Cognitive Development 47 (2018) 107–116

Contents lists available at ScienceDirect

Cognitive Development

journal homepage: www.elsevier.com/locate/cogdev

Hysteresis-induced changes in preverbal infants' approximate number precision

Jinjing (Jenny) Wang[a,*], Melissa E. Libertus[a,b], Lisa Feigenson[a]

[a] Department of Psychological and Brain Sciences, Johns Hopkins University, 3400 N Charles Street, Baltimore, MD 21218, USA
[b] Department of Psychology and Learning Research and Development Center, University of Pittsburgh, 3939 O'Hara Street, Pittsburgh, PA 15260, USA

ARTICLE INFO

Keywords:
Approximate number system
Hysteresis
Malleability

ABSTRACT

Infants represent the approximate number of items in visual and auditory arrays. These number representations are noisy: for example, whereas 6-month-olds discriminate numerosities that differ by a 1:2 ratio (e.g., 8 vs. 16 dots), they fail to discriminate a 2:3 ratio (e.g., 8 vs. 12 dots) until 9 months old. How should we understand the nature of the representations underlying this performance? One possibility is that the precision of approximate number representations is fixed at a given age; alternatively, precision may be dynamic and context dependent. Here we asked whether one aspect of context—prior numerical experience—influences preverbal approximate number precision. We familiarized 6-month-old infants with pairs of images containing different numerosities. Critically, as trials progressed, the ratio of the two numerosities within each pair also gradually progressed– either from highly discriminable ratios to ratios that became harder to discriminate, or vice versa. After this ordered numerical training, we tested infants' ability to discriminate numerosities differing by a challenging 2:3 ratio, with which infants of this age typically fail. In three experiments, we found that 6-month-old infants successfully discriminated the 2:3 ratio after starting with easy ratios and progressing to hard ones, but not after starting with hard ratios and progressing to easy ones, despite experiencing identical numerosities across these two conditions. This "numerical hysteresis" effect was feedback-dependent: infants only succeeded with the scaffolded training when they received trial-by-trial feedback. Together, these results provide evidence for a temporary modulation of infants' number sense.

1. Introduction

Preverbal infants have an intuitive sense of number. For example, 6-month-old infants who are habituated to arrays of eight dots (or sounds) dishabituate when later shown 16, and vice versa (Lipton & Spelke, 2003; Xu & Spelke, 2000). When presented with two streams of images, one repeatedly showing images of 8 dots, the other alternating between 8 and 16 dots, 6-month-old infants look longer at the numerically changing stream than the numerically constant stream (Libertus & Brannon, 2010). These findings and others, which obtain when controlling for non-numerical stimulus dimensions such as cumulative surface area, individual dot size, and array density, show that 6-month-olds represent the numerical difference between, for example, 8 and 16.

This ability to discriminate numerosities without verbally counting is underwritten by the Approximate Number System, or ANS (Dehaene, 1997; Feigenson, Dehaene, & Spelke, 2004), a hallmark feature of which is ratio-dependence. That is, the discriminability of two numerosities represented by the ANS depends on the numerosities' ratio. Six-month-old infants exhibit this classic

* Corresponding author at: Department of Psychological and Brain Sciences, Johns Hopkins University, Ames 128, 3400 N Charles Street, Baltimore, MD, 21218, USA.

E-mail address: jenny.wang@jhu.edu (J.J. Wang).

https://doi.org/10.1016/j.cogdev.2018.05.002
Received 17 September 2017; Received in revised form 20 April 2018; Accepted 8 May 2018
Available online 14 May 2018
0885-2014/

performance signature—they reliably discriminate 8 dots from 16, and 16 from 32 (a 1:2 ratio), but under the same testing conditions they fail to discriminate 8 from 12 or 16 from 24 (a 2:3 ratio) (Feigenson, 2011; Libertus & Brannon, 2010; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005). This ratio-dependence does not reflect visual limitations, as 6-month-olds show the same precision when presented with auditory sequences (Lipton & Spelke, 2003). Notably, ANS precision improves over development. Whereas 6-month-old infants fail to discriminate numerosities that differ by a 2:3 ratio, at 9 months of age they succeed in both the visual and auditory modalities (Libertus & Brannon, 2010; Lipton & Spelke, 2003; Xu & Arriaga, 2007). This improvement in ANS precision continues throughout childhood, with precision eventually peaking in early adulthood (Halberda & Feigenson, 2008; Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Odic, Libertus, Feigenson, & Halberda, 2013), although individual differences in precision are observed across development (Halberda, Mazzocco, & Feigenson, 2008; Libertus & Brannon, 2010; Starr, Libertus, & Brannon, 2013).

How should we understand the nature of approximate number precision? Current views of the ANS suggest that a stimulus (e.g., 16 dots) will, if presented repeatedly, trigger a range of mental activations that form a distribution with some mean (e.g., at approximately 16) and standard deviation. It is this standard deviation (i.e., representational noisiness) of the distribution of activations that is thought to change with development, and that varies between individuals. This conceptualization often assumes ANS precision to be relatively fixed for a given person, or for a given age group. Yet some evidence suggests that ANS precision is more dynamic than this.

One piece of support for dynamic ANS precision comes from the enhancing effects of redundant information. In one study, researchers compared infants' ability to discriminate the numerosities of stimuli that were synchronously presented in two sensory modalities (a ball seen to bounce varying numbers of times, and that made a sound each time it bounced) with their ability when the same stimulus was presented only visually. Whereas 6-month-old infants failed to discriminate numerosities differing by a 2:3 ratio when stimuli contained only visual numerical information (as in previous work that also used a single modality), infants succeeded when presented with synchronous visual and auditory input (Jordan, Suanda, & Brannon, 2008). Further work found that even multiple sources of information from a single sensory modality can benefit numerical performance. Baker, Mahamane, & Jordan, 2014 observed that after habituating to a ball of a constant size bouncing a constant number of times, 6-month old infants dishabituated to a ball of a novel size, bouncing a novel number of times (i.e., infants detected a change in numerosity that was confounded with a change in object area), even with a challenging 2:3 ratio, despite failing to detect a 2:3 change in numerosity or surface area alone.

A second line of evidence for the dynamic nature of ANS precision comes from studies that borrowed an approach from dynamical psychophysics to ask whether trial order affects numerical discrimination (Odic, Hock, & Halberda, 2014; Wang, Odic, Halberda, & Feigenson, 2016). Five-year-old children saw two arrays of dots and had to decide which array had more without counting. Some of the trials contained ratios that were easily discriminable (e.g., 5 blue dots versus 10 yellow); others contained ratios that were much more difficult (e.g., 9 blue dots versus 10 yellow). Instead of ratio difficulty varying randomly across trials, as in previous studies using this method, difficulty changed systematically as the task unfolded. Some children started with the easiest ratios (e.g., 1:2) and gradually moved to the hardest ones (e.g., 9:10), whereas others started with the hardest ratios and gradually moved to the easiest ones. All children received trial-by-trial feedback telling them whether their choices were correct. In this task, children who started with the easier ratios performed significantly better than children who started with the harder ratios, despite all children ultimately completing the same number of trials at each ratio, and all other aspects of the presentation being identical (Odic et al., 2014; Wang et al., 2016). Notably, even when the feedback was reversed (i.e., when correct responses triggered negative auditory feedback, and incorrect responses triggered positive feedback), children who first experienced the easier ratios exhibited improved ANS precision. However, when feedback was removed altogether, the effect disappeared (Odic et al., 2014). These findings suggest that the recent history of one's ANS discriminations affects representational precision—a phenomenon termed "ANS hysteresis," following the dynamical psychophysics literature.

Can recent ANS experience also produce precision shifts in infants? One reason to be doubtful is that ANS hysteresis might rely on conscious, top down cognition. For example, a possible interpretation of the findings of Odic et al. (2014) and Wang et al. (2016) is that receiving feedback changed the way children felt about their own ability to perform the task. Children who began with easier discriminations (and who therefore started the task with more positive feedback) might have felt more confident and positive ("I'm pretty good at this game!"), and hence could have been more motivated or attentive than children who began with harder discriminations. The disappearance of ANS hysteresis without feedback is consistent with this view. Although Odic et al. (2014) found that "reversed feedback" still increased ANS performance in children who started with easier numerical ratios, it is conceivable that children recognized the feedback as reversed and discounted or mentally un-reversed it. One way to test whether self-perception caused the ANS hysteresis effect is to test preverbal babies, who have a less developed sense of their own task competence. If a conscious sense of self-efficacy underlies ANS hysteresis, then the effect should not obtain with infants.

A second reason to ask whether ANS hysteresis is observed in infants is to better understand the conditions under which ANS precision is malleable. Jordan and colleagues found that 6-month-olds showed enhanced precision when multiple sources of quantity information were available (e.g., when habituation and test stimuli differed in both visual and auditory numerosity, or in both visual numerosity and surface area) (Baker et al., 2014; Jordan et al., 2008). But numerical information often is only available from a single sensory modality—for example, a distant flock of birds can be visually enumerated, but is likely to be silent. And approximate number can be represented in the absence of any correlated visual features like area (e.g., estimating the number of ideas proposed in a meeting). These issues led us to ask whether, early in life, ANS precision still can be changed when intersensory information is unavailable, and when numerosity changes are unconfounded from changes in continuous dimensions.

In the present study we asked whether numerical precision is modulated by infants' prior history of numerical discrimination. Infants saw two screens on which dot arrays were continuously flashed. On one screen, the numerosity of the arrays remained constant across flashes (although other aspects of the arrays varied); on the other screen, numerosity alternated between two values. The dependent measure was infants' relative visual interest in the two screens; a significant preference for the numerically changing screen would indicate that infants successfully discriminated the numerosities in the alternating arrays (Libertus & Brannon, 2010; Ross-sheehy, Oakes, & Luck, 2003). Critically, infants in Experiment 1 saw arrays that, on the numerically changing screen, started with highly discriminable ratios, and over the course of the experiment gradually progressed towards ratios that were much harder to discriminate. Following this ordered familiarization, during which infants were rewarded with music for looking at the numerically changing screen, all infants were tested with an unreinforced 2:3 ratio—with which, as reviewed above, 6-month-old infants typically fail, including in the change detection task (Xu & Spelke, 2000; Feigenson, 2011; Libertus & Brannon, 2010; Lipton & Spelke, 2003; Xu et al., 2005). We predicted that if hysteresis affects ANS precision even in infancy, infants would succeed at this challenging discrimination after experiencing the "easy-first" ANS trial ordering. Next, in Experiment 2 we asked whether infants' success in Experiment 1 was driven by trial order or merely by experiencing many different numerical ratios over the course of the experiment. We showed infants the identical trials in a reversed order: infants started with the hardest ratios and gradually progressed towards easier ones, and then were tested with the challenging 2:3 ratio discrimination. Finally, in Experiment 3 we examined the role of feedback on numerical hysteresis in infants; infants saw the same "easy-first" trial ordering as in Experiment 1, but without contingent reinforcement during the familiarization. To preview our findings, we found that experience with numerical discriminations that gradually increased in difficulty, coupled with trial-by-trial feedback, enhanced infants' numerical approximation performance.

2. EXPERIMENT 1

In Experiment 1, 6-month old infants first were familiarized to a series of approximate number discriminations that first differed by an easy-to-discriminate ratio, and gradually progressed to a hard-to-discriminate ratio. Throughout this familiarization, infants received musical reinforcement for looking at the numerically changing array. We probed for an effect of ANS hysteresis by then testing infants on a challenging 2:3 ratio discrimination (with which they have failed in previous studies (Feigenson, 2011; Libertus & Brannon, 2010; Lipton & Spelke, 2003; Xu & Spelke, 2000; Xu et al., 2005), without reinforcement.

2.1. Method

2.1.1. Participants

Thirty-two full-term infants between 5 and 7 months old participated (14 girls; mean age 6 months 14 days (*SD* = 22 days). Of these, 22 were identified by their parent(s) as White, six as Black, one as Asian, one as Hispanic, and one as "another race/ ethnicity." One infant's race and ethnicity were not identified. Children received a small gift (e.g., t-shirt, book, or toy) to thank them for their participation, and parents provided written informed consent prior to the experiment. Two additional infants were tested but were excluded from analysis due to fussiness (n = 1) or equipment failure (n = 1).

2.1.2. Design and stimuli

We modeled our design on previous infant change detection tasks (Libertus & Brannon, 2010; Ross-sheehy et al., 2003). On each trial infants saw two streams of images as shown in Fig. 1. One of the image streams contained arrays that were always constant in numerosity (Numerically Constant stream), whereas the other stream contained arrays that alternated between two numerosities (Numerically Changing stream). Within each trial, the ratio between these two alternating numerosities did not vary. However, across trials this ratio varied among the following: 6 vs. 36 (1:6 ratio), 9 vs. 27 (1:3 ratio), 10 vs. 20 (1:2 ratio), 15 vs. 25 (3:5 ratio), and 12

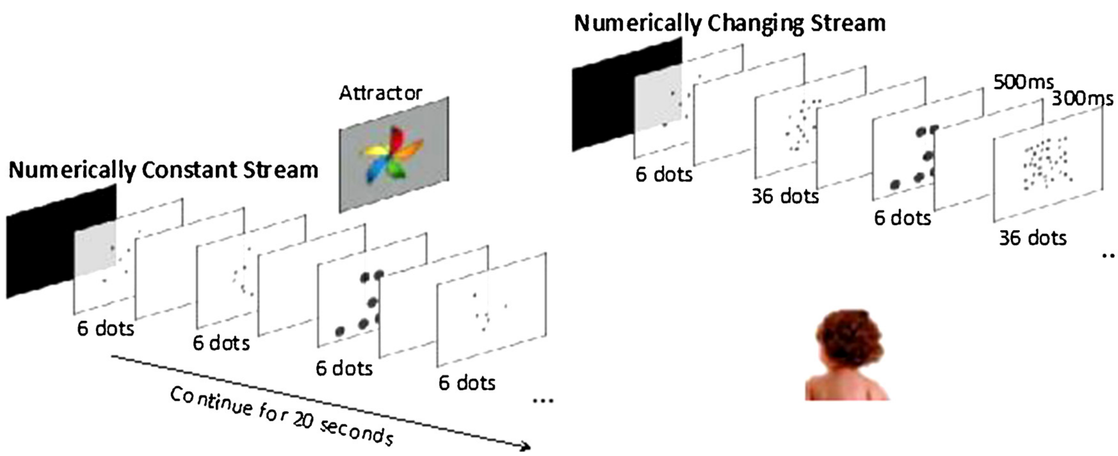

Fig. 1. On each trial infants saw two streams of images on separate monitors. The Numerically Constant Stream contained repeated presentations of a single numerosity (e.g., 6 dots), whereas the Numerically Changing Stream alternated between two numerosities that differed by a fixed ratio (e.g., 6 and 36 dots). Each trial continued for 20 s.

vs. 16 (3:4 ratio). All infants saw two trials of each of these five ratios: one with the Numerically Changing stream on the left-hand monitor and one with the Numerically Changing stream on the right-hand monitor. This resulted in a total of 10 training trials. Critically, the training trials started with the easiest numerical discrimination (1:6 ratio) and monotonically progressed to the hardest discrimination (3:4 ratio). After the 10 training trials, infants saw two test trials involving a 2:3 ratio (14 vs. 21 dots), one with the Numerically Changing stream on the left and the other with the Numerically Changing stream on the right. Each of the 10 training trials and two test trials lasted for 20 s.

The arrays in both the Numerically Changing and Numerically Constant streams contained items that varied in size and spatial position. For half of the infants, stimuli were solid black circles (as in previous studies by Libertus & Brannon, 2010; Xu & Spelke, 2000; Xu et al., 2005); for the other half, stimuli were rainbow-colored squares (as in previous studies by Brannon, 2002; Suanda, Tompson, & Brannon, 2008). This allowed us to ask whether any effects of ANS hysteresis would be seen across multiple stimulus types. Average circle diameter and average square side length were both 1.25 cm, ranging from 0.38 cm to 2.27 cm. Average density was 0.03 items/cm^2, ranging from 0.007 items/cm^2 to 0.12 items/cm^2. Following the design of Libertus and Brannon (2010), the arrays were presented for 500 ms followed by 300 ms of blank screen. On half of the 500-ms presentations, the images shown on the two monitors were identical (i.e., contained arrays of the same numerosity, size, and spatial configuration). These alternated with presentations in which number differed between the monitors. Of these, one-third were matched across the two screens on total surface area, one-third were matched on individual item size, and one-third were matched on total perimeter. Orthogonally, half of the differing arrays were matched on density. These controls ensured that the Numerically Changing stream was no more variable than the Numerically Constant stream in these non-numerical stimulus dimensions (see Libertus & Brannon, 2010). For half of the infants in each condition, the Numerically Constant stream contained the larger number in the alternating pair (e.g., 36 items, when the Numerically Changing stream consisted of 6 and 36 items), for the other half it contained the smaller number (e.g., 6 items).

2.1.3. Apparatus and procedure

Infants sat on a caregiver's lap approximately 100 cm from the midpoint of two 17-inch computer screens spaced 104 cm apart, facing a small light-up fan between the screens. Caregivers were instructed to look at the back of infants' heads or to close their eyes. An experimenter stood concealed by a curtain, where she could see infants' faces on a hidden monitor but could not see the stimuli.

At the beginning of each trial, the fan illuminated and spun to attract infants' attention to the midpoint between the two screens. As soon as infants had fixated this attractor, the experimenter initiated the trial. The side of the Numerically Changing stream (left versus right monitor) alternated between trials, and the side on which the Numerically Changing stream started was counterbalanced between infants.

Given that ANS hysteresis in older children appears to require response feedback (Odic et al., 2014), we incorporated feedback into the infant change detection paradigm. As each trial progressed, an experimenter (who was blind to the side on which the Numerically Changing stream was presented) watched the monitor feed of infants' faces and pressed a key to indicate whether infants were looking at the left monitor, right monitor, or neither. A custom-made program recorded this input. During the training trials, whenever infants looked at the Numerically Changing stream for longer than 500 ms, music played from a centrally located speaker until infants looked away from that stream, at which time the music stopped immediately until infants looked back at that stream. After the 10 training trials, infants received two 20-second test trials that were unreinforced: music played throughout, regardless of where infants looked.

Infants' looking was recorded digitally. An experienced observer coded infants' looking off-line using a program written in RealBasic (Libertus, 2008). A second observer coded the videos from 25% of the participants' testing sessions. Reliability between the two observers was 0.97.

2.2. Results and discussion

Our primary question was whether infants would show significant discrimination of the challenging 2:3 numerical ratio at test. To examine this we first calculated a numerical change preference score that reflected infants' relative interest in the Numerically Changing versus Constant streams (Fig. 2). For each infant, we computed the difference in looking to the Numerically Changing versus Numerically Constant streams, and divided this by the sum of looking to either screen ((looking to Numerically Changing – looking to Numerically Constant)/ (looking to Numerically Changing + looking to Numerically Constant)); we then averaged this proportional preference score across infants. Infants' numerical change preference score was significantly greater than zero, $M = 17.44\%$, $SD = 27.86\%$, $t(31) = 3.54$, $p = .001$, *Cohen's* $d = 0.63$, indicating that infants looked significantly longer to the Numerically Changing stream, and therefore successfully discriminated the 2:3 test ratio (Fig. 3).

Next we asked whether infants' numerical change preference scores were influenced by stimulus type, size of the Numerically Constant numerosity (recall that for half of the participants the Numerically Constant stream contained the smaller of the two alternating numerosities in the Numerically Changing stream, and for the other half it contained the larger), or spatial congruity between training and test trials (whether the Numerically Changing stream was on the same side during the last training trial and the first test trial). A 2 (Stimuli: black circles vs. rainbow squares) x 2 (Size of Numerically Constant Numerosity: larger vs. smaller) x 2 (Spatial Congruity: same vs. different side) ANOVA on infants' preference scores during the test trials revealed no effect of Stimuli, $F(1,24) = 0.052$, $p = .82$, $\eta_p^2 = .002$, no effect of Size of Numerically Constant Numerosity, $F(1,24) = 0.11$, $p = .75$, $\eta_p^2 = .004$, no effect of Spatial Congruity, $F(1,24) = 2.07$, $p = .16$, $\eta_p^2 = .080$, and no interactions among these, Fs < 1, ps $> .43$. The lack of an effect of Spatial Congruity shows that infants did not merely carry over their direction of looking from the last training trial (in which reinforcement was given while infants were looking at the Numerically Changing stream) to the first test trial (in which there was no

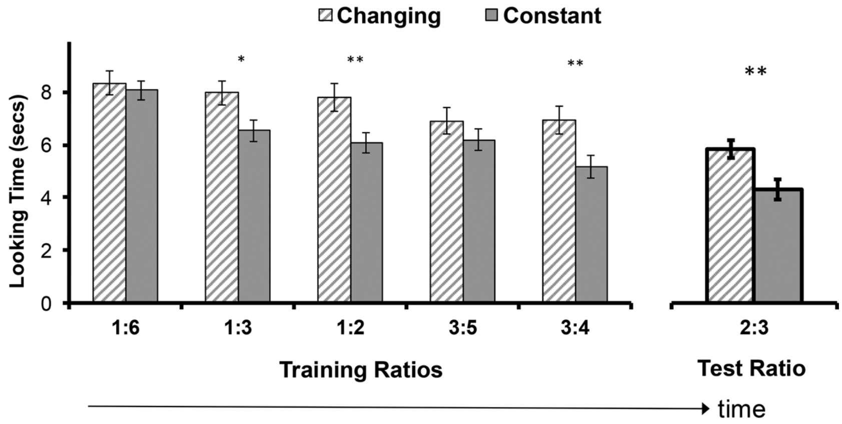

Fig. 2. Average looking at the Numerically Changing and Numerically Constant streams during the training trials and the 2:3 ratio test trials in Experiment 1. Error bars represent the standard error of the mean. ***p* < .01. **p* < .05.

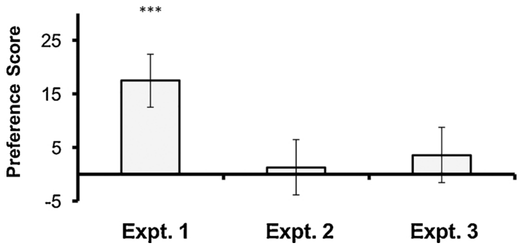

Fig. 3. Average numerical preference score in the 2:3 ratio test trials of Experiments 1–3. Error bars represent the standard error of the mean. ****p* = .001.

reinforcement), because infants showed similar discrimination of the 2:3 ratio regardless of whether it appeared in a new location or not. Twenty-five out of 32 infants preferred the Numerically Changing stream at test (binomial exact test $p = .002$).

In Experiment 1 we found that 6-month-old infants were able to discriminate numerosities differing by a 2:3 ratio after experiencing a training sequence in which easily discriminable ratios were followed by a series of successively harder-to-discriminate ratios. This shows that early approximate number precision is malleable, and can be affected by infants' recent history with large numerical quantities. Our findings with preverbal infants suggest that the previously observed effect of ANS hysteresis in older children (Odic et al., 2014; Wang et al., 2016) may not be due to a conscious sense of self-efficacy or view of task competence, but rather to changes within the ANS, or in the way in which internal ANS signals are used to guide behavior.

Which aspects of the training experience in Experiment 1 were responsible for infants' success with the difficult 2:3 ratio? It might be that merely experiencing many different numerical ratios sharpens infants' discrimination abilities. In previous experiments using the change detection paradigm, infants saw just a single ratio throughout the testing session (Libertus & Brannon, 2010). In studies using the habituation paradigm, infants also typically experience many exemplars of a single numerosity, and then are tested on their ability to discriminate this from a novel numerosity (e.g., Cordes & Brannon, 2009; Lipton & Spelke, 2003; Xu & Arriaga, 2007; Xu & Spelke, 2000; Xu et al., 2005; Zosh, Halberda, and Feigenson, 2011). In contrast, each infant in Experiment 1 saw a total of six different ratios (including 2:3), raising the possibility that simply seeing a diversity of numerical ratios improved discrimination performance. Alternatively, the particular ordering of ratios infants experienced—from easiest to hardest—may have been critical. Odic et al. (2014) and Wang et al. (2016) found that children showed better ANS precision when they made ANS discriminations in the Easy-First ratio order, compared to a reversed Hard-First order, and compared to a randomized order, even though the number of different ratios experienced and the difficulty of these ratios were identical across conditions. We tested these possibilities in Experiment 2 by presenting a separate group of infants with the same ANS trials as in Experiment 1, but using a Hard-First trial order.

3. EXPERIMENT 2

3.1. Method

3.1.1. Participants

Thirty-two full-term 5- to 7-month-old infants participated (15 girls; mean age: 6 months 11 days (*SD* = 17 days)). Of these, 22 were identified by their parent(s) as White, five as Black, one as Asian, one as Hispanic, and two were identified as being "another race/ ethnicity." The race and ethnicity of one infant were not identified. Two additional infants participated but were excluded from analysis due to fussiness (n = 1) or equipment failure (n = 1).

3.1.2. Design and procedure

The design, stimuli, and procedure were as in Experiment 1, except that the order of the training trials was reversed: trials started with the hardest numerical ratios (3:4) and monotonically progressed to easier ones (1:6). As in Experiment 1, after the 10 training trials infants saw two unreinforced test trials involving a 2:3 ratio (14 vs. 21 dots), one with the Numerically Changing stream on the left and the other with the Numerically Changing stream on the right (order counterbalanced across infants).

3.2. Results

We first examined infants' numerical change preference scores in the 2:3 ratio test trials. Infants' average preference score did not differ from zero, $M = 1.25\%$, $SD = 29.15\%$, $t(31) = 0.24$, $p = .81$, *Cohen's* $d = 0.043$ (Fig. 3).

Next we asked whether infants' numerical change preference scores were influenced by stimulus type, size of the Numerically Constant numerosity, or spatial congruity between training and test trials. A 2 (Stimuli: black circles vs. rainbow squares) x 2 (Size of Numerically Constant Numerosity: larger vs. smaller) x 2 (Spatial Congruity: same vs. different side) ANOVA on infants' preference scores during the test trials revealed no effect of Stimuli, $F(1,24) = 0.003$, $p = .95$, $\eta_p^2 < .001$, no effect of Size of Numerically Constant Numerosity, $F(1,24) = 0.82$, $p = .37$, $\eta_p^2 = .033$, no effect of Spatial Congruity, $F(1,24) = 0.024$, $p = .88$, $\eta_p^2 = .001$, and no significant interactions among these, Fs < 1, ps $> .34$. Only 17 out of 32 infants looked longer at the Numerically Changing stream (binomial exact test $p = .86$).

To further ask whether the order in which infants experienced numerical ratios affected their discrimination abilities, we compared the test trial performance of infants in Experiment 1 (Easy-First training) and Experiment 2 (Hard-First training). This revealed a significant main effect of Experiment on infants numerical change preference scores, $t(62) = 2.27$, $p = .027$, *Cohen's* $d = 0.57$, with infants significantly preferring the Numerically Changing Stream in Experiment 1 but not in Experiment 2.

Infants' overall length of looking during the training trials suggests that this performance difference did not stem from differences in attentiveness leading up to the test trials. A 5 (Training Trial Pair) x 2 (Experiment: Experiment 1 vs. Experiment 2) repeated measures ANOVA on infants' average looking (combined across the Numerically Changing and Numerically Constant streams) during the training trials revealed a main effect of Training Trial Pair, $F(4,248) = 36.35$, $p < .001$, $\eta_p^2 = .37$, with infants in both experiments decreasing their visual interest as the training progressed, and no effect of Experiment, $F(1,62) = 1.35$, $p = .25$, $\eta_p^2 = .021$, nor any Training Trial Pair x Experiment interaction, $F(4,248) = 0.65$, $p = .58$, $\eta_p^2 = .10$ (all repeated-measures analyses Greenhouse-Geisser corrected for non-sphericity).

The difference in infants' test performance also was not due to differences in looking preferences during training, nor to the total amount of training feedback received. A 5 (Training Ratio) x 2 (Experiment: Experiment 1 vs. Experiment 2) ANOVA on infants' numerical change preference scores during training revealed no effect of Training Ratio, $F(4,248) = 1.38$, $p = .24$, $\eta_p^2 = .022$, no effect of Experiment, $F(1,62) = 0.15$, $p = .70$, $\eta_p^2 = .002$, and no Training Ratio x Experiment interaction, $F(4,248) = 1.13$, $p = .34$, $\eta_p^2 = .018$. Infants in Experiment 1 heard an average of 5.98 s of musical feedback per training trial ($SD = 4.93$), and infants in Experiment 2 heard an average of 5.54 s ($SD = 4.84$), $t(62) = 1.10$, $p = .28$, Cohen's $d = 0.28$. These results indicate that infants did not exhibit significantly better training performance in Experiment 1 than Experiment 2 (although Figs. 2 and 4 suggest that infants in Experiment 1 did discriminate more training ratios than infants in Experiment 2). Nonetheless, infants in Experiment 1 appeared to successfully use their training experience to succeed in the unreinforced test trials, whereas infants in Experiment 2 did not (see General Discussion).

Finally, infants' surprising success at discriminating numerosities differing by a 2:3 ratio in Experiment 1 is unlikely to be due to age effects. Although infants in our experiments were slightly older on average (6 months, 14 days) than infants in previous studies who failed to discriminate 2:3 ratios (e.g., 6 months, 0 days in Xu & Spelke, 2000; 6 months, 2 days in Lipton & Spelke, 2003; 6 months, 1 day in Libertus & Brannon, 2010), only infants in Experiment 1 succeeded, and infants in our two experiments did not

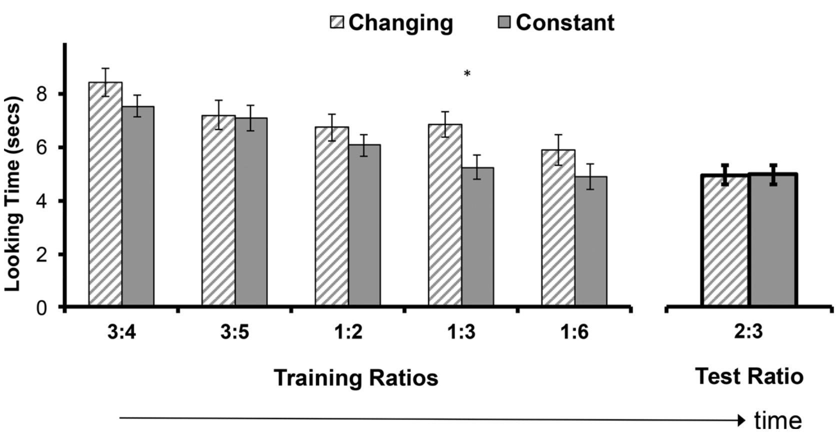

Fig. 4. Average looking at the Numerically Changing and Numerically Constant streams during the training trials and the 2:3 ratio test trials in Experiment 2. Error bars represent the standard error of the mean. *$p < .05$.

significantly differ in age (average age of infants in Experiment 1: 6 months, 14 days; average age of infants in Experiment 2: 6 months, 11 days), $t(62) = 0.05$, $p = .48$, Cohen's $d = 0.18$. Furthermore, we found no correlation between the age of infants in Experiment 1 and their numerical change preference score ($r = -0.18$, $p = .32$).

Our finding that preverbal infants exhibit ANS hysteresis suggests that self-perception of mathematical competence is not required to drive the effect. But ANS hysteresis might still require a certain degree of task engagement—in particular, it might rely on participants receiving feedback on their task-relevant behavior. The older children in the studies by Odic et al. (2014) and Wang et al. (2016) always made explicit numerical judgments (verbally identifying or pointing to an array), and immediately received an auditory signal after every trial telling them whether they were correct. We deliberately maintained a similar aspect of feedback in our infant task, rewarding infants with music when they looked at the Numerically Changing stream throughout the training portion of the study (but not during the critical test trials). Was this reinforcement a contributor to the ANS hysteresis effect in infants?

We addressed this in Experiment 3. A separate group of infants was tested under the same conditions as infants in Experiment 1. They saw a Numerically Constant stream on one monitor and a Numerically Changing stream on the other, in which the flashed arrays gradually progressed from easily discriminable ratios to hard-to-discriminate ratios. As in Experiment 1, infants in Experiment 3 also heard music during the training. However, this music was no longer tightly temporally linked to infants' looking at the Numerically Changing stream. For half the infants we inserted a short delay between their looking at the Changing stream and the music onset, as well as between their looking away from the Changing stream and the music offset. For the other infants we played music continuously throughout the training phase. These manipulations allowed us to ask, using two different ways of decoupling infants' behavior from the feedback they received, whether contingent feedback is required for infants to show a hysteresis-induced enhancement of ANS precision.

4. EXPERIMENT 3

4.1. Method

4.1.1. Participants

Thirty-two full-term 5- to 7-month-old infants participated (15 girls; mean age: 6 months 9 days (SD = 20 days)). Of these, 24 were identified by their parent(s) as White, three as Black, three as Asian, and one as Hispanic. The race and ethnicity of one infant were not identified. No infants were excluded from the data analysis.

4.1.2. Design and stimuli

The design and stimuli were identical to Experiment 1, with infants experiencing an Easy-First numerical training sequence followed by two 2:3 ratio test trials, this time using only rainbow-colored squares. We reduced the contingency of the feedback in two ways. Whereas in Experiment 1 music played as soon as infants looked at the Numerically Changing stream, and stopped playing as soon as infants looked away from that screen, in Experiment 3 this contingency was either weakened or removed altogether. In the Delayed Feedback condition (16 infants) we inserted a 500-ms delay before both the onset and offset of the feedback. As such, sometimes there was no music playing while infants were looking at the Numerically Changing stream, and sometimes there was music playing while infants were looking at the Numerically Constant stream or were distracted (i.e., not looking at either screen). In the No Feedback condition (16 infants) music was played throughout the training trials, regardless of where infants were looking.

4.2. Results and discussion

We first analyzed infants' average numerical change preference score in the 2:3 ratio test trials. Infants' preference scores did not differ from zero, $M = 3.56\%$, $SD = 29.06\%$, $t(31) = 0.69$, $p = .49$, *Cohen's* $d = 0.12$ (Fig. 3).

Next we asked whether infants' preference scores were influenced by feedback type, size of the numerically constant numerosity, or spatial congruity between training and test trials. A 2 (Feedback: delayed vs. no feedback) x 2 (Size of the Numerically Constant Numerosity: larger vs. smaller) x 2 (Spatial Congruity: same vs. different side) ANOVA on infants' preference scores in the test trials revealed no effect of Feedback, $F(1,24) = 0.23$, $p = .64$, $\eta_p^2 = .010$, no effect of Size of the Numerically Constant Numerosity, $F(1,24) = 0.86$, $p = .39$, $\eta_p^2 = .035$, and no effect of Spatial Congruity, $F(1,24) = 2.93$, $p = .10$, $\eta_p^2 = .11$. There was a marginally significant Feedback x Spatial Congruity interaction, $F(1,24) = 3.80$, $p = .063$, $\eta_p^2 = .14$, and no other significant interactions, Fs < 2.4, ps > .13. Only 17 out of 32 infants looked longer at the Numerically Changing stream (binomial exact test $p = .86$).

To more closely examine the effect of feedback, we compared infants' numerical preference scores in the 2:3 ratio test trials of Experiment 1 (in which they received Easy-First training with contingent feedback) and Experiment 3 (in which they received Easy-First training with delayed or no feedback). This revealed a marginally significant main effect of Experiment, $t(62) = 1.95$, $p = .056$. Infants preferred the Numerically Changing Stream in Experiment 1, but not in Experiment 3 (Fig. 3).

To further understand how feedback influences infants' numerical discrimination, we compared infants' numerical preference scores in the training trials of Experiments 1 and 3. A 5 (Training Trial Pair) x 2 (Experiment: Experiment 1 vs. Experiment 3) ANOVA on infants' numerical preference scores during training found no effect of Training Trial Pair, $F(4,248) = 1.95$, $p = .12$, $\eta_p^2 = .031$, and no effect of Experiment, $F(1,62) = 0.017$, $p = .90$, $\eta_p^2 < .001$, but did reveal a significant Training Trial Pair x Experiment interaction, $F(4,248) = 4.93$, $p = .002$, $\eta_p^2 = .074$. Infants in Experiment 1 (with contingent feedback) showed increasing preference scores as training progressed, whereas infants in Experiment 3 (without contingent feedback) showed the opposite pattern (Figs. 2 and 5), decreasing their numerical preference over time. This raises the possibility that the presence of contingent feedback early in

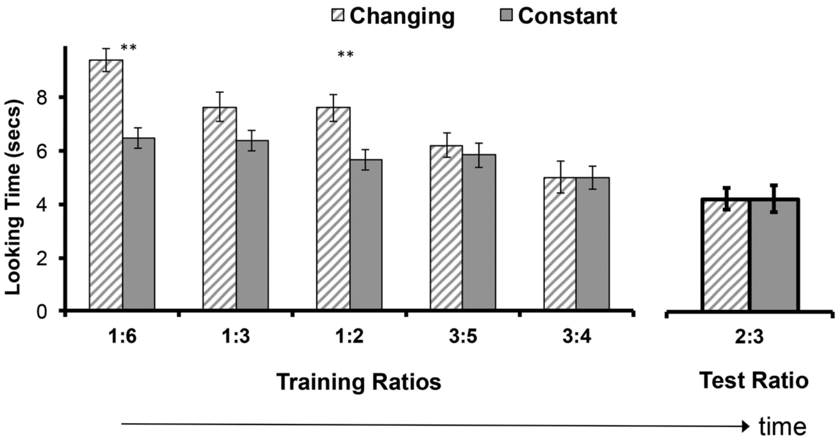

Fig. 5. Average looking at the Numerically Changing and Numerically Constant streams during the training trials and the 2:3 ratio test trials in Experiment 3. Error bars represent the standard error of the mean. $^{**}p < .01$.

the training trials of Experiment 1 disrupted infants' spontaneous preference for the Numerically Changing stream when it contained an extremely discriminable ratio (1:6). However, as the trials progressed, infants in Experiment 1 (but not Experiment 3) appeared to gradually discover the link between the contingent feedback and the numerical changes, and used this to help direct their looking to the numerically changing stimulus in the subsequent test trials.

5. General discussion

In the present study we asked whether infants' approximate number precision is influenced by their prior history with numerical stimuli. Several previous studies found that 6-month-old infants fail to discriminate quantities that numerically differ by a 2:3 ratio (Feigenson, 2011; Libertus & Brannon, 2010; Lipton & Spelke, 2003; Xu & Spelke, 2000; Xu et al., 2005), implicating relatively coarse ANS precision at this age. Here we show that when 6-month-old infants first experienced a series of numerical quantities that initially differed by large ratios, but that gradually came to differ by much smaller, hard-to-discriminate ratios, they succeeded at discriminating this challenging 2:3 ratio (Experiment 1). Because infants failed to discriminate the 2:3 ratio when the training sequence was reversed (Experiment 2), the scaffolded order in which numerical ratios were experienced, rather than the mere prior exposure to the numerical quantities, appeared to underlie infants' surprising success. Contingent feedback was required for this hysteresis-induced precision enhancement. When infants saw the scaffolded Easy-First training order, but experienced feedback that was decoupled from their looking behavior, they showed no precision enhancement (Experiment 3).

Previous research shows that extended training in non-symbolic number comparison over a period of days or weeks can lead to improved numerical performance in children and adults (DeWind & Brannon, 2012; Libertus et al., 2018). In addition to this targeted ANS training, long-term experience with formal mathematics education also seems to improve ANS precision (Nys et al., 2013; Piazza, Pica, Izard, Spelke, & Dehaene, 2013; but see Castronovo & Göbel, 2012 and Sullivan, Frank, & Barner, 2016 for contradictory findings). Our findings, and those of Odic et al. (2014) and Wang et al. (2016), show that ANS precision also can be modulated by a very brief dose of experience: just a few minutes of training resulted in a measurable enhancement. It is unknown how long this enhancement lasts, although our speculation is that it is short-lived. If so, a major contribution of ANS hysteresis may be in helping us understand the nature of a foundational system for thinking about quantity, rather than in any practical benefit (e.g., in the classroom).

The mechanism underlying ANS hysteresis remains ripe for future study. The finding that infants and children show better numerical discrimination when they begin with easier trials and progress to hard ones is reminiscent of the idea of scaffolding in the education literature, whereby learners are given instructional support at the start of a learning task, and this support is gradually removed. However, scaffolding does not have a single agreed-upon definition (Pea, 2004), and has been used to refer to what are likely distinct processes. For example, scaffolding sometimes involves grading the amount of social support a learner receives, and often involves tailoring the learning sequence to the trajectory of the individual learner. Neither of these applies to the case of ANS hysteresis studied here.

Our results do suggest that ANS hysteresis probably does not rely on a well-developed sense of self-efficacy, since it is unlikely that infants have access to such meta-cognitive awareness (Flavell, 1999; Sodian, Thoermer, Kristen, & Perst, 2012). However, another possibility is that ANS hysteresis in part reflects children's understanding of the task. Beginning the task with very easy numerical discriminations and gradually progressing to harder ones may have helped infants come to link the musical feedback with their behavior—i.e., with looking at changes in number, as opposed to looking at changes in density, or cumulative area, or spatial configuration. This increased attention to the dimension of numerosity could have reduced infants' uncertainty about which aspect of the arrays to respond to during the critical 2:3 test trials, thereby boosting their performance (for a related proposal with different predictions, see Cantrell, Boyer, Cordes, & Smith, 2015). In contrast, infants who started the task with quite hard-to-discern numerical differences could have experienced less certainty about which aspect of the stimulus was relevant (i.e., which aspect determined reinforcement). These infants may have sometimes attended to non-numerical aspects of the test arrays, thereby leading to poorer group discrimination performance. Note that in previous experiments that contained only a single numerical contrast, infants also failed to discriminate a 2:3 numerical ratio (e.g., Feigenson, 2011; Libertus & Brannon, 2010; Lipton & Spelke, 2003; Xu & Spelke, 2000; Xu et al., 2005). This suggests that it was the gradual progression from easy numerical discriminations to hard ones that helped infants focus on numerosity, rather than the progression from hard discriminations to easy ones distracting them from numerosity.

Another possibility proposed by Odic et al. (2014) is that ANS hysteresis reflects changes in children's internal confidence in their numerical decisions—not self-confidence in the task or the domain (e.g., "I am good at math"), but rather a sense of certainty when perceiving one of two arrays as more numerous. This view is in some ways similar to the interpretation given to observations of hysteresis in neural activity among participants who passively viewed different stimulus orderings, without performing any particular task (Kleinschmidt, Büchel, Hutton, Friston, & Frackowiak, 2002). Such results have been interpreted in terms of top-down signals, generated by recent previous experience, influencing subsequent lower levels of visual processing. In our task, infants' successful perception of distinct numerosities in the early trials of Experiment 1 may have biased their later processing of less-discriminable ratios, for example by adjusting the decision criterion used to direct looking (i.e., for two distinct distributions to be represented by the Approximate Number System, some bimodality must be present in the activations; how much bimodality is required in order for an observer to perceive two numerosities rather than one is a matter of setting an internal criterion, which may shift over time). A more precise decision criterion would result in less variability in infants' perception of distinct numerosities. However, unlike the perceptual hysteresis studied by Kleinschmidt et al. (2002), which required no response at all on the part of participants, the ANS hysteresis we observed was feedback-dependent. It may be that perceptual hysteresis and ANS hysteresis both involve shifts of unconscious decision criteria, but differ in their reliance on task engagement. It may also be that younger children (e.g. the infants in the present experiment, and the 5-year olds tested by Odic et al., 2014 and Wang et al., 2016) required some kind of task, and feedback on their performance in this task, to maintain sufficient attention to exhibit the effect. Adults could plausibly be less sensitive to such external factors.

Our results, combined with those of previous investigations (Baker et al., 2014; Cantrell et al., 2015; Jordan et al., 2008; Odic et al., 2014; Wang et al., 2016) suggest a reframing of the way in which ANS precision is often discussed. This precision is not fixed at a given age, nor for a given observer, even though on average ANS precision does increase throughout childhood (Halberda & Feigenson, 2008) and people vary in their ANS precision (Halberda et al., 2008). Instead, ANS precision fluctuates within an individual according to a variety of factors. These include recent history making approximate number discriminations (as in the current experiments), and the availability of redundant input (Baker et al., 2014; Jordan et al., 2008). Other factors, too, could plausibly affect ANS precision and should be tested, including the salience or perceived value of the stimuli, and the observer's recent history with symbolic number computations. Important to note when considering these possibilities is that "ANS precision" here is broadly construed. It could refer to either the width of the distributions within the Approximate Number System, or to the criterion used to individuate these distributions. Either would result in a change in behavioral performance in a numerical discrimination task, and it is this performance that we find is sensitive to recent numerical experience, from early in life.

In sum, our results provide evidence for a temporary modulation of infants' approximate number precision. Prior experience with easy numerical discriminations, followed by a gradual increase in difficulty, as well as contingent feedback on numerical responses led to improved number discrimination. Our findings add to the evidence suggesting that ANS precision is more dynamic than previously thought.

Acknowledgements

This research was conducted at Johns Hopkins University and was funded by National Institutes of Health grant R01 HD057258. The authors declare no conflict of interest that might influence the research. The authors gratefully thank all the participants of this study. We thank Kathleen Cracknell, Eva Lason, Arthi Rameshkumar, Janay Smith, and Samantha Tuepker for helping with data collection.

References

Baker, J. M., Mahamane, S. P., & Jordan, K. E. (2014). Multiple visual quantitative cues enhance discrimination of dynamic stimuli during infancy. *Journal of Experimental Child Psychology, 122*, 21–32.

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition, 83*(3), 223–240.

Castronovo, J., & Göbel, S. M. (2012). Impact of high mathematics education on the number sense. *PLoS One, 7*(4), e33832.

Cantrell, L., Boyer, T. W., Cordes, S., & Smith, L. B. (2015). Signal clarity: An account of the variability in infant quantity discrimination tasks. *Developmental Science, 18*(6), 877–893.

Cordes, S., & Brannon, E. M. (2009). Crossing the divide: Infants discriminate small from large numerosities. *Developmental Psychology, 45*(6), 1583.

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

DeWind, N. K., & Brannon, E. M. (2012). Malleability of the approximate number system: Effects of feedback and training. *Frontiers in Human Neuroscience, 6*, 68.

Feigenson, L. (2011). Predicting sights from sounds: 6-month-olds' intermodal numerical abilities. *Journal of Experimental Child Psychology, 110*(3), 347–361.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307–314.

Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology, 50*(1), 21–45.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the" number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology, 44*(5), 1457.

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, 109*(28), 11116–11120.

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

Jordan, K. E., Suanda, S. H., & Brannon, E. M. (2008). Intersensory redundancy accelerates preverbal numerical competence. *Cognition, 108*(1), 210–221.

Kleinschmidt, A., Büchel, C., Hutton, C., Friston, K. J., & Frackowiak, S. J. (2002). The neural structures expressing perceptual hysteresis in visual letter recognition. *Neuron, 34*(4), 659–666.

Libertus, K. (2008). *Preferential looking coder.* Retrieved August 5th, 2008, from . . http://www.duke.edu/~kl41.

Libertus, M. E., & Brannon, E. M. (2010). Stable individual differences in number discrimination in infancy. *Developmental Science, 13*(6), 900–906.

Libertus, M. E., Odic, D., Feigenson, L. & Halberda, J. (submitted). Training visual approximate number sense improves auditory number sense and school math ability.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense large-number discrimination in human infants. *Psychological Science, 14*(5), 396–401.

Nys, J., Ventura, P., Fernandes, T., Querido, L., Leybaert, J., & Content, A. (2013). Does math education modify the approximate number system? A comparison of schooled and unschooled adults. *Trends in Neuroscience and Education, 2*(1), 13–22.

Odic, D., Hock, H., & Halberda, J. (2014). Hysteresis affects approximate number discrimination in young children. *Journal of Experimental Psychology: General, 143*(1), 255.

Odic, D., Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Developmental change in the acuity of approximate number and area representations. *Developmental Psychology, 49*(6), 1103.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences, 13*(3), 423–451.

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.

Ross-sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development, 74*(6), 1807–1822.

Sodian, B., Thoermer, C., Kristen, S., & Perst, H. (2012). Metacognition in infants and young children. *Foundations of Metacognition,* 119–133.

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), 18116–188120.

Suanda, S. H., Tompson, W., & Brannon, E. M. (2008). Changes in the ability to detect ordinal numerical relationships between 9 and 11 months of age. *Infancy, 13*(4), 308–337.

Sullivan, J., Frank, M. C., & Barner, D. (2016). Intensive math training does not affect approximate number acuity: Evidence from a three-year longitudinal curriculum intervention. *Journal of Numerical Cognition, 2*(2), 57–76.

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

Xu, F., & Arriaga, R. I. (2007). Number discrimination in 10-month-old infants. *British Journal of Developmental Psychology, 25*(1), 103–108.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*(1), B1–B11.

Xu, F., Spelke, E. W., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8*(1), 88–101.

Zosh, J. M., Halberda, J., & Feigenson, L. (2011). Memory for multiple visual ensembles in infancy. *Journal of Experimental Psychology: General, 140*(2), 141.