; see Table 1). An additional 22 children were excluded because they did not complete more than a third of the tasks, mostly as a result of inattentiveness. All children were individually tested at the local science museum – Vancouver's Telus ScienceWorld LivingLab – in a dedicated, sound-attenuated

TABLE 1 Performance by each age group across the five discrimination tasks

Age group	N	Mean age (SE)	Number accuracy (n; SE)	Area accuracy (n; SE)	Density accuracy (n; SE)	Length accuracy (n; SE)	Time accuracy (n; SE)
3-year-olds	24	3.33 (0.08)	54.95 (24; 2.18)	62.43 (11; 4.46)	48.66 (7; 3.97)	61.14 (16; 3.62)	52.51 (13; 2.99)
5-year-olds	34	4.91 (0.11)	58.32 (34; 2.53)	70.58 (17; 3.66)	51.65 (16; 4.28)	77.56 (18; 3.60)	53.56 (17; 2.99)
7-year-olds	49	7.05 (0.08)	70.95 (49; 1.83)	79.84 (28; 1.55)	53.61 (26; 2.25)	82.05 (22; 2.29)	63.07 (20; 2.73)
9-year-olds	47	8.95 (0.08)	77.70 (47; 1.18)	86.40 (26; 1.38)	64.38 (26; 2.03)	89.47 (22; 1.83)	70.63 (21; 3.15)
11-year-olds	31	11.32 (0.21)	81.25 (31; 1.17)	87.13 (17; 1.54)	68.12 (15; 1.59)	90.82 (16; 1.38)	79.24 (14; 1.84)
Adults	15	20.06 (0.44)	81.73 (15; 1.27)	91.88 (15; 1.39)	72.11 (15; 1.52)	94.38 (15; 1.02)	79.91 (15; 2.53)

room. Parents waited outside the room while the child completed the task. All children were given a sticker as a prize for participating. An additional 15 undergraduates participated for course credit at the University of British Columbia.

2.2 | Methods and procedures

All stimuli were presented on a 11.3″ Macbook Air using custom-made Psychtoolbox-3 scripts (Brainard, 1997). These scripts are freely available for download and for future research use (http://odic.psych.ubc.ca/scripts/). Children were seated in front of the computer with the experimenter seated next to them. In order to reduce the potential effects of motor development on our results, the experimenter always pushed the buttons and the child was asked to respond to each trial verbally or by pointing.

We tested discrimination performance on five dimensions: approximate number, surface area, density, line length, and time, each described in detail below and illustrated in Figure 1. Because pilot testing showed that preschoolers could not complete all five tasks in a single sitting, we randomly split all children into one of two conditions: children either completed Number/Area/Density or Number/Length/Time tasks1 . Each of these conditions began with three trials from each dimension (e.g. three trials of Number, followed by three trials of Area, followed by three trials of Density) that allowed the experimenter to explain the task to the child. Subsequently, the trials were randomized and trials of the three dimensions were fully intermixed; this allowed us both to control for any task-order effects and to keep the child's attention for longer. Adult participants completed all five tasks in an intermixed order; unlike children, however, adults were allowed to push their own buttons rather than respond verbally or by pointing.

Participants received auditory feedback from the computer throughout the entire experiment. In general, children took about 5–8 minutes to complete the experiment, and adult participants took about 8–12 minutes to complete the experiment.

2.2.1 | Number task

Participants were shown displays of spatially separated blue and yellow dots within two rectangular frames, as shown in Figure 1, and asked to identify the side with more dots. Each set was associated with a cartoon character (e.g. Spongebob or a Smurf). The dots stayed on the screen for 500 milliseconds. The ratio of the dots was varied to control for difficulty, and could take the following values: 2.0 (20 vs. 10 dots), 1.5, 1.2 or 1.06. Each ratio was presented eight times, yielding a total of 32 Number trials. In an effort to control for children's use of non-numeric dimensions, we controlled for the average and the cumulative area of the dots. In addition, by partialling out the individual differences in the four non-numeric discrimination tasks, we can also control for any contributions of area, density, length, and time perception to the Number task. On half of the trials, the total surface area was congruent with the number of dots (i.e. the set with more dots had a higher cumulative surface area; Congruent trials), and on the other half of the trials, the total surface area was incongruent with the number of dots (i.e. the set with more dots had a lower cumulative surface area; Incongruent trials). As discussed above, previous work by Gilmore and colleagues (2013) has suggested that the difference between these two types of trials may be indicative of children's inhibitory control.

2.2.2 | Area task

Participants were shown displays of two amorphous blobs – one blue and one yellow (Figure 1) – and were asked to identify the larger blob. Each blob was presented in a rectangular frame and was associated with a cartoon character. The blobs stayed on the screen for 500 milliseconds. The ratio of the blob sizes was varied to control for difficulty, and could take the following values: 2.0 (212 vs. 106 pixel^2), 1.5, 1.2 or 1.06. Each ratio was presented eight times, yielding a total of 32 area trials. Note that – in contrast to the previous work by Odic, Libertus, et al. (2013) – the two blobs were presented spatially separated in order to better match the task demands of the number task.

2.2.3 | Density task

Participants were shown two clouds of blue and yellow dots that varied in density (Figure 1). These trials were generated in one of two ways: on half the trials, we kept the number of dots constant at 100 for both sides, but varied the total circular convex hull within which the dots were drawn. On the other half of the trials, we kept the convex hull constant in a circle with a radius of 70 pixels, but varied the total number of dots. Density was defined as the number of dots per pixel of convex hull area. Thus, for example, a ratio of 2.0 could be instantiated in one of two ways: either by doubling the area of the convex hull from 15 393 pixels to 30 787 pixels (i.e. changing the radius from 70 to 99 pixels), or by halving the number of dots from 100 to 50; in either case, the number of dots/pixel is 0.003. As discussed in the Results section, participants performed identically on these two trial types. We found that children did not understand the word 'denser'. Thus, in order to help children understand the task, we gave them a background story about the dots being lemmings that needed to huddle together for warmth; children were asked to identify whether the blue or the yellow lemmings were warmer. The two clouds of dots were presented within rectangular frames and stayed on the screen for 500 milliseconds. The ratio of the densities was varied to control for difficulty, and could take the following values: 2.0 (0.006 vs. 0.003 dots/pixel), 1.5, 1.2 or 1.06. Each ratio was presented eight times, yielding a total of 32 density trials.

2.2.4 | Length task

Participants were shown a blue line and a yellow line drawn on the screen (Figure 1), and were asked to identify which line was longer. Each line was drawn within a rectangular frame associated with a cartoon character. The lines stayed on the screen for 500 milliseconds. The ratio of the line lengths was varied to control for difficulty, and could take the following values: 2.0 (100 vs. 50 pixel), 1.5, 1.2 or 1.06. In order to make sure that children did not simply compare the tops of lines, each line was randomly oriented along its center axis; the

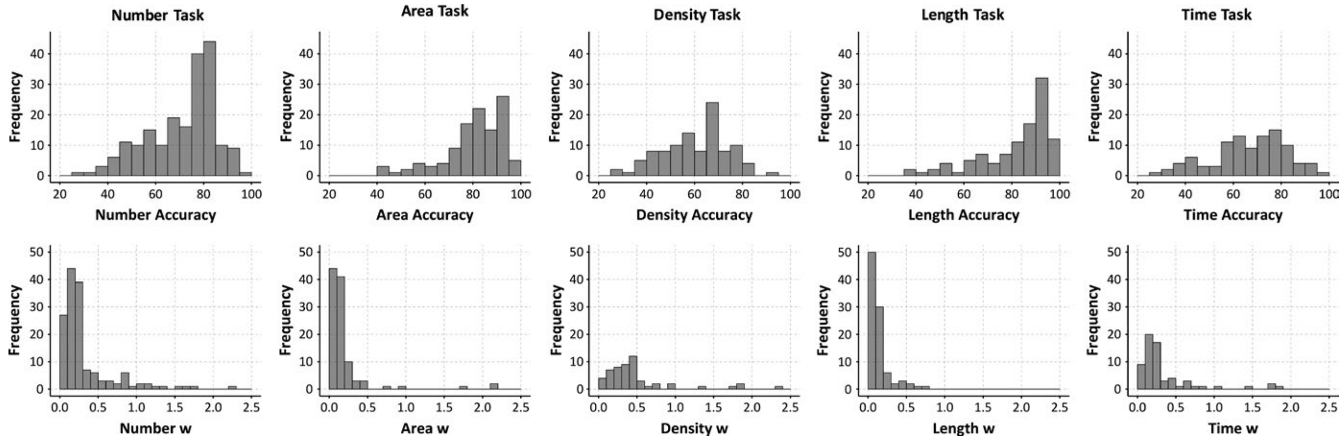

FIGURE 2 Histograms of average percentage correct and best-fit Weber fractions for each of the five tasks

difference in orientation was always at least 15 degrees. Each ratio was presented eight times, yielding a total of 32 length trials.

2.2.5 | Time task

Participants were introduced to two characters – a dog and a panda – who had a breath-holding competition. In counterbalanced order, each animal would animate to hold their breath for a particular duration lasting between 500 and 2000 milliseconds (see Figure 1). After each animal had held their breath, the children were asked to identify which animal held their breath for longer. The animals were presented within rectangular frames. The ratio of the durations was varied to control for difficulty, and could take the following values: 2.0 (1200 vs. 600 milliseconds), 1.5, 1.2 or 1.06. Each ratio was presented eight times, yielding a total of 32 time trials.

3 | RESULTS

We report our analyses in several steps. First, we test for standard ratio effects within each dimension and test whether any of the dimensions were more accurate than the others. Second, we model each participant's data to a standard psychophysical model used in the literature to estimate Weber fractions (w) and lapse/guessing rates for each dimension. Third, we model the changes in the accuracy and w in each dimension with age, including estimating the age of maturity for each dimension. Finally, and most importantly, we test whether the developmental changes in number can be accounted for by the development of area, density, length, and/or time, or by inhibitory control (the difference between Congruent and Incongruent trials). The correlations between dimensions while controlling for any age effects are reported in the Supplementary Material. All of the reported ANOVAs are corrected for sphericity.

3.1 | Ratio effects

The histograms and average accuracy for each dimension are presented in Table 1 and Figure 2 (top). Consistent with Weber's law, each dimension showed a clear ratio effect. A 3 (Task: Number, Area, Density) × 4 (Ratio: 2.0, 1.5, 1.2, 1.06) repeated-measures ANOVA with accuracy as the dependent variable (DV) showed a main effect of Ratio [$F(3, 300) = 115.92$; $p < .001$] and Task [$F(2, 200) = 168.34$; $p < .001$], and a significant Task × Ratio interaction [$F(6, 600) = 18.24$; $p < .001$]. Contrasts revealed that this main effect of Task was carried by significantly better performance in Number ($M = 72.54$; $SE = 1.21$) compared with Density [$M = 60.63$; $SE = 1.31$; $t(99) = 9.12$; $p < .001$], and by significantly better performance in Area ($M = 83.23$; $SE = 1.04$) compared with Density [$t(101) = 18.37$; $p < .001$] and compared with Number [$t(111) = -9.78$; $p < .001$].

An analogous ANOVA with Number, Time, and Length likewise showed a main effect of Ratio [$F(2, 182) = 73.82$; $p < .001$] and Task [$F(3, 273) = 92.34$; $p < .001$], and no Task × Ratio interaction [$F(6, 546) = 1.78$; $p = .10$]. Contrasts revealed that the main effect of Task was driven by significantly better performance in Number ($M = 73.76$; $SE = 1.47$) compared with Time [$M = 67.23$; $SE = 1.59$; $t(95) = 4.27$; $p < .05$], and by significantly better performance in Length ($M = 83.1$; $SE = 1.44$) compared with Time [$t(97) = 10.97$; $p < .001$] and compared with Number [$t(106) = -9.12$; $p < .001$].

An additional independent samples t-test showed no significant difference between Time and Density accuracy [$t(88) = 0.39$; $p = .69$], nor between Length and Area accuracy [$t(92) = 0.12$; $p = .90$]. We also found no effect of Condition on the Number task: children in the Number/Area/Density condition performed equivalently well ($M = 70.7\%$, $n = 84$, $SD = 15.65\%$) to the children in the Number/Length/Time condition [$M = 69.4\%$, $n = 99$, $SD = 13.44\%$; $t(181) = .62$, $p = .54$]. Finally, we did not find any difference between the area- vs. number-doubling Density Task trials [$t(89) = 0.37$; $p = .71$].

Taken together, these results suggest that all five dimensions showed ratio-dependent performance, consistent with Weber's law, and that Area and Length accuracy was superior to Number accuracy, which in turn was significantly better than Time and Density accuracy. These results broadly replicate previous work (Abreu-Mendoza & Arias-Trejo, 2015; Droit-Volet et al., 2008; Odic et al., 2016; Odic, Libertus, et al., 2013; Starr & Brannon, 2015a, 2015b) and further add information about children's perception of density.

3.2 | Weber fractions

Weber fractions – an index of the underlying precision of quantity representations – were modelled using a two-parameter psychophysical model previously used by Halberda and Feigenson (2008) and Pica and colleagues (2004). This model assumes that the underlying representations of number, time, density, etc. are normally distributed tuning curves with the single parameter *w* indexing their standard deviation (i.e. precision; for a review, see Halberda & Odic, 2014). In addition to this standard assumption, the model allows for a second parameter – the lapse rate – that can account for a constant percentage of trials that participants may have been guessing (e.g. a lapse rate of 0.10 indicates that participants were randomly guessing on 5% of trials, independent of ratio). More formally, Weber fractions and lapse rates were estimated using the equation

$$\text{Accuracy} = (1 - \text{lapse}) * \Phi\left(\frac{\text{Ratio} - 1}{w * \sqrt{1 + \text{Ratio}^2}}\right) + \frac{\text{lapse}}{2}$$

where Φ is the Gaussian cumulative distribution function. This model was fitted to each participant's data for each task using R's *mle2* function under the assumption of normally distributed errors, which converges on the best-fit parameters by minimizing the negative log-likelihood value.

The histogram and average *w* values and lapse rates for each dimension are presented in Table 2 and Figure 2. Consistent with previous work, we found that the model could not successfully fit every participant's data, most often because some participants were guessing randomly on all of the trials and thus had an accuracy of about 50%. As a result, these participants had either non-convergent models or unreasonable *w* estimates (i.e. values of more than 3.0) and extremely high lapse rates (i.e. 1.00, indicating pure guessing across all ratios). These children were excluded from any subsequent *w* analyses (see Table 2). In the remaining analyses, we always report data using both accuracy and *w* in order to maximize our sample and demonstrate that no reported finding is due to the excluded *w* data.

The average *w* values found in our five tasks are consistent with those previously measured in the literature for number (Halberda & Feigenson, 2008; Piazza et al., 2010), area (Odic, Libertus, et al., 2013; Odic, Pietroski, et al., 2013), time (Droit-Volet et al., 2008; Odic et al., 2016), and length (Droit-Volet et al., 2008; Starr & Brannon, 2015a). The results also provide the first estimates of density *w* in children. In replication of the above results with accuracy, we found that the Number *w* ($M = 0.32$; $SE = 0.001$) was significantly worse than the Area *w* [$M = 0.19$; $SE = 0.001$; $t(82) = 4.33$; $p < .001$] and Length *w* [$M = 0.13$; $SE = 0.001$; $t(78) = 4.57$; $p < .001$], but significantly better than the Density *w* [$M = 0.51$; $SE = 0.002$; $t(44) = 4.98$; $p < .001$], and Time *w* [$M = 0.35$; $SE = 0.001$; $t(55) = 2.47$; $p < .05$].

3.3 | Effects of age

We examined the effects of age in two ways. First, we examined pairwise correlations between each dimension and age as a continuous

TABLE 2 Best-fit Weber fractions (w) and lapse rates across the six age groups for the five discrimination tasks

	Number			Area			Density			Length			Time		
Age group	# Fit	*w* (*SE*)	Lapse rate (*SE*)	# Fit	*w* (*SE*)	Lapse rate (*SE*)	# Fit	*w* (*SE*)	Lapse rate (*SE*)	# Fit	*w* (*SE*)	Lapse rate (*SE*)	# Fit	*w* (*SE*)	Lapse rate (*SE*)
3-year-olds	11	0.98 (0.17)	.02 (0.02)	9	0.86 (0.30)	.06 (0.03)	2	0.72 (0.27)	.00 (0.00)	7	0.34 (0.10)	.13 (0.06)	2	1.27 (0.45)	.01 (0.00)
5-year-olds	17	0.59 (0.14)	.12 (0.04)	13	0.27 (0.07)	.05 (0.03)	3	0.40 (0.21)	.05 (0.05)	14	0.19 (0.05)	.03 (0.02)	4	0.91 (0.34)	.01 (0.01)
7-year-olds	35	0.31 (0.04)	.06 (0.02)	26	0.16 (0.01)	.07 (0.02)	7	0.83 (0.30)	.06 (0.06)	21	0.16 (0.03)	.03 (0.02)	12	0.52 (0.16)	.12 (0.04)
9-year-olds	44	0.21 (0.02)	.06 (0.01)	26	0.10 (0.01)	.02 (0.02)	16	0.63 (0.16)	.07 (0.03)	22	0.10 (0.01)	.007 (0.01)	19	0.25 (0.04)	.04 (0.02)
11-year-olds	28	0.16 (0.02)	.06 (0.02)	16	0.11 (0.01)	.004 (0.01)	13	0.35 (0.02)	.04 (0.02)	16	0.08 (0.01)	.02 (0.01)	13	0.16 (0.02)	.13 (0.04)
Adults	15	0.17 (0.02)	.001 (0.01)	15	0.08 (0.01)	.005 (0.01)	15	0.30 (0.03)	.12 (0.06)	15	0.05 (0.01)	.02 (0.01)	15	0.22 (0.04)	.01 (0.01)

variable, excluding adults (whose higher age values may disproportionally affect the correlations). Because w values are non-normally distributed (Figure 2; see also Inglis & Gilmore, 2014), we instead used Spearman rank correlations. We found very strong correlations between each dimension and age: Number accuracy ($r = .67$; $n = 185$; $p < .001$) and w (Spearman's rho = −.64; $n = 160$; $p < .001$), Area accuracy ($r = .63$; $n = 99$; $p < .001$) and w (Spearman's rho = −.67; $n = 97$; $p < .001$), Density accuracy ($r = .53$; $n = 90$; $p < .001$) and w (Spearman's rho = −.35; $n = 56$; $p < .001$), Length accuracy ($r = .62$; $n = 96$; $p < .001$) and w (Spearman's rho = −.64; $n = 86$; $p < .001$), and Time accuracy ($r = .64$; $n = 85$; $p < .001$) and w (Spearman's rho = −.64; $n = 63$; $p < .001$)2 . We found no significant correlations between age and lapse rates with the exception of the Area task [$r(109) = -.23$; $p < .05$]. The most likely explanation of this result is that our participants guessed in an all-or-none fashion independent of age, and participants who predominantly guessed were excluded owing to poor fits to the w model. As a result, lapse rates were excluded from future analyses.

Next, we grouped each participant into one of six age groups (Table 1 and Figure 3). A 6 (Age Group: 3, 5, 7, 9, 11, and Adults) × 3 (Task: Number, Area, Density) mixed-measures ANOVA over accuracy replicated the above main effect of Task [$F(2, 192) = 85.45$; $p < .001$], but also showed a main effect of Age Group [$F(5, 96) = 15.13$; $p < .001$] and an Age Group × Task interaction [$F(10, 192) = 2.00$; $p < .05$]. As can be seen in Table 1, the difference between Number, Area and Density accuracy increases and peaks at about age 7, then decreases and stabilizes at about age 11. An analogous ANOVA with Number, Length and Time showed a main effect of Task [$F(2, 184) = 47.76$; $p < .001$], Age Group [$F(5, 96) = 3.23$; $p < .01$] and an Age Group × Task interaction [$F(10, 184) = 2.94$; $p < .01$]. As with Area and Density, the difference in accuracy between Number, Length and Time peaks at about age 7, then decreases and stabilizes. These results held if the adults were excluded from the ANOVA. Together, these results show that each dimension improved with age, but that some dimensions improved faster than others.

A pair of mixed-level ANOVAs over w values revealed the same pattern of results (Figure 3). A 6 (Age Group: 3, 5, 7, 9, 11, and Adults) × 3 (Task: Number, Area, Density) mixed-measures ANOVA over w values showed a main effect of Task [$F(2, 92) = 11.72$; $p < .001$], Age Group [$F(5, 46) = 2.91$; $p < .05$], and an Age Group × Task interaction [$F(10, 92) = 2.24$; $p < .05$]. An analogous ANOVA over Number, Length and Time showed a main effect of Task [$F(2, 112) = 5.69$; $p < .001$], Age Group [$F(1, 56) = 7.88$; $p < .001$], and an Age Group × Task interaction [$F(10, 112) = 5.69$; $p < .001$]. Thus, Weber fractions also improve with age, but, once again, we find that some dimensions show a faster improvement than others.

3.4 | Logistic growth modelling

The analyses thus far demonstrate that developmental trajectories for the five dimensions are not identical, but they do not reveal what these trajectories actually are. In order to determine the trajectory for each dimension and estimate the approximate age of maturity,

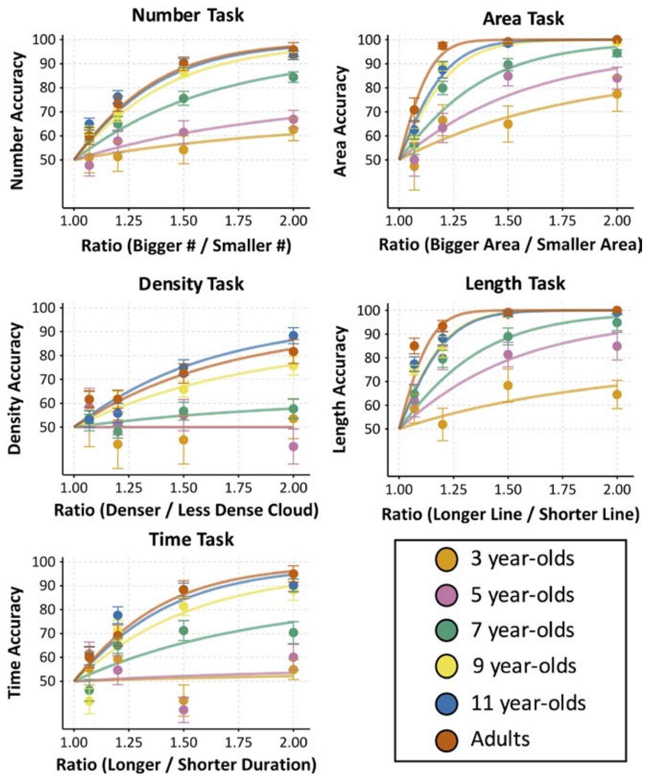

FIGURE 3 Average percentage correct across ratios and the best-fit model for each of the five tasks and for each of the six age groups. All five dimensions show improvement with age and ratio-dependent performance consistent with Weber's law. Bars indicate ±1 SD

we fitted the developmental data from each dimension to a series of logistic growth models. Logistic growth models assume that development begins at some age of onset and continues at a constant rate until an asymptote is reached and developmental growth peaks. We found that logistic growth models were superior to any other model we fitted to the accuracy or w data, including piecewise linear, log and power models. A major advantage of logistic growth models is that they estimate the growth rate – the speed at which development reaches maturity – independently from the peak itself. Thus, for example, if two dimensions have different peak accuracy levels, a non-logistic model (e.g. a power model) would mistake the lower asymptote in one dimension for evidence for continuing developmental growth.

We used the standard three-parameter logistic growth model typically used in the developmental literature (Marceau, Ram, Houts, Grimm, & Susman, 2011; Ram & Grim, 2015). This model assumes that average accuracy or w at each age is based on three parameters: a (peak accuracy or w at which development ends), $1/-b$ (the age of developmental onset), and c (the growth rate; lower values mean faster growth):

$$\text{Accuracy} = \frac{a}{1+\exp(-(b+\text{Age}*c))}$$

Note that because we did not measure accuracy or w in infancy, our ages of onset are probably slightly inflated.

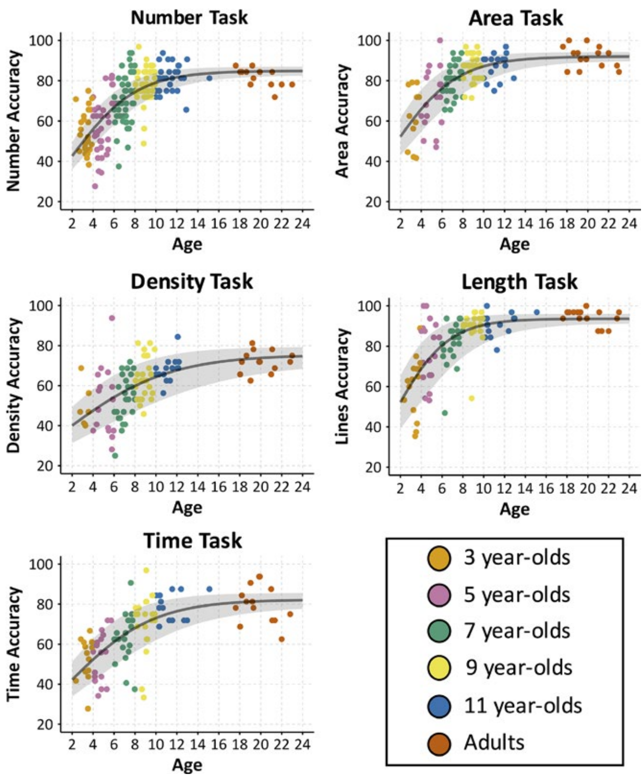

FIGURE 4 For each task, the average percentage correct for each participant and the best-fit logistic growth model. The gray shading indicates the 95% confidence interval around the model

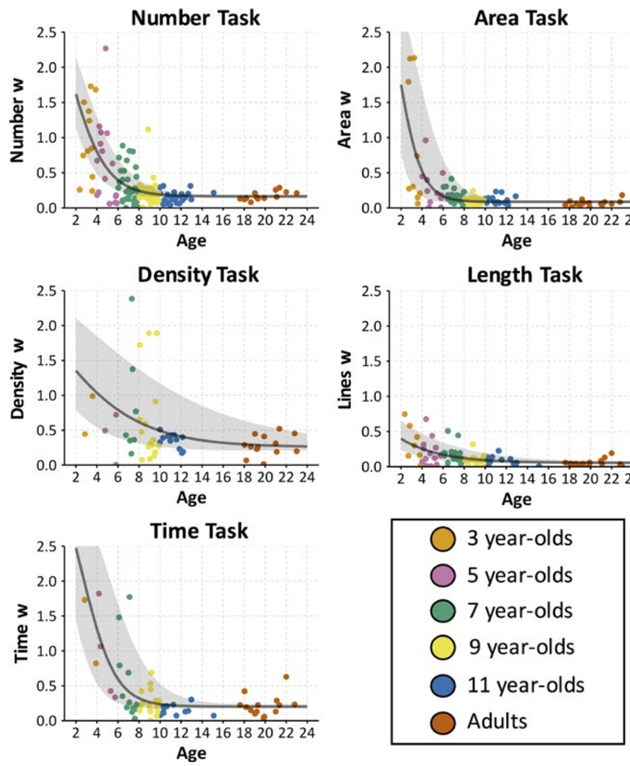

FIGURE 5 For each task, the Weber Fraction (w) for each participant fit by that model, along with the best-fit logistic growth model. The gray shading indicates the 95% confidence interval around the model

We fitted a separate logistic growth model over accuracy and *w* for each dimension. The best-fitting parameters were determined using R's *mle2* function, which converges on parameters that minimize the negative log-likelihood. In order to account for the heteroscedasticity inherent in the *w* and age data (i.e. accuracy and *w* values are more variable in younger children), the normally distributed regression error was allowed to scale with 1/Age (see also Odic, Im, Eisinger, Ly, & Halberda, in press). The best-fitting parameters are shown in Figures 4 and 5 and in Table 3. We found that these models were an excellent fit to all the data.

The estimated parameters appear reasonable and replicate the patterns reported above. For example, the estimated peak values are consistent with the ANOVAs: the peak accuracy and *w* for Density and Time are worse than the peaks for area and line length, with number falling somewhere between. Thus, the growth models suggest that the difference in accuracy and *w* among these five dimensions persists even once developmental growth is complete.

Importantly, even when these different peaks are accounted for, we see two distinct patterns in the developmental trajectories: whereas density and time show slow onsets and growth rates – not reaching peak performance until young adulthood – number, area and length develop significantly sooner, in late childhood and early adolescence. Statistically, we found that the age of maturity for the accuracy models was significantly later for density compared with line-length ($Z = 2.16$; $p = .01$), area ($Z = 1.82$; $p = .03$) and number ($Z = 1.65$; $p = .049$), and that age of maturity for time was marginally later than for line length ($Z = 1.56$; $p = .06$).

These results were even stronger for the logistic growth models over *w*: we found that the age of maturity for number was significantly earlier than for density ($Z = 11.51$; $p < .001$), earlier than for time ($Z = 3.22$; $p < .01$), and significantly later than for area ($Z = 3.60$; $p < .01$). Similarly, the age of maturity for area was significantly earlier than for density ($Z = 12.09$; $p < .001$) and time ($Z = 5.36$; $p < .01$), and the age of maturity for lines was significantly earlier than for density ($Z = 11.11$; $p < .001$) and time ($Z = 2.91$; $p < .01$).

The logistic growth model data suggest that the developmental trajectories for the five dimensions are not all identical: while ANS precision is at adult-like levels by adolescence, time and density do not fully develop until early adulthood; Length and Area, meanwhile, develop either at the same time as or slightly sooner than the ANS. Interestingly, because some dimensions develop and peak sooner than others, these results also suggest that the differences in accuracy between the five dimensions reach their peak in early childhood, and then, as Density and Time plateau, later stabilize in adolescence and adulthood.

3.5 | Partial correlations

Finally, we turn to the central question at hand – what drives the development of number accuracy and *w*? Although the developmental trajectory modelling confirms that Number develops at a different rate from the other dimensions, it does not provide evidence for any

TABLE 3 The parameters of the best-fit logistic growth models for each dimension across all participants

Dimension	DV	Peak value (a)	Age of onset (1/−b)	Growth rate (c)	Approximate age of maturity (SE)
Number	Accuracy	84.6 (2.14)	1.55 (5.59)	0.33 (0.04)	15.6 (2.14)
	w	0.16 (0.02)	1.58 (0.40)	0.56 (0.08)	11.71 (0.41)
Area	Accuracy	92.1 (2.16)	2.64 (4.11)	0.33 (0.06)	14.8 (2.17)
	w	0.09 (0.01)	0.68 (0.81)	0.89 (0.22)	8.37 (0.84)
Density	Accuracy	75.5 (3.90)	3.50 (3.79)	0.20 (0.05)	22.95 (3.91)
	w	0.26 (0.08)	2.35 (0.99)	0.23 (0.13)	24.13 (0.99)
Line length	Accuracy	93.7 (2.37)	1.81 (2.98)	0.39 (0.09)	13.02 (2.40)
	w	0.05 (0.01)	0.56 (0.41)	0.29 (0.07)	11.92 (0.46)
Time	Accuracy	82.3 (3.29)	2.22 (4.25)	0.25 (0.05)	19.39 (3.30)
	w	0.20 (0.03)	0.60 (0.82)	0.64 (0.14)	14.69 (0.83)

independence between Number and the development of the other four dimensions. In order to examine whether the development of Number is actually independent from the development of the other four dimensions, we carried out a set of partial correlations, examining whether Number continues to correlate with age even when the accuracy and *w* values of the other four dimensions are controlled for.

As shown in Figure 6, we found a significant correlation between Number accuracy and age, even when controlling for accuracy in both Area and Density ($r = .44$; $n = 85$; $p < .001$) and when controlling for accuracy in Length and Time ($r = .38$; $n = 80$; $p < .001$). Identical results were obtained with Number *w* values with Spearman partial correlations (Area/Density: rho = −.45, $n = 48$, $p < .001$; Length/Time: rho = −.37, $n = 52$, $p < .001$). These results remain identical when removing the group of 2- and 3-year-old children who are near chance performance (controlling for Area and Density: $n = 78$, $r = .37$, $p < .001$; controlling for Length and Time: $n = 67$, $r = .40$, $p < .001$). Thus, developmental improvements in the ANS appear to be independent of the development of non-numeric dimensions, including area, density, line length, and time, even over the large age range we tested.

We also found that each other dimension correlated with age even when controlling for the other two tested in that condition (Figure 6): Area accuracy correlated with age when controlling for Number and Density ($r = .21$; $n = 89$; $p < .05$), Density accuracy correlated with age when controlling for Number and Area ($r = .30$; $n = 89$; $p < .01$), Length correlated with age when controlling for Number and Time ($r = .30$; $n = 84$; $p < .01$), and Time correlated with age when controlling for Number and Length ($r = .37$; $n = 84$; $p < .001$). These results were near-identical – albeit slightly weaker – when examining *w* for each dimension: Area (rho = −.33; $p = .02$), Density (rho = −.25; $p = .08$) and Time (rho = −.35; $p < .05$). Length *w*, however, did not correlate with age when controlling for Time and Number ($r = -.16$; $n = 26$; $p = .26$); this result is probably because the *w* model eliminates children who guessed randomly, showing that line-length development probably occurs even earlier than the above accuracy analysis shows. As discussed below, this suggests that each dimension may, in turn, have an important and domain-specific developmental factor, and provides strong evidence against a generalized magnitude system.

As noted above, one consequence of this result is that domain-general factors, such as attention, parietal lobe development, or working memory, are unlikely to be the sole drivers of ANS development: because each discrimination task puts a significant load on these domain-general factors (e.g. the Time task required children to serially remember the duration of each tone before comparing it with the other), controlling for the four non-numeric dimensions should also incidentally control for the majority of these ancillary, domain-general individual differences.

3.6 | Role of inhibitory control

Recently, Gilmore et al. (2013) suggested that individual differences in the ANS might be largely reflective of differences in inhibitory control: Incongruent trials – those in which total surface area and number disagree in the answer – require the participant to actively inhibit the non-numeric dimensions competing for the response. Given that inhibitory control is well known to develop with age (Dowsett & Livesey, 2000; Munakata et al., 2011), could the developmental improvements in the ANS be merely reflecting these improvements in inhibitory control?

In order to assess whether the development of inhibitory control could account for our results, we examined children's performance on the Number task Congruent vs. Incongruent trials and correlated them with age. We found an approaching but non-significant difference between Congruent ($M = 71.4\%$; $SD = 15.4\%$) and Incongruent trials [$M = 69.14\%$; $SD = 18.4\%$; $t(197) = 1.56$, $p = .11$]. In addition, and despite the large sample size, we found no correlation between the Congruent/Incongruent difference and age (Figure 7; $r = -0.07$; $p = .35$). This difference remained non-significant even in our youngest sample of 3-year-olds [$t(23) = 0.03$; $p = .50$]. Similarly, a one-way ANOVA over all Age Groups with the Congruent/Incongruent difference failed to find a main effect of Age Group [$F(1, 196) = 1.80$; $p = .18$].

Our data point to the conclusion either that inhibitory control cannot account for the development of ANS precision, or alternatively that the Congruent/Incongruent difference is not a good measure of inhibitory control. In support of the measure, however, we did find a

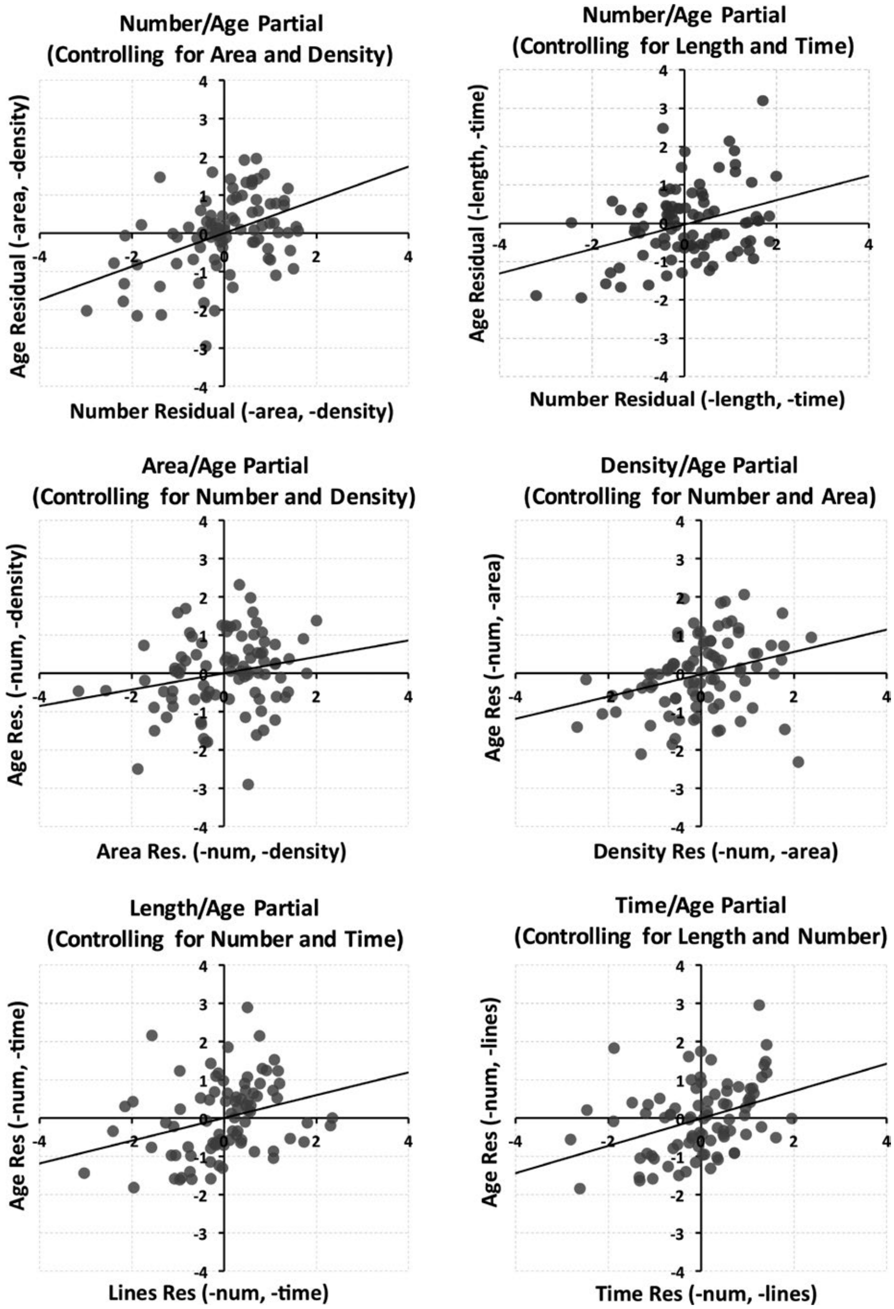

FIGURE 6 The partial correlations between each dimension and age when controlling for the other two dimensions. All six partial correlations are significant at $p < .05$

significant correlation between the Congruent/Incongruent difference in the Number task and Time task accuracy, even when controlling for Age [$r(82) = -.30$, $p < .01$], as would be expected given that the serial nature of the Time task places a higher load on working memory and executive control (Droit-Volet et al., 2008). We thus believe that the most likely conclusion is that the Congruent/Incongruent difference can (in part) depend on inhibitory control, but that this factor does not, in turn, drive the development of ANS precision.

4 | GENERAL DISCUSSION

From early in life, children have an intuitive and automatic representation of number – an Approximate Number System (ANS). Despite its ubiquity, the ANS undergoes significant development from infancy to adulthood. In order to determine which factors are the primary drivers of this development, we tested a large sample of 2- to 12-year-old children and adults on five discrimination tasks: number, area, density, length, and time. In addition to being the first to provide estimates of accuracy and Weber fractions for each of these dimensions across a broad age range, we report on three major findings. First, we find that the developmental trajectory of the ANS is distinct from the developmental trajectory of area, density, length, and time perception: while ANS precision is at adult-like levels by adolescence, time and density do not fully develop until early adulthood; length and area, meanwhile, develop either at the same time as or sooner than the ANS. Second, we find that ANS accuracy and Weber fractions continue to improve with age, even when individual differences in area, density, length,

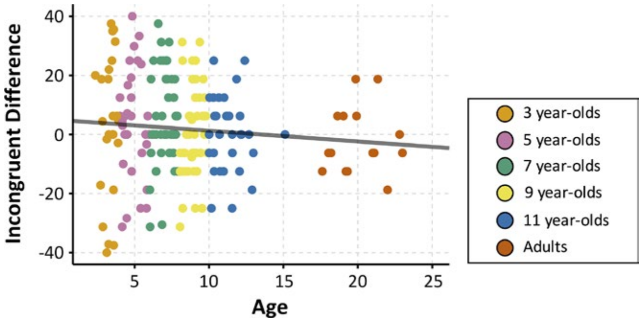

FIGURE 7 The correlation between Age and each participant's difference score between Congruent and Incongruent trial accuracy (thought to index inhibitory control). We found no significant correlation between age and this Incongruent trial difference

and time are controlled for. Hence, ANS development is independent from the four tested dimensions. Finally, we find that the difference between Incongruent and Congruent ANS trials – thought to index a participant's ability to inhibit non-numeric cues (Gilmore et al., 2013) – does not correlate with age and cannot explain the development of ANS precision.

Together, these results suggest that developmental improvements in the ANS are not driven by improvements in other non-numeric quantities, including area, density, length, and time. In turn, these results are incompatible with recent suggestions that ANS discrimination tasks are significantly biased by children's attention to or the use of non-numeric cues such as area, convex hull, etc. (e.g. Cantrell & Smith, 2013; Clayton, Gilmore, & Inglis, 2015; Defever et al., 2013; Gebuis & Reynvoet, 2012; Szucs et al., 2013). In other words, by controlling for the four dimensions most often claimed to be used during ANS tasks, we have shown that area, density, length, and time are not the sole or even the primary contributors to the developmental and individual differences in ANS precision. The distinct and independent developmental trajectories are also a challenge to the generalized magnitude system theory, and are instead more consistent with findings showing independence between number, space, and time, and place the relationships between various dimensions largely due to shared decision-making components (Anobile, Cicchini, & Burr, in press; Cappelletti, Freeman, & Cipolotti, 2011; Castelli, Glaser, & Butterworth, 2006; Odic et al., 2016; Odic & Halberda, 2015; Odic, Libertus, et al., 2013; Starr & Brannon, 2015b).

Our results are also inconsistent with the idea that improvements in the ANS are driven solely by domain-general improvements in attention, working memory, decision making, or parietal lobe maturation, as these factors would be shared between number and non-numeric quantity discrimination tasks (Droit-Volet et al., 2008; Odic et al., 2016; Pinel et al., 2004; Van Opstal & Verguts, 2013). Similarly, we found no evidence that ANS development is driven by children's improving ability to attend to numeric cues. Although it is possible that the ANS discrimination task places a load on some domain-general factor that is not used during area, density, length, or time discrimination, this explanation seems unlikely. Consider, for example, that we found no influence of – or even correlation with – performance on the density task, which even presented the stimuli as spatially separated blue and yellow dots. Ultimately, we do not wish to claim that domain-general factors have no impact on ANS performance, as other work has clearly shown that they do (Droit-Volet et al., 2008), but rather we believe that our work shows that these are not the only nor the primary factors that drive ANS development.

Instead, our results point to an important source of domain-specific maturation or experience that drives improvements in ANS precision. Our results do not, however, identify what these domain-specific factors are. Broadly speaking, these improvements may be related either to children's experience with number (e.g. learning to manipulate numbers in the context of mathematics), or to domain-specific maturation of the ANS itself. Previous work by Piazza et al. (2013), investigating ANS development in the Amazonian Munduruku tribe, showed that domain-specific education in mathematics significantly improves ANS precision. However, as the most dramatic changes in ANS precision occur prior to age 7 – when most children in our sample begin formal schooling – factors outside education must also play a role. One especially likely candidate is the maturation of the computations that encode the ANS (e.g. the object localization map proposed by Dehaene and Changeux, 1993), or the maturation of the specific neurons that encode number in the visual system and the parietal lobe. For example, Burr and Ross (2008; see also Anobile et al., in press; Odic & Halberda, 2015; Ross & Burr, 2010) argue that number is a primary visual feature, encoded by a set of dedicated neurons in the early visual cortex; as a result, the development of ANS precision may also be driven by the development and tuning of domain-specific neurons in the early visual system. Our work suggests that – whether at the level of encoding or of representations – the cognitive processes that most tightly track individual and developmental differences are dissociable for number compared with the other four dimensions. Exploring these possibilities will be an important avenue for future work.

Although the generalized magnitude system has often been hypothesized to persist into adulthood (e.g. Bueti & Walsh, 2009; Lu et al., 2011; Xuan et al., 2007), developmental psychologists have more recently claimed that newborns begin with a unified, 'one-bit' sense of magnitude that differentiates with experience and maturation (Cantrell & Smith, 2013; Lourenco & Longo, 2011). Although our data do not provide any direct evidence against a differentiation view, they do provide important caveats for such a theory. First, our results show that different dimensions develop at different rates – while area and length develop quickly, peaking in early childhood, density and time do not fully develop until young adulthood. Under the differentiation view, such a pattern would imply that children learn about length and area prior to learning about time and density. Hence, any potential learning mechanism that allows for differentiation between magnitudes will have to claim that learning about some spatiotemporal properties (e.g. area, length) is easier than learning about others. Second, our results provide a developmental trajectory for any potential differentiation: on average, we found that differences in accuracy and precision between the five tested dimensions peak at about age 7 (and subsequently decrease and stabilize). Finally, our work suggests that – once differentiated – each dimension follows its own trajectory, independent of the others. Thus, whatever the common currency or shared resources between number, area, density, length, and time are, they are not the ones primarily driving developmental or individual differences in accuracy and precision.

Interestingly, beyond identifying that ANS precision develops independently of the other four dimensions, we further found that each of the four non-numeric dimensions showed the same domain-specific developmental pattern as the ANS: area, density, length, and time all continued to develop with age, even when controlling for the ANS and the other tested dimension. These results broadly suggest that – to the extent that each dimension may be modular or independent from the other from preschool onwards – some non-shared factor drives the development of each. By analogy, consider that many of children's language abilities, such as knowledge of open-class vocabulary, develop significantly before many spatial manipulation and navigation abilities (Landau & Ferrara, 2013; see Karmiloff-Smith et al., 2004 for a discussion of distinct developmental trajectories within face perception). In a similar way, children's perception of area and length may mature significantly earlier compared with density and time.

One crucial point concerns whether children relied on any non-numeric cues during the Number discrimination task. Traditionally, differences between the congruent and incongruent trials have been used to suggest the use of non-numeric cues during ANS discrimination tasks (Dakin et al., 2011; Defever et al., 2013; Gebuis & Reynvoet, 2012; Gebuis & Van Der Smagt, 2011; Hurewitz, Gelman, & Schnitzer, 2006). We, however, failed to observe this difference at any age, including for the youngest section of our sample who performed at- or near-chance on the Number task. Furthermore, we find that individual differences in ANS acuity continue to correlate with age, even when controlling for the other dimensions. One possibility is that children in our sample used non-numeric cues in a haphazard and inconsistent way that could not be detected through the reported congruency effects and that, in addition, this use yields a non-linear relationship between number and the four dimensions that could not be statistically partialled out. Such an alternative explanation leaves much to be answered, such as what aspect of number performance continues to develop when controlling for individual differences in area, density, length, and time; and, if children do not use non-numeric cues on every trial, what they use instead, etc. Although the results presented here are most consistent with the idea of a dedicated system for approximating number, we do not have direct evidence against these alternative accounts and they remain open routes of inquiry.

An important limitation of this work is that we could study the children only cross-sectionally. Although we do not expect strong cohort effects, a true understanding of developmental trajectories would ideally use longitudinal data. Longitudinal data would also capture whether a developmental boost in one dimension leads to a cascade of changes in another.

In conclusion, our results suggest that the development of ANS precision is independent of children's perception of area, density, length, or time. These results place the development of number, space, and time representation into a broader and richer developmental context – showing both the similarities and differences in how these representations mature – and provide ample opportunity for future work aimed at understanding how number representations are actually extracted from a visual scene, and how domain-specific ANS representations may contribute to a variety of other cognitive abilities, including mathematics.

ACKNOWLEDGEMENTS

Thank you to Andy Baron, the LivingLab, and the Vancouver Telus Scienceworld museum for allowing us to test children in the museum. Thank you to all the parents and children who participated in the study. Thank you to Mimi Zhang and Nicole Yang for help in data collection, and to Edwina Picon, Johnny Jewel and Aaron Maine for help in the writing of this manuscript.

NOTES

[1] Owing to an experimenter error, eight children were running in a Number/Area/Lines condition. Excluding these children changes none of the results reported here, and in order to maximize our sample all of these children were kept in the analyses.

[2] An inspection of our figures and tables reveals that our group of 2- and 3-year-olds was generally around chance-levels, suggesting that they may not have understood the task. Nevertheless, removing this group of children did not change any of our results (e.g. the correlation between ANS and age becomes $r = .61$, $p < .001$, when this group of 24 participants is removed).

REFERENCES

Abreu-Mendoza, R.A., & Arias-Trejo, N. (2015). Numerical and area comparison abilities in Down syndrome. *Research in Developmental Disabilities*, *41–42*, 58–65.

Anobile, G., Cicchini, M., & Burr, D. C. (in press). Number as a primary perceptual attribute: a review. *Perception*.

Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Brannon, E.M., Lutz, D., & Cordes, S. (2006). The development of area discrimination and its implications for number representation in infancy. *Developmental Science*, *9*, F59–F64.

Brannon, E.M., & Terrace, H.S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, *282*, 746–749.

Brannon, E.M., Wusthoff, C.J., Gallistel, C.R., & Gibbon, J. (2001). Numerical subtraction in the pigeon: Evidence for a linear subjective number scale. *Psychological Science*, *12*, 238–243.

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 1831–1840.

Burr, D.C., & Ross, J. (2008). A visual sense of number. *Current Biology*, *18*, 425–428.

Cantlon, J.F., & Brannon, E.M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, *17*, 401–406.

Cantrell, L., & Smith, L.B. (2013). Open questions and a proposal: A critical review of the evidence on infant numerical abilities. *Cognition*, *128*, 331–352.

Cappelletti, M., Freeman, E.D., & Cipolotti, L. (2011). Numbers and time doubly dissociate. *Neuropsychologia*, *49*, 3078–3092.

Cappelletti, M., Gessaroli, E., Hithersay, R., Mitolo, M., Didino, D., Kanai, R., … & Walsh, V. (2013). Transfer of cognitive training across magnitude dimensions achieved with concurrent brain stimulation of the parietal lobe. *The Journal of Neuroscience*, *33*, 14899–14907.

Castelli, F., Glaser, D., & Butterworth, B. (2006). Discrete and analogue quantity processing in the parietal lobe: A functional MRI study. *Proceedings of the National Academy of Sciences, USA*, *103*, 4693–4698. http://doi.org/10.1073/pnas.0600444103

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.

Cordes, S., & Brannon, E.M. (2008). Quantitative competencies in infancy. *Developmental Science*, *11*, 803–808.

Dakin, S.C., Tibber, M.S., Greenwood, J.A., Kingdom, F.A.A., & Morgan, M.J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences, USA*, *108*, 19552–19557.

Defever, E., Reynvoet, B., & Gebuis, T. (2013). Task- and age-dependent effects of visual stimulus properties on children's explicit numerosity judgments. *Journal of Experimental Child Psychology*, *116*, 216–233.

Dehaene, S., & Changeux, J.-P. (1993). Development of elementary numerical abilities: A neural model. *Journal of Cognitive Neuroscience*, *5*, 390–407.

DeWind, N.K., Adams, G.K., Platt, M.L., & Brannon, E.M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition*, *142*, 247–265.

Dormal, V., & Pesenti, M. (2012). Processing numerosity, length and duration in a three-dimensional Stroop-like task: Towards a gradient of processing automaticity? *Psychological Research*, *77*, 116–127.

Dowsett, S.M., & Livesey, D.J. (2000). The development of inhibitory control in preschool children: Effects of "executive skills" training. *Developmental Psychobiology*, *36*, 161–174.

Droit-Volet, S., Clément, A., & Fayol, M. (2008). Time, number and length: Similarities and differences in discrimination in adults and children. *Quarterly Journal of Experimental Psychology*, *61*, 1827–1846.

Droit-Volet, S., & Wearden, J.H. (2001). Temporal bisection in children. *Journal of Experimental Child Psychology*, *80*, 142–159.

Durgin, F.H. (1995). Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 149.

Emerson, R.W., & Cantlon, J.F. (2015). Continuity and change in children's longitudinal neural responses to numbers. *Developmental Science*, *18*, 314–326.

Emmerton, J. (1998). Numerosity differences and effects of stimulus density on pigeons' discrimination performance. *Learning and Behavior*, *26*, 243–256.

Fabbri, M., Cancellieri, J., & Natale, V. (2012). The A Theory Of Magnitude (ATOM) model in temporal perception and reproduction tasks. *Acta Psychologica*, *139*, 111–123.

Feigenson, L., Dehaene, S., & Spelke, E.S. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.

Fischer, M.H., Castel, A.D., Dodd, M.D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience*, *6*, 555–556.

Frank, M.C., Everett, D.L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, *108*, 819–824.

Gebuis, T., & Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General*, *141*, 642.

Gebuis, T., & Van Der Smagt, M.J. (2011). False approximations of the approximate number system. *PLoS One*, *6*, e25405.

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*, e67374.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, *306*, 496.

Grondin, S., & Killeen, P. (2009). Tracking time with song and count: Different Weber functions for musicians and nonmusicians. *Attention, Perception, and Psychophysics*, *71*, 1649–1654.

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*, 1457–1465.

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, USA*, *109*, 11116–11120.

Halberda, J., & Odic, D. (2014). The precision and internal confidence of our approximate number thoughts. In D.C. Geary, D. Berch, & K. Koepke (Eds.), *Evolutionary origins and early development of number processing* (pp. 305–332). Chicago: Academic Press.

de Hevia, M.D., Izard, V., Coubart, A., Spelke, E.S., & Streri, A. (2014). Representations of space, time, and number in neonates. *Proceedings of the National Academy of Sciences, USA*, *111*, 4809–4813.

Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences, USA*, *103*, 19599–19604.

Inglis, M., & Gilmore, C. (2014). Indexing the approximate number system. *Acta Psychologica*, *145*, 147–155.

Izard, V., Sann, C., Spelke, E.S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences, USA*, *106*, 10382–10385.

Jordan, K.E., & Brannon, E.M. (2006). The multisensory representation of number in infancy. *Proceedings of the National Academy of Sciences, USA*, *103*, 3486–3489.

Karmiloff-Smith, A., Thomas, M., Annaz, D., Humphreys, K., Ewing, S., Brace, N., … & Campbell, R. (2004). Exploring the Williams syndrome face-processing debate: The importance of building developmental trajectories. *Journal of Child Psychology and Psychiatry*, *45*, 1258–1274.

Keller, L., & Libertus, M. (2015). Inhibitory control may not explain the link between approximation and math abilities in kindergarteners from middle class families. *Developmental Psychology*, *685*.

Landau, B., & Ferrara, K. (2013). Space and language in Williams syndrome: Insights from typical development. *Cognitive Science*, *4*, 693–706.

Libertus, M.E., Odic, D., & Halberda, J. (2012). Intuitive sense of number correlates with math scores on college-entrance examination. *Acta Psychologica*, *141*, 373–379.

Lindskog, M., Winman, A., & Juslin, P. (2014). The association between higher education and approximate number system acuity. *Frontiers in Psychology*, *5*, 462.

Lourenco, S.F., & Longo, M.R. (2010). General magnitude representation in human infants. *Psychological Science*, *21*, 873–881.

Lourenco, S.F., & Longo, M.R. (2011). Origins and development of generalized magnitude representation. *Space, Time, and Number in the Brain: Searching for the Foundations of Mathematical Thought*, 225–244.

Lu, A., Mo, L., & Hodges, B.H. (2011). The weight of time: Affordances for an integrated magnitude system. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1855–1866.

Marceau, K., Ram, N., Houts, R.M., Grimm, K.J., & Susman, E.J. (2011). Individual differences in boys' and girls' timing and tempo of puberty: Modeling development with nonlinear growth models. *Developmental Psychology*, *47*, 1389–1409.

Marques, J.F., & Dehaene, S. (2004). Developing intuition for prices in euros: Rescaling or relearning prices? *Journal of Experimental Psychology: Applied*, *10*, 148–155.

Mayer, D.L., & Dobson, V. (1982). Visual acuity development in infants and young children, as assessed by operant preferential looking. *Vision Research*, *22*, 1141–1151.

Meck, W.H., & Church, R.M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*, 320–334.

Munakata, Y., Herd, S.A., Chatham, C.H., Depue, B.E., Banich, M.T., & O'Reilly, R.C. (2011). A unified framework for inhibitory control. *Trends in Cognitive Sciences*, *15*, 453–459.

Nieder, A. (2005). Counting on neurons: The neurobiology of numerical competence. *Nature Reviews. Neuroscience*, *6*, 177.

Nieder, A., & Miller, E.K. (2004). Analog numerical representations in rhesus monkeys: Evidence for parallel processing. *Journal of Cognitive Neuroscience*, *16*, 889–901.

Odic, D., & Halberda, J. (2013). The independence of visual number and area processing: evidence from psychophysics, development, and eye-tracking. *Journal of Vision*, *13*, 1259–1259.

Odic, D., & Halberda, J. (2015). Eye movements reveal distinct encoding patterns for number and cumulative surface area in random dot arrays. *Journal of Vision*, *15*, 5–5.

Odic, D., Im, H.Y., Eisinger, R., Ly, R., & Halberda, J. (in press). PsiMLE: A maximum-likelihood approach to estimating psychophysical scaling and variability more reliably, efficiently, and flexibly. *Behavior Research and Methods*.

Odic, D., Libertus, M.E., Feigenson, L., & Halberda, J. (2013a). Developmental change in the acuity of approximate number and area representations. *Developmental Psychology*, *49*, 1103–1112.

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.

Odic, D., Pietroski, P., Hunter, T., Lidz, J., & Halberda, J. (2013b). Young children's understanding of "more" and discrimination of number and surface area. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 451–461.

Piazza, M., Facoetti, A., Trussardi, A.N., Berteletti, I., Conte, S., Lucangeli, D., … & Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*, 33–41.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, *44*, 547–555.

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*, 1037–1043.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian Indigene group. *Science*, *306*, 499–503.

Piffer, L., Agrillo, C., & Hyde, D.C. (2011). Small and large number discrimination in guppies. *Animal Cognition*, *15*, 215–221.

Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, *41*, 983–993.

Ram, N., & Grim, K. (2015). Growth curve modeling and longitudinal factor analysis. In W. Overton & P.C.M. Molenaar (Eds.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (7th edn). Hoboken, NJ: Wiley.

Ratcliff, R., Love, J., Thompson, C.A., & Opfer, J.E. (2012). Children are not like older adults: a diffusion model analysis of developmental changes in speeded responses. *Child Development*, *83*, 367–381.

Roitman, J.D., Brannon, E.M., & Platt, M.L. (2007). Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS Biology*, *5*, e208.

Ross, J., & Burr, D.C. (2010). Vision senses number directly. *Journal of Vision*, *10*, 10.

Starr, A., & Brannon, E.M. (2015a). Developmental continuity in the link between sensitivity to numerosity and physical size. *Journal of Numerical Cognition*, *1*, 7–20.

Starr, A., & Brannon, E.M. (2015b). Evidence against continuous variables driving numerical discrimination in infancy. *Frontiers in Psychology*, *6*.

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, USA*, *110*, 18116–18120.

Szucs, D., Nobes, A., Devine, A., Gabriel, F.C., & Gebuis, T. (2013). Visual stimulus parameters seriously compromise the measurement of approximate number system acuity and comparative effects between adults and children. *Frontiers in Psychology*, *4*. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3715731/

Tudusciuc, O., & Nieder, A. (2007). Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. *Proceedings of the National Academy of Sciences, USA*, *104*, 14513–14518.

Van Opstal, F., & Verguts, T. (2013). Is there a generalized magnitude system in the brain? Behavioral, neuroimaging, and computational evidence. *Frontiers in Psychology*, *4*. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3711208/

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*, 483–488.

Witt, J.K., Linkenauger, S.A., Bakdash, J.Z., & Proffitt, D.R. (2008). Putting to a bigger hole: Golf performance relates to perceived size. Psychonomic Bulletin and Review, 15, 581–585.

Wood, G., Willmes, K., Nuerk, H.-C., & Fischer, M.H. (2008). On the cognitive link between space and number: A meta-analysis of the SNARC effect. *Psychology Science Quarterly*, *50*, 489–525.

Xenidou-Dervou, I., De Smedt, B., van der Schoot, M., & van Lieshout, E.C.D.M. (2013). Individual differences in kindergarten math achievement: The integrative roles of approximation skills and working memory. *Learning and Individual Differences*, *28*, 119–129.

Xu, F., & Spelke, E.S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11.

Xuan, B., Zhang, D., He, S., & Chen, X. (2007). Larger stimuli are judged to last longer. *Journal of Vision*, *7*, 2.

Yuodelis, C., & Hendrickson, A. (1986). A qualitative and quantitative analysis of the human fovea during development. *Vision Research*, *26*, 847–855.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Odic D. Children's intuitive sense of number develops independently of their perception of area, density, length, and time. *Dev Sci*. 2017;00:e12533. doi:10.1111/desc.12533.