Testing the Emergentist Theory of Number Perception Development

Miranda N. Long (mlong@psych.ubc.ca)
Department of Psychology at the University of British Columbia, 2136 West Mall,
Vancouver, BC V6T 1Z4 Canada

Darko Odic (darko.odic@psych.ubc.ca)
Department of Psychology at the University of British Columbia, 2136 West Mall,
Vancouver, BC V6T 1Z4 Canada

Abstract

The origin of abstract concepts remains a central question in cognitive science. Empiricists argue that abstract concepts can be learned through a small set of domain-general cognitive mechanisms, whereas nativists argue that most abstract concepts are not learned at all. To explore these opposing views, we examine the abstract concept of number. Drawing on a recent "emergentist" proposal derived from machine learning models, we tested whether stimuli that align with experiences critical for an empiricist view of number perception result in more accurate and less variable numeric estimations. Five- to nine-year-olds were tested on stimuli that conformed to the natural statistics of objects (i.e., number was correlated with features such as cumulative area and spatial position, and the distribution of stimuli was power-law based) or those that did not comply with these statistics. We find little-to-no evidence for predictions of the emergentist theory when presenting children with natural stimuli.

Keywords: Number; nativism; empiricism; computational modelling.

Introduction

An incredible feat of the human mind is its ability to reason and represent concepts abstractly. How exactly we acquire such concepts is under considerable debate (for review, see Laurence & Margolis, 2024). Empiricists argue that abstract concepts are learned from a small set of domain-general cognitive mechanisms while nativists believe they aren't learned but emerge via special purpose learning mechanisms or maturation. We investigated one abstract concept that has compelling evidence on both sides: the concept of number.

A nativist perspective of number perception claims that number is represented via a domain-specific module, the Approximate Number System (ANS; Dehaene, 1997; Feigenson et al., 2004). To support this claim, evidence from comparative psychology (Brannon and Terrace, 1998; Howard et al., 2019; Rugani et al., 2015), cross-cultural studies (Dehaene et al., 2008; Ferrigno et al., 2017; Frank et al., 2008; Piazza et al., 2013), and developmental work with newborns and infants (Izard et al., 2009; Xu and Spelke, 2000) all show a sensitivity to changes in cardinality with little-to-no experience with formal arithmetic, number words, etc. Developmental improvements in number perception are seen as maturational. Additionally, research disentangling number from non-numeric variables (e.g., cumulative area, density, etc.; DeWind et al., 2015; Starr et al., 2017) and work highlighting specific neurons and brain regions that correspond with changes in cardinality (Piazza et al., 2004; Nieder, 2005; Nieder & Miller, 2004) demonstrate that number is a specific psychological representation and not part of a set of domain-general sense of quantity.

The empiricist perspective claims that number emerges via domain-general learning mechanisms and experience. For example, the "the emergentist perspective" (Testolin et al., 2020; Zorzi & Testolin, 2018) of number perception argues that number is a byproduct of how the brain learns to *compress* rich sensory input. This can be demonstrated in simple neural networks that are trained – without supervision – to re-create their own inputs using layers smaller than the input. Testolin et. al (2020) show that a neural network that learns to re-create objects in natural scenes *spontaneously* develops number-specific neurons with developmental improvements in this network closely mirror those of children. This opens the possibility that humans similarly learn about number through passive experience with real-world scenes (see also S. Y. Chen et al., 2018; Creatore et al., 2021; Kim et al., 2021; Stoianov & Zorzi, 2012; Testolin, Dolfi, et al., 2020; Zorzi & Testolin, 2018).

The emergentist perspective has been very successful in accounting for the existing developmental data on number perception (see Zorzi & Testolin, 2018 for a review). However, a stronger test would be to derive novel predictions about how number perception should function given emergentist models and then test these in human observers.

Odic and Oppenheimer (2022) did exactly this. A key prediction of emergentist models is that number perception should be superior when viewing stimuli that resemble natural scenes on which number perception was theoretically trained on. That is, the closer the test stimuli are to the training stimuli, the better number perception should be. Contrary to this, however, they find that adult observers are no better – and are sometimes worse – at estimating number in real-world photographs compared to artificial stimuli.

While offering a preliminary test of the emergentist account, these results can still be integrated within the emergentist theory in two ways. First, an emergentist could argue that adults have gained sufficient experience with all kinds of visual stimuli, so that the preference for natural properties may have vanished; children, therefore, offer a more robust test. Second, while Odic and Oppenheimer (2022) tested one type of natural statistic – the correlations between number and non-numeric features such as size, the spatial distribution of objects in real-world scenes, etc. – this

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

is not the only relevant statistic. An even more prominent feature of natural scenes is that number is distributed in a strongly right-skewed power-law distribution (i.e., there are substantially more natural scenes with less than 10 objects compared to more than 10 objects; Piantadosi, 2016; Testolin et al., 2020). Therefore, perhaps a sensitivity to natural statistics can be detected in some natural features, like stimuli adhering to a power-law distribution, but not others.

A recent study (Sanford & Halberda, 2024) extended some of the above work to 5 – 9-year-old children, showing that children also do no better on number perception tasks using real-world photographs vs. using images from children's counting books (which often *reverse* the relationship between number and the size of objects). However, this study did not examine any developmental effects, instead comparing children as a large group on real-world vs. counting-book images. Additionally, their work did not test any sensitivity to the distributional information of number in natural scenes.

In the current study, we tested whether 5 – 9-year-olds perform better on a number perception task when testing stimuli: (1) mirror features of natural scenes (e.g., the correlation between number and cumulative, sense of depth, and spatial location of objects, etc.) and (2) have a power-law distribution of the number of objects over the course of the experiment (see Figure 1). To do this, all children were shown displays used to train and test emergentist neural network models (Testolin et al., 2020) and asked how many objects they saw. Importantly, some displays were based on natural scene images where the natural correlations between number and cumulative area, a sense of depth, and the spatial location of objects were all maintained but other features such as color or object identity were removed ("Natural Stimuli"). Other displays were artificial, in that objects had no correlation with cumulative area and were placed in a random position ("Artificial Stimuli"). Additionally, half of the participants saw stimuli where the number of objects presented followed a right skewed distribution (i.e., "Power-law Distribution"), similar to natural scenes, whereas the other half of participants completed the experiment where the number of objects were presented equally as many times (i.e., "Uniform Distribution"). The task was structured around two Distribution between-subject conditions (Uniform vs. Power-Law) and two Stimuli within-subject conditions (Natural vs. Artificial), controlling for order of Stimuli.

If humans develop number perception like the emergentist account suggests, then children should (1) be most accurate and less variable in their estimations when both features (i.e., Natural Stimuli + Power-law Distribution) are present in the stimuli, (2) be least accurate and more variable when these features are both absent (i.e., Artificial Stimuli + Uniform Distribution), and (3) perform somewhere in the middle if only one of the natural stimuli features were present. We did not have a prediction as to whether just a Power-law Distribution of numbers (paired with Artificial Stimuli) or just the Natural Stimuli (paired with Uniform Distribution) would produce a stronger effect, though – because we control for order – we can detect any such differences. Additionally, given that the emergentist account suggests that number perception is learn via experience, we predicted that older children (i.e., those with more experience with the features of natural scenes) would show stronger effects (i.e., better accuracy, lower variability) for conditions where the features of natural scenes were present.

Method

Participants

One-hundred-and-sixteen children (63 girls, 53 boys) were recruited from the Vancouver area. Children were between the ages of 5- and 9-years-old (M = 6;11 [years; months], range = 5;0 – 8;11).

Individual responses for each participant were excluded from analysis if they were (1) zero or above 64 (i.e., double the maximum number shown) and (2) 3 standard deviations above or below the participant's mean estimate. Additionally, participants were excluded and replaced if they (1) did not complete the task (n=17), (2) the removal of outlier responses removed more than 10% of their responses (n=14), (3) participants' responses produced a linear slope that was not significantly higher than 0 at $p < .05$ (n=12), and (4) participants' estimated slopes were 3 SD higher or lower than the rest of their answers (n = 1). These exclusion rates are consistent with previous work on number estimation in young children who recently acquired number words and such criteria are used to flag children that are not paying attention (Dramkin & Odic, 2024; Libertus et al., 2020).

Materials

All participants saw two stimuli sets that deviated in whether stimuli were based on natural scenes or were artificial. The images in the Natural stimuli set were created by Testolin et al. (2020) from the datasets used for the PASCAL detection challenge. Original images contained pictures of natural scenes (e.g., cows in a field) with rectangular bounding boxes around objects. Testolin et al. (2020) replaced these bounding boxes with non-overlapping white rectangles on a black background, maintaining natural correlations between number and cumulative area, $r(75{,}072) = .23$, $p < .001$, a sense of depth, and the spatial location of objects while removing other features of the scene that were not the focus of the current study. Images in the Artificial stimuli set (referred to as the 'S&Z dataset' by Testolin et al., 2020) contained white rectangles where object position was random and the correlation between cumulative area and number was controlled, $r(41{,}598) = .00$, $p = .972$, both of which are not characteristic of natural scenes. A subset of the stimuli used in Testolin et al. (2020) was randomly sampled and used in the current study. This random selection process maintained a marginal correlation between number and cumulative area in the Natural stimuli set, $r(39) = .30$, $p = .059$, and a lack of correlation between number and cumulative are in the Artificial stimuli set, $r(39) = .06$, $p < .714$. Like Testolin et al.'s (2020) study with neural networks, we restricted our experiment to numbers up to 32. Additionally, numbers 1 to

6 were excluded to eliminate the possibility of subitizing. Figure 1 shows examples of the stimuli used.

While each participant saw one block of the Natural Stimuli images and one block Artificial Stimuli images (order was counterbalanced), each participant was randomly assigned to one of two conditions based on the distribution of numbers presented in the experiment: Power-law (i.e., based on natural scenes) and Uniform (i.e., more artificial). In the Power-law distribution condition, each number of items from 7 to 32 was approximately represented following a power-law function (See Figure 1). Specifically, there were nine trials with 7 objects, six trials with 9, five trials with 11, four trials with 13, three trials with 16, two trials with 19, one trial for 23, 27, and 32. In the Uniform distribution condition, the number of objects was uniformly distributed: there were four trials with 7, 9, 13, 16, 19, 23, 27, and 32 objects each. In total, children saw 32 trials per stimuli category, totaling 64 trials total in the experiment.

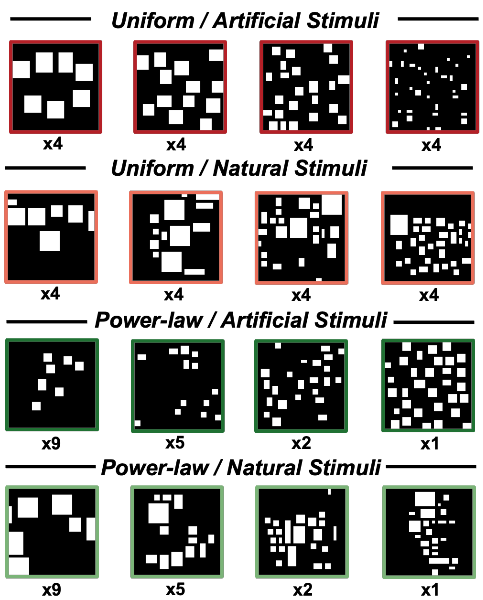

Figure 1: Examples of stimuli. Children either saw stimuli following a natural (i.e., power-law) or artificial (i.e., uniform) distributions of numbers with one block containing images of natural stimuli and one block containing images of artificial stimuli. Blocks were counterbalanced.

Procedure

Participants were tested in-person in a quiet room on an iMac running PsychoPy. Children were first randomly assigned to either the Power-law or Uniform Distribution condition. No participant completed both Distribution conditions. To keep child participants engaged and maintain their attention over the entire experiment, an experimenter introduced the experiment as a game where children were tasked with helping a construction worker build new buildings by estimating the number of bricks he was using. The task included two blocks of 32 trials for a total of 64 trials over the entire experiment. One block included images with Natural Stimuli whereas another block contained images of Artificial Stimuli. For each trial, stimuli were presented for 500 ms which is fast enough to prevent counting. Participants verbally reported how many "bricks" they saw as the experimenter typed out the response. No feedback was given to prevent calibration and anchoring.

Results

Individual differences in number estimation tasks are characterized by two signatures: estimation accuracy (i.e., β) and estimation variability (i.e., σ). To derive estimate accuracy and variability, we used the PsiMLE method (Odic et al., 2016). When interpreting the first parameter, β (i.e., estimate accuracy), the further β is from 1.0, the worse the participant is at matching their estimates to the target number. Values below 1.0 demonstrates underestimation whereas values above 1.0 demonstrate overestimation. A typical β value in the domain of number is around 0.80 for adults (Krueger, 1984; Odic et al., 2016) but tends to be lower in young children (Libertus et al., 2020). The second parameter, σ, models estimation variability: when σ is closer to 0.0 then participants' responses are less variable whereas when σ is higher, then participants' responses are more variable. A typical σ value in number estimation is around 0.20 for adults (Cordes et al., 2001) but tends to be higher in young children (Libertus et al., 2020). If the emergence of the ANS is in line with the empiricist account, then we should expect participants in the Natural Distribution and Natural Stimuli conditions to produce β values closer to 1.0 compared to all other conditions and produce the lowest σ values (see Odic & Oppenheimer, 2022 for simulations).

We first examined the raw estimates across target numbers. Figure 2A shows the average response across all target numbers with the dotted line denoting perfect performance in the task. As can be seen in Figure 2A, the average estimate increases with target number for all conditions, highlighting that participants were able to extract numeric information from displays.

Table 1: Estimation accuracy and variability.

	Accuracy (β)		Variability (σ)	
	M	*SD*	*M*	*SD*
Uniform + Artificial	0.65	0.31	0.31	0.11
Uniform + Natural	0.63	0.37	0.31	0.12
Power-law + Artificial	0.71	0.38	0.25	0.09
Power-law + Natural	0.64	0.39	0.27	0.11

To quantify if a Power-law Distribution and Natural Stimuli produced more accurate numerical estimations, we examined estimation accuracy (indexed as β; see Table 2) across all four conditions. A 4 (Age Group: 5-, 6-, 7-, 8-

years-olds) by 2 (Distribution: Power-Law, Uniform) by 2 (Stimuli: Natural, Artificial) mixed effects ANOVA with β as the dependent variable revealed no main effect of Age Group, $F(3, 214) = 0.36$, $p = .785$, $\eta^2_p = 0.005$, no main effect of Distribution, $F(1, 214) = 0.64$, $p = .424$, $\eta^2_p = 0.002$, no main effect of Stimuli, $F(1, 214) = 0.00$, $p = .989$, $\eta^2_p = 0.000009$, no interaction between Age Group and Distribution, $F(3, 214) = 0.71$, $p = .545$, $\eta^2_p = 0.01$, no interaction between Age Group and Stimuli, $F(3, 214) = 0.52$, $p = .670$, $\eta^2_p = 0.001$, no interaction between Distribution and Stimuli, $F(1, 214) = 0.31$, $p = .580$, $\eta^2_p = 0.01$, and no interaction between Age Group, Distribution, and Stimuli, $F(3, 214) = 0.25$, $p = .861$, $\eta^2_p = 0.003$. Results suggest no advantage for the Power-law Distribution or the Natural Stimuli conditions, contrary to the emergentist account of number perception (Figure 2B). Additionally, these results held across all age groups. To make sure our exclusion criteria did not disproportionately remove lower-performing children from our sample, we conducted those same analyses with all children but found no differences.

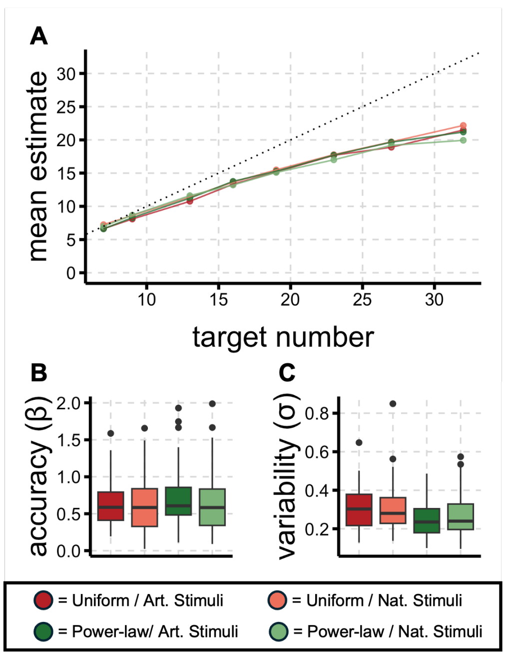

Figure 2: (A) Mean estimates across numbers. Dotted line denotes 100% accuracy. (B) Mean accuracy (β) across conditions. (C) Mean variability (σ) across conditions.

Because our original analyses treated age as a discrete variable, there is the potential that we were not capturing more nuanced developmental differences. To address this, we treated age as a continuous variable and examine whether continuous age and accuracy (β) are related across all four conditions. For evidence in line with the empiricist account of number perception, we would expect a more positive relationship between continuous age and accuracy (β) for the conditions that correspond with features of real-world scenes (i.e., Power-law Distribution, Natural Stimuli). We find no relationship between continuous age and accuracy (β) for the Power-law Distribution/Natural Stimuli condition, $r(55) = 0.02$, $p = .864$, no relationship for the Power-law Distribution/Artificial Stimuli condition, $r(55) = 0.09$, $p = .520$, no relationship for the Uniform Distribution/Natural Stimuli condition, $r(57) = 0.08$, $p = .524$, and no relationship for the Uniform Distribution/Artificial Stimuli condition, $r(57) = 0.06$, $p = .629$ (Figure 3).

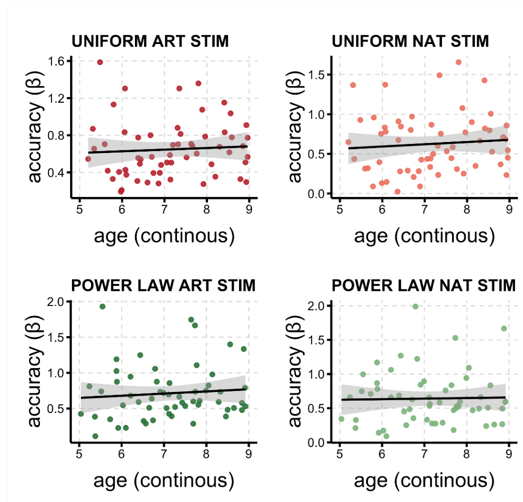

Figure 3: Correlation between accuracy (β) and continuous age for all four conditions.

Different cognitive processes can yield equally accurate solutions but may differ in how long it takes to reach such solutions. To test whether the time it took to complete each trial differed between Natural and Artificial stimuli, we ran a 2 (Distribution: Power-Law, Uniform) by 2 (Stimuli: Natural, Artificial) mixed effects ANOVA with reaction time as the dependent variable. Results revealed no main effect of Distribution, $F(1, 226) = 0.81$, $p = .369$, $\eta^2_p = 00357$, no main effect of Stimuli, $F(1, 226) = 0.55$, $p = .458$, $\eta^2_p = 0.00244$, no interaction between Distribution and Stimuli, $F(3, 226) = 0.004$, $p = .953$, $\eta^2_p = 0.00002$, suggesting no differences in process timing among all the conditions tested.

One alternative explanation for the lack of accuracy differences (β) is that the novelty of the artificial stimuli recruited extra attentional resources. To test this, we compared estimation accuracy on the first and last trial of the Artificial Stimuli condition. If novelty had enhanced accuracy, we'd expect the first trial to show higher accuracy in the Artificial Stimuli condition. However, a paired samples t-test revealed the opposite: accuracy was higher on the last trial ($M = 0.80$, $SD = 0.37$) than the first trial ($M = 0.70$, $SD = 0.28$), $t(113) = -2.53$, $p = .01$.

Next, we examined estimation variability (indexed as σ; see Table 2) across all conditions. A 4 (Age Group: 5-, 6-, 7-, 8-years-olds) by 2 (Distribution: Power-law, Uniform) by 2 (Stimuli: Natural, Artificial) mixed effects ANOVA with σ as the dependent variable revealed a main effect of Age Group, $F(3, 214) = 10.65$, $p < .001$, $\eta^2_p = 0.13$, a main effect of Distribution, $F(1, 214) = 11.59$, $p = .001$, $\eta^2_p = 0.05$, no main effect of Stimuli, $F(1, 214) = 0.22$, $p = .643$, $\eta^2_p = 0.001$, no interaction between Age Group and Distribution, $F(3, 214) = 1.62$, $p = .185$, $\eta^2_p = 0.02$, no interaction between Age Group and Stimuli, $F(3, 214) = 0.71$, $p = .547$, $\eta^2_p = 0.01$, no interaction between Distribution and Stimuli, $F(1, 214) = 0.81$, $p = .371$, $\eta^2_p = 0.004$, and no interaction between Age Group, Distribution, and Stimuli, $F(3, 214) = 0.52$, $p = .667$, $\eta^2_p = 0.007$. The main effect of Age Group is not surprising given that previous research demonstrates that children get less variable and more precise in the numeric estimation over development (Halberda & Feigenson, 2008) and can be equivalently explained by the nativist account (i.e., via maturation) and empiricists (i.e., via experience). The main effect of Distribution suggests an advantage for stimuli that exhibit one of the features of real-world scenes (i.e., power-law distribution of objects), consistent with an emergentist account of number perception. However, this effect should also increase with age given that older children have more experience, yet we do not observe an interaction between Age Group and Distribution. To make sure our exclusion criteria did not disproportionately remove lower-performing children from our sample, we conducted those same analyses with all children and found that there was no longer a main effect of Distribution. All other results were replicated.

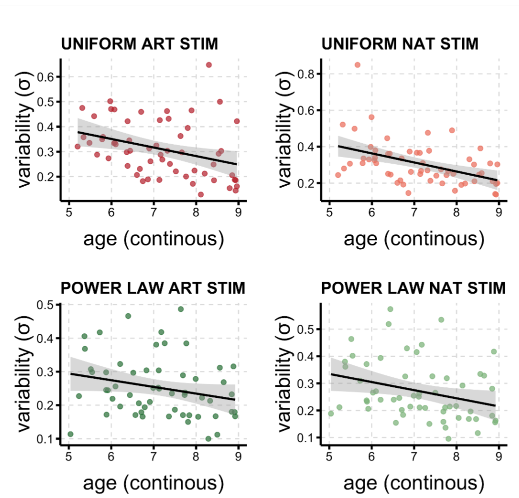

Figure 4: Correlation between variability (σ) and continuous age for all four conditions.

We treated age as a continuous variable and examine whether continuous age and variability (σ) are related across all four conditions. For evidence in line with the empiricist account of number perception, we would expect a more negative relationship between continuous age and accuracy (σ) for the conditions that correspond with features of real-world scenes (i.e., power-law distribution, natural stimuli). We find a negative relationship between continuous age and variability (σ) for the Power-law Distribution/Natural Stimuli condition, $r(55) = -0.30$, $p = .026$, a marginal relationship for the Power-law Distribution/Artificial Stimuli condition, $r(55) = -0.24$, $p = .070$, a negative relationship for the Uniform Distribution/Natural Stimuli condition, $r(57) = -0.46$, $p < .001$, and a negative relationship for the Uniform Distribution/Artificial Stimuli condition, $r(57) = -0.34$, $p = .008$ (Figure 4). Given that the relationships appear to be higher in the Uniform Distribution conditions, this can be interpreted as evidence that children are more rapidly developing number estimation variability for Uniform compared to Power-law distributions.

Because children always saw both the natural and artificial stimuli in two separate testing blocks, we tested if the order in which participants saw these stimuli (i.e., natural following artificial, artificial following natural) contributed to any of our findings. A 4 (Age Group: 5-, 6-, 7-, 8-years-olds) by 2 (Order: Artificial/Natural, Natural/Artificial) by 2 (Stimuli: Natural, Artificial) mixed effects ANOVA with β as the dependent variable revealed no main effect of Age Group, $F(3, 214) = 0.35$, $p = .787$, $\eta^2_p = 0.005$, no main effect of Order, $F(1, 214) = 0.58$, $p = .446$, $\eta^2_p = 0.002$, no main effect of Stimuli, $F(1, 214) = 0.00$, $p = .988$, $\eta^2_p = 0.00001$, an interaction between Age Group and Order, $F(3, 214) = 7.66$, $p < .001$, $\eta^2_p = 0.10$, no interaction between Age Group and Stimuli, $F(3, 214) = 0.43$, $p = .732$, $\eta^2_p = 0.006$, no interaction between Order and Stimuli, $F(1, 214) = 3.11$, $p = .080$, $\eta^2_p = 0.01$, and no interaction between Age Group, Order, and Stimuli, $F(3, 214) = 0.11$, $p = .956$, $\eta^2_p = 0.01$. As seen in Table 3, this interaction between Age Group and Order was carried by the fact that young children who began with the Natural Stimuli condition (independent of distribution) showed lower β values (i.e., they were *worse* than when starting with the Natural Stimuli condition).

Table 3: β values across age and order of stimuli

	Natural/Artificial		Artificial/Natural	
	M	*SD*	*M*	*SD*
5	0.43	0.25	0.90	0.43
6	0.66	0.40	0.56	0.32
7	0.75	0.40	0.62	0.32
8	0.66	0.25	0.70	0.37

A 4 (Age Group: 5-, 6-, 7-, 8-years-olds) by 2 (Order: Artificial/Natural, Natural/Artificial) by 2 (Stimuli: Natural, Artificial) mixed effects ANOVA with σ as the dependent variable revealed a main effect of Age Group, $F(3, 214) = 9.86$, $p < .001$, $\eta^2_p = 0.12$, no main effect of Order, $F(1, 214) = 1.20$, $p = .275$, $\eta^2_p = 0.006$, no main effect of Stimuli, $F(1, 214) = 0.21$, $p = .651$, $\eta^2_p = 0.001$, no interaction between Age Group and Order, $F(3, 214) = 0.54$, $p = .653$, $\eta^2_p = 0.008$, no interaction between Age Group and Stimuli, $F(3, 214) = 0.80$, $p = .494$, $\eta^2_p = 0.01$, no interaction between Order and Stimuli, $F(1, 214) = 2.62$, $p = .107$, $\eta^2_p = 0.01$, and no interaction between Age Group, Order, and Stimuli, $F(3, 214) = 0.70$, $p = .554$, $\eta^2_p = 0.01$. Results suggest that order did not influence the variability of responses.

Discussion

The current study further examined the origin of abstract concepts by testing whether a relatively recent empiricist theory (i.e., the emergentist account) can be validated in children. Using number as a case study for abstract concepts, we tested the prediction that children would perform better on a number perception task (i.e., more accurate and less variable responses) when the testing stimuli (1) were derived from real-world images that contained natural correlations between number and cumulative area, a sense of depth, and the spatial location of objects, and (2) followed a power-law distribution of objects, another common feature of natural scenes. We see no advantage for accuracy when participants saw either the natural stimuli or power-law distribution and when participants saw both features of natural scenes. We find a small but significant effect for the Power-law Distribution condition in variability of responses, with variability being lower in the Power-law Distribution condition compared to the Uniform Distribution condition. We find no developmental effects, except that the youngest children show lower (i.e., worse) estimation if the first block they completed was the Natural Stimuli condition.

Our results are contrary to the emergentist account. Replicating previous work in adults (Odic & Oppenheimer, 2023) and children (Sanford and Halberda, 2024), we find no advantage in number perception for stimuli that matched features of real-world scenes. Unlike previous work, we show that there is also no developmental change in this effect. One might argue that children no longer learn much about numbers from passive experience (e.g., they've had sufficient exposure), but fully trained "adult" models still show an advantage in both estimation accuracy and variability with natural stimuli (see Odic & Oppenheimer, 2023) *and* the developmental period between 5 and 9 years is rife with ongoing development of the ANS (Halberda & Feigenson, 2008; Odic, 2016). Furthermore, younger children did *worse* on when they began with stimuli that matched features of real-world scenes, contrary to what would be expected if number sense emerges from scenes such as these.

While we found no effects of stimuli type (i.e., Natural vs. Artificial) on performance, we did find an effect for the distribution of target numbers across the experiment. The Power-law Distribution showed an overall better (i.e., lower) variability, consistent with emergentist models. However, we find no interactions with age. This effect – by itself – is not predicted by model stimulations in Odic & Oppenheimer (2022), which show strong correlations in network performance on accuracy and variability. Nevertheless, if the pattern is reliable, this does suggest that children's number perception is more sensitive to the *distributional* features of real-world scenes. For example, children may learn to expect that most daily experiences contain fewer than 10 objects and are therefore more precise when stimuli conform to that expectation. Convergent work has shown that 5-year-olds possess a low-number prior – an expectation that an array of objects should contain a small number (Long & Odic, 2025). This is consistent with work on the emergence of visual priors (Girshick et al., 2011; Stocker & Simoncelli, 2006), some are believed to be a product of experience with natural scenes (e.g., the "light-from-above" prior; Adams et al., 2004). Future emergentist models must be revised to produce the pattern of separating variability from accuracy performance to fully account for this developmental trend and account for the lack of developmental effects for power-law distributions.

Although the effect for the power-law distribution on variability aligns with the empiricist account of number perception, this effect alone can also theoretically be explained by the nativist account. Rather than children utilizing this distributional information of natural scenes to learn the concept of number over a lifetime of experiences, humans may be tuned to environmental pressures and develop innate priors about this distributional information that coincide with a specialized system for number (i.e., the ANS). This may explain why small cardinalities are represented with substantially higher precision than larger ones, even in infants (Piantadosi, 2016). The lack of an age effect in our study further supports this nativist claim compared to the explanations of the empiricist account. Future research should examine younger children to determine if developmental differences expected from the empiricist account of number emerge earlier, or if evidence suggests this distributional information is an innate prior.

Why might real-world distributions benefit number perception whereas the stimuli based on natural scenes do not? One possibility is that information like the correlations between number and cumulative area, sense of depth, spatial location of object, etc. may create an additional challenge in isolating number from other cues in everyday scenes (DeWind et al., 2015). A system that is specialized for number perception (e.g., through evolutionary pressures) would potentially do better by de-confounding number from other features. Instead, cues like the distributional information of natural scenes may instead be more relevant, unique, and less conflicting than other types of information for learning the concept of number.

Although our work showcases how novel predictions can be derived from computational models to test new ideas and theories, one limitation of the current work is that we did not test computational models with the same intermixed research design we used with children. Existing models have been trained on either Natural and Power-Law Distributions or Artificial and Uniform Distributions, but not on the intermediate conditions. Perhaps some of these variants could account for the differences observed here.

Our results can apply only to the ANS and numbers larger than 4, as we only tested numbers outside of the subitization range. It's unclear whether correlations between number and cumulative area, a sense of depth, and the spatial location of objects benefit smaller numbers and if this system relies on distributional information from natural scenes, as suggested by the two-number systems theory (Feigenson et al., 2004).

Where abstract concepts originate remains a key issue in cognitive science, with differing views from empiricists and nativists. In examining the abstract concept of number, we find results that are contrary to the empiricist account. These findings contribute to the ongoing debate of how abstract concepts may be learned and represented.

References

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the "light-from-above" prior. *Nature Neuroscience, 7*, 1057–1058.

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

Chen, S. Y., & Fang, M. (n.d.). Can Generic Neural Networks Estimate Numerosity Like Humans?

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

Creatore, C., Sabathiel, S., & Solstad, T. (2021). Learning exact enumeration and approximate estimation in deep neural network models. *Cognition, 215*, 104815.

Dehaene, S. (1997). The Number Sense. New York: OxfordUniversity Press.

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or Linear? Distinct Intuitions of the Number Scale in Western and Amazonian Indigene Cultures. *Science, 320*, 1217–1220.

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.

Dramkin, D., & Odic, D. (2024). Children dynamically update and extend the interface between number words and perceptual magnitudes. *Developmental Science, 27*, e13433.

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

Ferrigno, S., Jara-Ettinger, J., Piantadosi, S. T., & Cantlon, J. F. (2017). Universal and uniquely human factors in spontaneous number perception. *Nature Communications, 8*, 13968.

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., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review. *Acta Psychologica, 171*, 17–35.

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience, 14*, 926–932.

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.

Howard, S. R., Avarguès-Weber, A., Garcia, J. E., Greentree, A. D., & Dyer, A. G. (2019). Numerical cognition in honeybees enables addition and subtraction. *Science Advances, 5*, eaav0961.

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

Kim, G., Jang, J., Baek, S., Song, M., & Paik, S.-B. (2021). Visual number sense in untrained deep neural networks. *Science Advances, 7*, eabd6127.

Krueger, L. E. (1984). Perceived numerosity: A comparison of magnitude production, magnitude estimation, and discrimination judgments. *Perception & Psychophysics, 35*, 536–542.

Laurence, S., & Margolis, E. (2024). The Building Blocks of Thought: A Rationalist Account of the Origins of Concepts (1st ed.). Oxford University Press Oxford.

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

Libertus, M. E., Odic, D., Feigenson, L., & Halberda, J. (2020). Effects of Visual Training of Approximate Number Sense on Auditory Number Sense and School Math Ability. *Frontiers in Psychology, 11*, 2085.

Long, M. N., & Odic, D. (2025). Evidence for a Low Number Prior in Children's Intuitive Number Sense. *Child Development*.

Nieder, A. (2005). Counting on neurons: The neurobiology of numerical competence. *Nature Reviews Neuroscience, 6*, 177–190.

Nieder, A., & Miller, E. K. (2004). A parieto-frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences, 101*, 7457–7462.

Odic, D., Im, H. Y., Eisinger, R., Ly, R., & Halberda, J. (2016). PsiMLE: A maximum-likelihood estimation approach to estimating psychophysical scaling and variability more reliably, efficiently, and flexibly. *Behavior Research Methods, 48*, 445–462.

Odic, D., & Oppenheimer, D. M. (2023). Visual numerosity perception shows no advantage in real-world scenes compared to artificial displays. *Cognition, 230*, 105291.

Piantadosi, S. T. (2016). A rational analysis of the approximate number system. *Psychonomic Bulletin & Review, 23*, 877–886.

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.

Rugani, R., Vallortigara, G., Priftis, K., & Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' mental number line. *Science, 347*, 534–536.

Sanford, E. M., & Halberda, J. (2024). Non-numerical features fail to predict numerical performance in real-world stimuli. *Cognitive Development, 69*, 101415.

Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience, 9*, 578–585.

Starr, A., DeWind, N. K., & Brannon, E. M. (2017). The contributions of numerical acuity and non-numerical stimulus features to the development of the number sense and symbolic math achievement. *Cognition, 168*, 222–233.

Stoianov, I., & Zorzi, M. (2012). Emergence of a "visual number sense" in hierarchical generative models. Nature Neuroscience, 15(2), 194–196.

Testolin, A., Dolfi, S., Rochus, M., & Zorzi, M. (2020a). Visual sense of number vs. Sense of magnitude in humans and machines. *Scientific Reports, 10*, 10045.

Testolin, A., Dolfi, S., Rochus, M., & Zorzi, M. (2020b). Visual sense of number vs. Sense of magnitude in humans and machines. *Scientific Reports, 10*, 10045.

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

Zorzi, M., & Testolin, A. (2018). An emergentist perspective on the origin of number sense. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*, 20170043.