Looking for more food or more people? Task context influences basic numerosity perception

Michele Fornaciai[1,*], Abigail Farrell[1], Joonkoo Park[1,2]

[1] Department of Psychological and Brain Sciences, University of Massachusetts, Amherst, MA, USA.
[2] Commonwealth Honors College, University of Massachusetts, Amherst, MA, USA.

* Corresponding author.
Michele Fornaciai
(mfornaciai@umass.edu)
University of Massachusetts
Department of Psychological and Brain Sciences
135 Hicks Way, Amherst, 01003
USA

ABSTRACT

Approximate numerical magnitude (or numerosity) is thought to represent one of the fundamental sensory properties driving perceptual choices. Recent studies indicate that numerosity judgment on a dot array is primarily driven by its numerical magnitude, largely independent from its other non-numerical visual dimensions. Nevertheless, these findings do not preclude the possibility that non-numerical cues such as size or spacing of a dot array influence numerosity judgment. Here, we test the hypothesis that numerosity judgment is influenced by non-numerical dimensions of a dot array depending on the context to which those non-numerical cues could be useful. Participants were asked to choose the more numerous of two dot arrays in two different contexts that differed only in one aspect. In one condition, the task was framed as choosing a set with more fruits to consume. In the other condition, the task was framed as choosing a group with more people to join. The results demonstrate that the influence of non-numerical cues – and particularly of the dimension of size – was significantly smaller when participants made quantitative choices about people than when they made choices about food, illustrating that the representation of discrete magnitude is more pronounced in the former case. These findings suggest that the information pooled to reach a decision about numerosity is flexibly determined according to the context and the goals of such judgment.

KEYWORDS

Numerosity perception; numerical cognition; number sense; non-numerical cues.

INTRODUCTION

From an evolutionary point of view, the ability to rapidly estimate the approximate number of items in a set (a.k.a. numerosity) has an adaptive value. For instance, approximate numerical abilities would be advantageous for several activities spanning from foraging to social decisions and to fight-or-flight decisions determining survival. That said, modern theories of numerical cognition posit that our approximate numerical abilities have deep ontogenetic and phylogenetic roots (e.g. Gelman & Cordes, 2001; Dehaene, 2011). This idea is empirically supported by studies demonstrating that this *number sense* is widespread across animal species (Agrillo et al., 2008; Pepperberg, 2006; Piantadosi & Cantlon, 2017; Rugani et al., 2015) and present from birth in human newborns (Izard et al., 2009; Xu, 2003; Xu & Spelke, 2000).

One crucial question, however, is whether numerosity is processed by a dedicated perceptual system independently from other non-numerical magnitude dimensions, or whether and to what extent other continuous visual cues are used to derive the representation of numerosity. According to the aforementioned lay evolutionary story, there is no particular reason for the perceptual system to be specifically sensitive to one unique magnitude dimension (e.g., number, which is a discrete magnitude, as opposed to total mass, which is a continuous magnitude).

Along this line of reasoning, some authors raise the hypothesis that numerosity is processed via other visual attributes, like texture-density (e.g. Durgin, 2008) or some unknown combination of other continuous magnitude dimensions (e.g. Gebuis et al., 2016; Leibovich et al., 2017).

While this issue between the numerical versus non-numerical nature of the brain's magnitude system has been a polarizing topic in the past years, there is a growing amount of evidence supporting the idea that numerosity is a fundamental perceptual attribute, not reducible to combinations of other non-numerical cues (Anobile et al., 2016; Cicchini et al., 2016; Fornaciai et al., 2016; DeWind et al., 2015; Park et al., 2016; Park, 2017; Fornaciai & Park, 2017; Fornaciai et al., 2017). One critical contribution came from DeWind and colleagues (2015), who developed an innovative method to quantify the relative contributions of various magnitude dimensions to one's performance in a numerosity judgment task. Most prior work on numerosity judgment attempted to de-correlate numerical magnitude from other non-numerical magnitudes of a dot array, which is physically impossible. In contrast, DeWind and colleagues identified three orthogonal dimensions (*numerosity*, *size*, and *spacing*) that serve as a basis for most, if not all, magnitude dimensions of a dot array, and used a generalized linear model to quantify how each of these basic dimensions contribute to one's numerical judgment. They found that numerosity is the primary source of information driving one's performance in a numerosity discrimination task, with very little influence of size and spacing. More crucially, subsequent studies have now repeatedly demonstrated that brain responses (arising from early visual cortex) to dot-array stimuli even in passive viewing paradigms are strongly modulated by the numerical magnitude of the stimuli, with little contributions from other dimensions (Park et al., 2016; Park, 2017; Fornaciai & Park, 2017; Fornaciai et al., 2017). These results bolster the idea that discrete numerosity information gets extracted very early in the brain largely independent from non-numerical information of a visual scene.

These recent findings, however, do not preclude the possibility that non-numerical cues such as size or spacing of a dot array influence numerosity judgment. Moreover, it is easy to imagine a real-life situation where judgment based on non-numerical cues would be more advantageous. Consider the lay evolutionary story again. In the case of foraging for food, aggregate size of the items is perhaps more important for survival than merely the number of items, although in the case of making social decisions like joining a group of people, the number of people may be more important for survival than the aggregate body size of the people. Such reasoning leads to the hypothesis that numerosity perception is supported by a flexible mechanism exploiting different numerical and non-numerical dimensions according to the context and goals of the judgment.

To investigate how numerical and non-numerical information contributes to numerosity perception as a function of task context, we examined approximate numerical abilities in human observers by simulating more realistic tasks in order to contextualize numerical choices. We devised two conditions using nearly identical stimuli, but framed in different ways. In one condition, participants were instructed to perform a numerosity discrimination task choosing a set with more food items (i.e. more apples), while in the other condition participants had to choose a group with more people. In both cases, the stimuli were systematically constructed to span identical ranges of numerosity and non-numerical dimensions.

METHODS

Participants

Two hundred twenty-one subjects took part in the study (154 females, mean age = 20.2 ± 1.5 years). All participants had normal or corrected-to-normal vision, provided written informed consent prior to participating in the study, and were compensated for their time with course credit. Experimental procedures were approved by the University of Massachusetts Institutional Review Board, and were in line with the declaration of Helsinki.

Apparatus and stimuli

The study was conducted in a large computer lab, with groups of 1 to 8 participants (most typically around 3 or 4 participants) tested in parallel during each session, although each participant completed the study individually. Stimuli were dot arrays constructed using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for Matlab (R2013b, The Mathworks, Inc.), and presented on a monitor screen encompassing 37 x 30 degrees of visual angle from a distance of about 57 cm (resolution = 1280 x 1024, frame rate = 60 Hz).

Dot arrays comprised orange dots enclosed in a black outline, with two small lines (length scaled as function of dot size, 10-14 pixel) added to characterize the dots according to the specific task context (see Fig. 1B). Namely, in one case ("food" condition) the lines were arranged to resemble the stem of an apple, while in the other condition ("people" condition) the two lines were arranged to resemble two eyes. Such simple features used to differentiate the stimuli in the two conditions were chosen to keep low level information (contrast, edges) as similar as possible.

Stimuli were constructed, following the design previously used by DeWind et al. (2015), to span similar ranges across three orthogonal dimensions: *numerosity*, *size*, and *spacing*. Numerosity comprised 5 levels, evenly spaced in a $\log_2$ scale: 12, 14, 17, 20, 24 dots. The non-numerical dimension of size is derived by combining the log-scaled values of the area of the individual items and the overall area occupied by the items. The non-numerical dimension of spacing is derived by combining the log-scaled values of the area of the invisible circular field in which the items are drawn (field area) and the sparsity of the items (the inverse of item density). More specifically, the dimension of size (Sz) refers to the dimension along which both the total area of the items (TA) and their individual area (IA) change at the same rate, while numerosity (N) is kept constant: $\log(Sz) = \log(TA) + \log(IA)$. The dimension of spacing (Sp) refers to the dimension along which both field area of the stimuli (FA) and the sparsity ($Spar$) of the items change concurrently, while numerosity is held constant: $\log(Sp) = \log(FA) + \log(Spar)$.

Furthermore, based on the dimensions of size and spacing, two additional attributes can be defined: apparent closeness (AC) and coverage (Cov). Apparent closeness (AC) represents the overall scaling of the dots independently from numerosity – i.e., an increase in apparent closeness is equivalent to increasing both size and spacing at the same rate, and is defined as $\log(AC) = \frac{1}{2}\log(Sz) + \frac{1}{2}\log(Sp)$. Coverage ($Cov$) represents the total area of the items (TA) divided by the field area of the stimuli (FA), and is defined as $\log(Cov) = \frac{1}{2}\log(Sz) - \frac{1}{2}\log(Sp)$.

One important characteristic of this peculiar design is that all the non-numerical dimensions of such stimuli (i.e., *IA*, *TA*, *FA*, *Spar*, *AC*, *Cov*) can be defined as a linear combination of the three orthogonal dimensions of numerosity, size, and spacing. For more details about this design, see also DeWind et al. (2015) and Park et al. (2016). Across the experiment, different numerosities were tested an equal amount of times, while the levels of the other dimensions were randomly chosen (independently for each of the stimuli) in each trial.

Regarding the specific values of the different attributes, the smallest individual area of the dots (*IA*) was set to ~1018 pixel^2, corresponding to a diameter of 1.04 deg (36 pixel), while the largest individual area was ~2290 pixel^2, corresponding to a diameter of 1.55 deg (54 pixel). On the other hand, the smallest field area (*FA*) was set to 101,787 pixel^2, encompassing a diameter of 10.4 deg (360 pixel), while the largest *FA* was 152,053 pixel2, corresponding to a diameter of 12.7 deg (440 pixel).

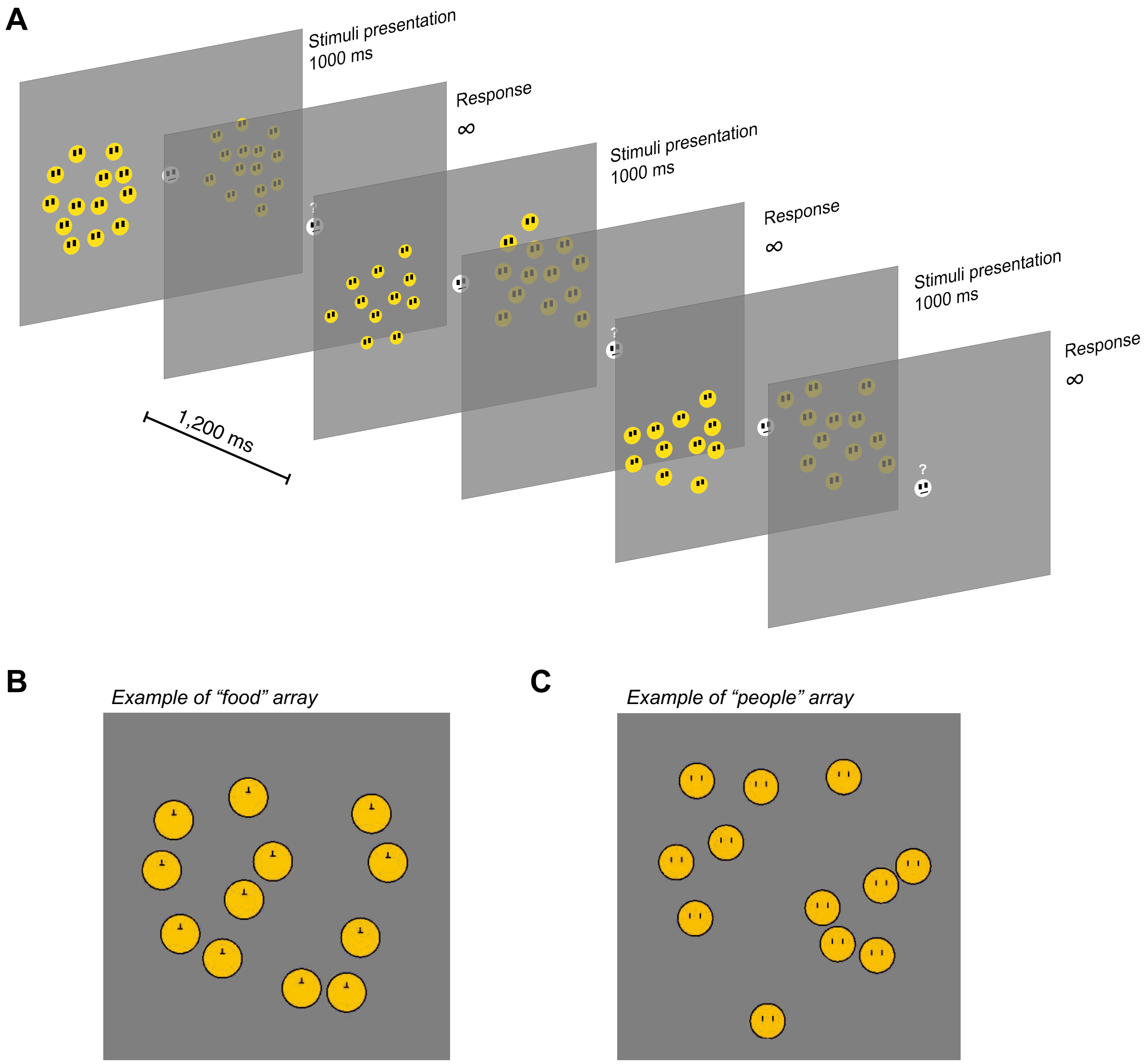

FIGURE 1. Procedure and stimuli. (A) Example of the experimental procedure. Participants completed two conditions, where different contexts were provided. With the exception of the context and small details of the stimuli displayed (see panel B), the basic task was identical for both conditions. On each trial, two arrays, one on each side of the screen, were presented for 1000 ms, and participants were asked to choose the more numerous stimulus. After providing a response, the next trial started after 1200 ms. Note that the stimuli are not depicted in scale (B) An exemplary array of "apples" in the "food" condition. (C) An exemplary array of "people" in the "people" condition. Note that the parameters (numerosity, size, and spacing) of the two sample stimuli were randomly drawn from the set of parameters used in the actual experiment,

and in this case the numerosity happened to be identical between the images in panels B and C but not the size and spacing of the arrays.

Procedure

Each participant completed two different conditions, each comprising 5 blocks of 70 trials. At the beginning of the experiment, a fictional character ("Jack") was introduced and shown on the screen as a smiley face throughout all the instruction screens and during the experiment. Participants were told that the experiment will involve helping the character to solve some problems. In both conditions, participants performed a two-alternative forced-choice numerosity discrimination task, where the subject had to choose one of two stimuli presented on the right and the left part of the screen (horizontal eccentricity = 8.25 deg). To avoid confusion about the boundary of the two stimuli, the character was depicted at the center of the screen throughout the task. Each pair of stimuli was presented for 1000 ms, and participants were free to look at the stimuli during the presentation. Afterwards, a question mark appeared above the character, and participants were instructed to choose one of the two stimuli according to the specific task, by pressing the appropriate key on a standard keyboard.

In the "food" condition, the task was framed as a foraging expedition, and participants had to help the character choose the set with more apples. Specifically, the instructions were as follows: "Jack is very hungry and must go foraging for food in the forest. Jack finds two patches of apples on either side of his path. However, Jack will only be able to collect the apples from one of the two patches. In this condition, your task is to help Jack chose the patch with more apples." In the "people" condition, the task required participants to help the character choose the group with more people. This task was framed as if the character had to escape a predator, and needed to join the larger group to be safer. Specifically, the instructions were as follows: "Jack is being pursued by a predator and must join another group of his kind for protection. Jack finds two groups of which he could join either. However, the two groups are not traveling together and Jack can only join one. In this condition, your task is to help Jack choose the group with more members." Besides these instructions, provided at the beginning of each task, the stimuli were differentiated only by two lines, arranged to resemble two eyes or the stem of an apple (Fig. 1B). Once a response was provided, the next trial started after 1200 ms.

Within each participant, the two conditions were performed in a random order. Each block took approximately 5 minutes and participants were instructed to rest their eyes between blocks if they wished. The entire procedure took approximately 50 minutes to complete (see Fig. 1A for a depiction of the experimental procedure).

Data analysis

In order to assess participants' overall performance in the tasks, we first computed participants' accuracy and precision in the numerosity discrimination task, separately for each of the two conditions (food or people). To achieve these measures, subjects' responses as a function of the difference in numerosity between the two stimuli presented in each trial were fitted with a cumulative Gaussian function, following the maximum-likelihood method (Watson, 1979). The point of subjective equality (PSE), representing the difference in numerosity between the two stimuli yielding chance-level responses, was defined as the median of the best-fitting Gaussian curve to all the data of a given subject in a given condition. The just noticeable difference (JND), representing the minimum difference in the stimuli detectable by a subject, was defined as the standard deviation of the Gaussian function. Participants with insufficient levels of performance (i.e. JND > 6) were excluded from further analysis. This criterion led to the exclusion of 21 participants, leaving a total of 200 participants. Note that this relatively high number of excluded participants may be due to little close supervision from the experimenters, as we ran the experiment on relatively large groups of participants. Such lack of close supervision is likely to have been occasionally resulted in poorly motivated participants not performing the task as instructed (i.e. pressing keys at random).

In order to assess the contribution of numerical and non-numerical magnitude dimensions on behavioral responses, the data were analyzed by modeling responses as a function of different visual attributes of the stimuli presented on each trial (DeWind et al., 2015). This model was indeed specifically designed to take into account the role of other non-numerical continuous attributes, in order to include their influence on numerical judgments when modeling participants' performances (eq. 1; adapted from DeWind et al., 2015). A generalized linear model was fitted to the data, which included regressors for the dimensions of *numerosity*, *size*, and *spacing* (see *Apparatus and stimuli* for details about the construction of such dimensions).

$$p(ChooseRight) = (1-\gamma)\left(\frac{1}{2}\left(1+erf\left(\frac{\log_2(rnum)-(\frac{-\beta side-\beta size\log_2(rsize)-\beta spacing\log_2(rspacing)}{\beta num})}{\sqrt{2}\frac{1}{\beta num}}\right)\right)-\frac{1}{2}\right)+\frac{1}{2} \quad (1)$$

More specifically, the parameters of the model included the log ratios of *numerosity*, *size*, and *spacing* (i.e. the ratio of the values of the different dimensions of the two stimuli presented on each trial, r_{num}, r_{size}, $r_{spacing}$). The model then fitted the behavioral responses (expressed as *p(ChooseRight*), representing the probability of choosing the stimulus on the right as more numerous) to estimate the regressors (β_{num}, β_{size}, $\beta_{spacing}$) of the log-ratio parameters. Additionally, the parameter γ represents the guessing rate – i.e. the proportion of trials where participants may have provided random responses due to a lapse of attention. However, as our task was relatively slowly-paced, we set the guessing term to zero. For more details about the model and a comparison with other models, see DeWind et al. (2015).

Finally, we tested whether each of the three beta estimates (i.e., the contribution of number, size, and spacing on judgment) differed in the two conditions. Importantly, because the order of the two conditions were given randomly across participants, we reasoned that the beta estimates could be modulated by that order (e.g., whether one performs the "food" condition first or the "people" condition first). As such, we used a repeated measure ANOVA with task condition (food vs. people) as a within-subject variable and condition order (food was given first vs. people was given first) as a between-subject covariate, allowing a full factorial design. Inspired by this model, we also analyzed individual participant's PSE and JND as a function of both task condition and condition order, with which we begin the Results section.

RESULT

We first assessed participants' overall performance in terms of accuracy (PSE) and precision (JND). A repeated-measures ANOVA with task condition as a within-subject factor and condition order as a between-subject factor was run separately on PSE and JND. There were no significant effects of condition ($F(1,198) = 1.144$, $p = 0.286$), order ($F(1,198) = 0.56$, $p = 0.813$), and the interaction ($F(1,198) = 1.144$, $p = 0.286$) on PSE. Similarly, there were no significant effects of condition ($F(1,198) = 1.016$, $p = 0.315$), order ($F(1,198) = 1.466$, $p = 0.227$), and the interaction ($F(1,198) = 1.016$, $p = 0.315$) on JND. These results indicate that there was a negligible systematic difference in the general measures of performance across the two tasks.

The central goal of this study was to assess the extent to which numerical and non-numerical visual attributes contribute to perceptual performance in different task contexts. To answer this question, we quantified the contribution of numerosity, size, and spacing of a dot array to participant's perceptual judgments following DeWind et al. (2015) and compared those estimates of contribution between the two task contexts. Specifically, a repeated-measures ANOVA with task condition as a within-subject factor and condition order as a between-subject factor was run on each of the three beta estimates. Regarding the dimension of numerosity, we observed negligible effect of condition ($F(1,198) = 1.145$, $p = 0.286$), negligible effect of order ($F(1,198) = 0.011$, $p = 0.917$), and only a weak effect of the interaction between condition and order ($F(1,198) = 3.31$, $p = 0.070$). Similarly, the dimension of spacing did not show any effect of condition ($F(1,198) = 0.238$, $p = 0.626$), order ($F(1,198) = 0.853$, $p = 0.357$), or interaction ($F(1,198) < 0.01$, $p = 0.993$). More strikingly, however, the dimension of size revealed a statistically significant effect of condition ($F(1,198) = 17.84$, $p < 0.001$) and interaction ($F(1,198) = 34.88$, $p < 0.001$), although with no main effect of order ($F(1,198) = 0.011$, $p = 0.917$). The strong effect of interaction warrants posthoc comparisons. As shown in Figure 2, the effect of interaction was captured by a slight but nonsignificant difference in the beta estimates for size between the "food" and the "people" condition in participants who performed the "food" condition first (Fig. 2A-B; posthoc paired t-test, $t(101) = 1.27$, $p = 0.20$) and at the same time a more negative beta estimate for size in the "food" condition in participants who performed the "people" condition first (Fig. 2C-D; posthoc paired t-test, $t(97) = -6.71$, $p < 0.001$). One possible explanation for this pattern is that participants exhibit a carry-over effect where their implicit strategy in the second half of the experiment is influenced by the implicit strategy that they have developed throughout the first half of the experiment, but that this carry-over effect is asymmetric in the two conditions. We return to this point in the discussion.

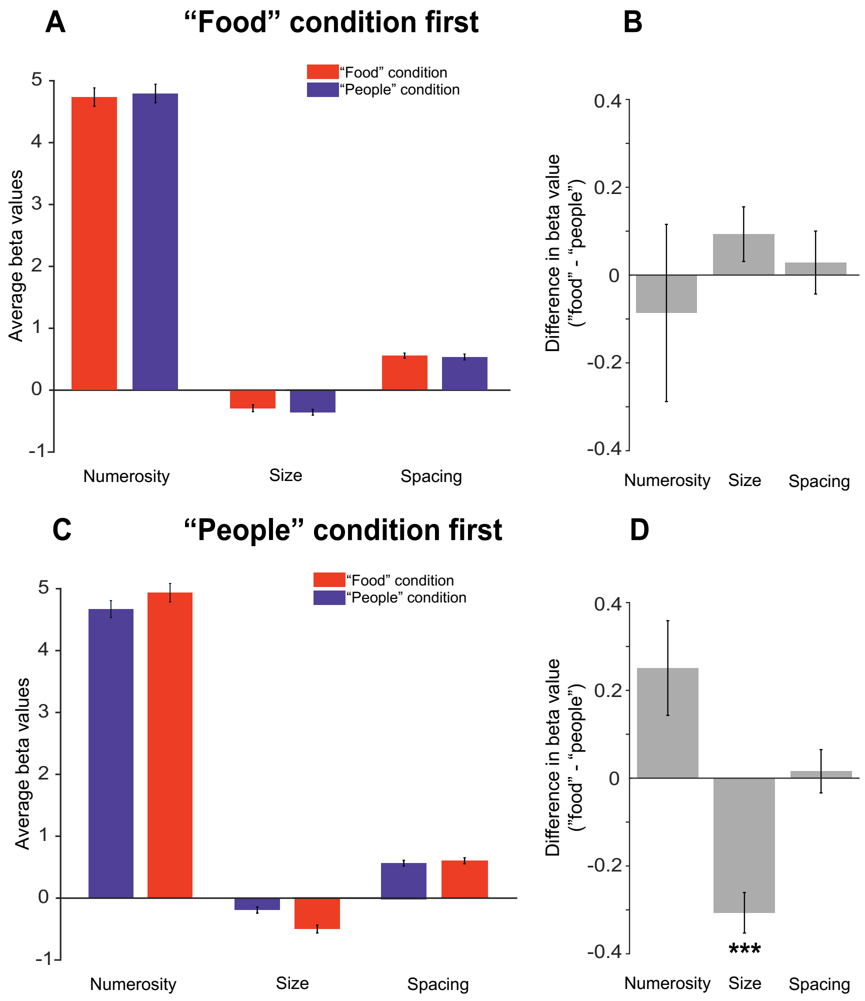

FIGURE 2. Comparison of the contributions of numerosity, size, spacing to the behavioral judgment across task conditions and order. (A) Beta estimates for the dimensions of numerosity, size, and spacing obtained with the generalized linear model, across the two conditions, in the cases where the “food” condition was performed first. (B) Differences in beta values between the two conditions, for each of the three orthogonal dimensions, for the “food” condition first case. (C) Beta estimates for numerosity, size, and spacing across the two conditions, in the cases where the “people” condition was performed first. (D) Differences in beta values between the two conditions when the “people” condition was performed first. Bars corresponding to different conditions in panels A and C are reported in the order in which they were performed. The difference in beta values reported in panels C and D represent beta values in the “food”

condition minus beta values in the "people" condition. Error bars represent SEM. *** $p < 0.001$.

As a potential carry-over effect makes it difficult to interpret the influence of the task context in the within-subject analysis, we assessed the between-subject effects of task context by considering exclusively the first task performed by the participants. Doing so, we found larger beta estimates (in the negative direction) in the food condition (-0.29 ± 0.05) than in the people condition (-0.17 ± 0.05) ($t(198) = -1.74$, two-tailed $p = 0.083$; Cohen's $d = 0.20$; see Fig. 2A and 2C), indicating a smaller bias from the dimension of size in those who looked for more people compared to those who looked for more food.

DISCUSSION

Humans, as well as many other animal species, are endowed with an intuitive sense of number that allows for a rapid and approximate estimation of numerical magnitude of a set of objects in a visual scene. A growing amount of evidence suggests that such number sense could be considered a basic perceptual ability, underpinned by a dedicated brain system (e.g. Burr & Ross, 2008; Park et al., 2016; Anobile et al., 2016; Nieder, 2016). Indeed, several recent studies have demonstrated that using multi-dimensional stimuli modulated along numerical and non-numerical dimensions, numerosity represents the most relevant information driving behavioral responses in explicit numerical tasks (DeWind et al., 2015) and even driving brain responses in passive-viewing paradigms (Park et al., 2016; Fornaciai & Park, 2017; Fornaciai et al., 2017). However, most studies investigating approximate, non-symbolic numerical abilities are usually performed in a laboratory setting (but see Piantadosi & Cantlon, 2017, for a work examining quantitative abilities in wild baboons), employing generic stimuli and tasks with little or no personal meaning. If our number sense has evolved as an adaptive strategy, a better understanding of the mechanism underlying numerosity perception would be achieved by considering a situation more relevant to a judgment in real life. We therefore aimed in this study to test how context (e.g., foraging or joining social group) may influence numerosity judgment.

In the present study, we achieved this aim by employing a basic numerical task (two-alternative forced-choice numerosity discrimination), but framed in two different ways. Participants were asked to help a fictional character solve some specific problems. In one case, the character had to find food, while in the other case the character had to choose a group of people to join. In both cases, however, the key task instructions were identical, as participants were instructed to choose the side with "more" items (either apples or people). This paradigm provides two advantages. First, despite the fact that it is still a laboratory experiment, the different task contexts more closely represent relevant tasks that have to be accomplished in the real environment. Second, employing the exact same task instructions but only varying the context in which the task is framed allows us to assess whether and to what extent the context itself affects numerosity discrimination performance.

Our results show that this is indeed the case: the specific context systematically affects the extent to which different magnitude dimensions are exploited to guide behavior, although an interaction between condition and the order in which the conditions are performed suggests asymmetries between the effect of different task contexts. More specifically, while on the one hand our results show that numerosity is the primary source of information driving numerosity discrimination judgments – in line with previous studies (e.g. DeWind et al., 2015) – the dimension of size contributes differently between the two task contexts. Furthermore, such differences are modulated by the task order. Namely, when participants perform the "food" condition first, there is little difference between the average beta estimates for the dimension of size in the two conditions. Conversely, when participants perform the "people" condition first, there is a sharp difference between the two conditions: the contribution of size is much closer to zero in the "people" condition, while it is much stronger in the "food" condition.

These different patterns of the effect of size could be explained by simultaneously considering (1) an asymmetric effect of task context, with one condition (i.e., the "food" condition) providing a stronger modulation compared to the other, and (2) a carry-over effect from the first condition performed in a session to the second one. Evidence for the first idea above comes from the between-subject analysis in which the "food" condition elicited a stronger bias in size than the "people" condition. Such differences in the strength of modulation then could result in asymmetric carry-over effect. That is, when those who performed the "people" condition first were then given the "food" condition, their numerical judgment was drastically biased by the size dimension. Conversely, when those who performed the "food" condition first were then given the "people" condition, the bias of the size dimension may have been already strong in the first part of the experiment and was carried over to the second part of the experiment. To understand the asymmetric carry-over effect in the effect of size, it is first worth asking whether it is the food condition that results in the implicit use of the size cue (in the more negative direction) or whether it is the people condition that results in the implicit use of the size cue (in the less negative direction). Previous studies employing the same modeling approach to basic numerical judgments (DeWind et al., 2015; Starr et al., 2017) give the clue to this question. Those previous studies demonstrated that the effect of size in numerosity judgment is very close to zero, more similar to the current results of the people condition (when performed first). That said, it is likely that the substantially more negative effect of size in the food condition is driven by an implicit use of the size cue, and that the less negative (closer to zero) effect of size in the people condition represents a more default, neutral use of the size cue. In addition to the effect of size, the effects of spacing across the present and previous studies were very similar. The effect of numerosity was much larger in the present result; however, that difference may easily be explained by the differences in the difficulty of the task where the present study was much easier with a substantially longer viewing time (i.e., 1000 ms vs. 250 ms in previous studies). This asymmetric carry-over effect can then be explained by considering the use of a specific implicit strategy in one context (i.e. food), as opposed to the use of a default, neutral, strategy in the other case (i.e. people). if a non-numerical cue (i.e. size) was implicitly used to aid judgments in the first condition (food), such an implicit strategy may remain in effect in the second condition. In contrast, a decision free from such an implicit strategy at the beginning of the session (i.e. the people task in the people first condition) could be altered with a different task that puts demand on developing such a strategy. In other words, an implicit strategy once employed, and only when it is employed, may continue to be exploited throughout the session.

These results overall indicate that the specific contributions drawn from non-numerical dimensions depend on the specific task at hand. Interestingly, judging the number of members in a group as in the "people" condition shows relatively smaller bias from non-numerical dimensions (specifically by the size dimension), suggesting that such kind of perceptual choice is more heavily driven by numerical as opposed to non-numerical information. Conversely, when judging the amount of food, perceptual choices are more easily biased by non-numerical information, and particularly by the size of the individual items. In a somehow counterintuitive way, the bias provided by the dimension of size in the "food" condition is in the negative direction, suggesting that smaller size is more often selected when choosing the patch with more food items. Nevertheless, this pattern is plausible considering the nature of the task. The task required participants to choose the patch with more apples to collect. In this context, a negative weight for the size dimension might reflect the fact that it is easier to collect a more numerous set of smaller items than larger items. More specifically, one way to interpret this pattern of results is by considering the possibility of the fictional character to grasp and collect the food items. Previous work in the literature concerning grasping movement reports that attempting to grasp a large object (reaching the limits of graspable size), poses severe limitations to the grasping movement (Bootsma et al., 1994). If participants project such kind of limitations of grasp movements to the fictitious character involved in the task, this may explain the negative contribution of size to numerical judgments. Indeed, the fictitious character was similar in size to the dots, which makes sense in the people condition but is less realistic in the food condition. Thus, the preference for smaller items might be driven by the need of avoiding food items potentially very difficult to grasp or carry (i.e. items almost as big as the character itself), due to the physical limitations of grasp movements. According to this interpretation, participants would thus be more prone to choose items that appear to be more likely graspable by the character. Interestingly, if this is correct, it may be possible to reverse the effect of size by modulating relative size of the character and the food items – an interesting possibility that should be addressed by future studies.

Alternatively, a second interpretation could be advanced in light of previous results showing a peculiar negative relation between object size and perceived numerosity – i.e. whereby smaller items tend to be slightly overestimated. For instance, in an early study by Ginsburg & Nicholls (1988) where item size was modulated along with numerosity, participants systematically tended to overestimate smaller items and underestimate larger items. A relative overestimation of smaller-sized items has more recently been reported by other studies (Tokita & Ishiguchi, 2010, 2013) making it possible that this peculiar effect may represents a baseline feature of numerosity perception. Although this negative relation between size and perceived numerosity seems in contrast with reports by DeWind et al. (2015) and Starr et al. (2017), it should be noted that unlike DeWind et al. (2015) and Starr et al. (2017) studies that found a negative relation between size and perceived numerosity utilized a much longer presentation duration (Ginsburg & Nicholls, 1988; Tokita & Ishiguchi, 2013) allowing a much more deliberate choice, as in our current design. If that negative relation between object size and perceived numerosity is a peculiar feature of numerosity perception, an implicit strategy in the use of the size cue could have been employed in the people condition rather than in the food condition. In this scenario, the effect of size in the food condition would represent a baseline tendency to overestimate sets with smaller items, while the lack of effect of size in the people condition would represent a suppression of such bias to achieve a better estimate of the number of individuals in a group. In reality, of course, the inconsistencies between different studies makes it difficult to strongly support one interpretation over the other, and it is plausible that the results of our study reflect a combination of two implicit strategies: a more pronounced negative effect of size for food, and a suppression of the effect of size for people. While the main goal of our work was to demonstrate that different task contexts give rise to different behavioral patterns in a simple numerosity perception task, clarifying which context provides the stronger contribution and how different strategies are carried over to different contexts represents an interesting open question for future studies.

Besides the general effect of the two task framings, one additional interesting question is whether the contributions of non-numerical cues could be further modulated by other internal variables. For instance, when estimating amounts of food, it naturally follows that estimation may be modulated by a participant’s hunger level. How internal state in combination with task context may influence a seemingly simple perceptual task is an interesting question that should be addressed in future studies.

Another important question following from these findings is: what are the neural bases of such flexible use of numerical and non-numerical information for perceptual decision making? Does it represent flexibility of the sensory mechanisms extracting magnitude information from a visual scene, possibly enabled by top-down influences determining which information to be extracted in early visual areas? Or, does it represent flexibility at the decision stage, where different information might be exploited to guide behavior? Previous studies focusing on the neural correlates of numerosity perception highlight a complex stream of processing stages, showing both a cascade (i.e. a series of processing stages emerging over time; Park et al., 2016; Fornaciai et al., 2017) and feedback dynamics (i.e. as suggested by potential interactions across multiple perceptual systems in numerosity processing; Fornaciai & Park, 2017; Fornaciai & Park, under review). More specifically, previous work has demonstrated evidence for numerosity processing in subcortex (Collins et al., 2017), as early as V2/V3 in cortex (Fornaciai et al., 2017; Fornaciai & Park, under review), and higher-level regions such as the intraparietal sulcus (Piazza et al., 2004; Harvey et al., 2016) and in prefrontal areas (Viswanathan & Nieder, 2013; Nieder, 2016).

The first possibility is that the most relevant information for the task at hand is directly encoded starting from the earliest level of numerosity processing (e.g. V2/V3 as found in Fornaciai et al., 2017). Indeed, according to Lennie (1998), the primary visual cortex (V1) might contain a multi-dimensional representation of the visual scene, encoding several visual attributes that are relayed to specific areas and used to serve different aspects of visual sensory processing. In this view, while V1 might contain a representation of all the magnitude dimensions of a dot array stimulus, later areas such as V3 might exploit the most relevant information according to the specific task. Then, how does the information get selected at such an early stage? One plausible explanation is that feedback from higher-level areas (either on a trial-by-trial basis, or developed over the course of the experiment) determines what kind of information is preferentially processed starting from the earliest level of numerosity processing. Indeed, there is evidence showing that the information represented in early visual areas could be determined by higher-level influences (e.g. Lee et al., 2002; Gilbert & Li, 2013). Interestingly, according to a recent theoretical framework proposed by Roelfsema & de Lange (2016), early visual cortex might represent a cognitive blackboard used by high-level regions for read and write operations. The flexibility provided by such a mechanism indeed fits well with the idea that information is selected on the basis of current goals. According to the contextual information provided by the task context, decision-related high-level areas (i.e. possibly the dorsolateral prefrontal cortex, related to decision control; Rahnev et al., 2011; Rahnev 2017) might directly determine which information is selectively used across all the processing stream.

The second possibility is that different sources of information are exploited at the decision level in order to guide behavior according to the current goals, but without any change in the information encoded at earlier levels. However, in this context, such explanation appears less likely, as it would require different sensory information to be preserved throughout multiple processing stages. Indeed, according to Lennie (1998), only very early visual areas such as V1 contains an exhaustive multidimensional representation of the visual scene, while downstream in the processing hierarchy only the results of computations carried out at each specific level is passed on to the next level. Also, neurophysiological evidence demonstrates that even at the earliest level of numerosity processing (feed-forward processing in early visual cortex, at or before 100 ms after stimulus onset; Fornaciai et al., 2017), brain responses are already driven by specific contributions from different magnitude dimensions (and mostly by numerosity). Even if these studies exploited very different tasks, this evidence seems to support the idea that only the relevant information is passed on throughout the visual processing stream. According to this reasoning, the most likely explanation is that the information extracted and processed starting from early visual stages is determined by higher-level influences modulating sensory activity according to the specific context.

By considering this latter interpretation, these results provide important implications for the idea of a visual sense of number. Indeed, while current frameworks of numerosity perception regard it as a very basic perceptual function (i.e. Anobile et al., 2016; Cicchini et al., 2016), our results show that the information used to process and represent approximate numerosities is not hardwired in sensory processing, but flexibly determined according to the goals of the task at hand. This finding thus expands previous reports concerning the contributions of numerical and non-numerical dimensions to numerical judgments in more general experimental contexts (i.e. estimating number of dots; DeWind et al., 2015). Moreover, by considering the neural correlates of numerosity processing pinpointed in early visual areas (i.e. V2 and V3; Fornaciai et al., 2017), these results also add novel evidence to the literature documenting the remarkable plasticity of early visual sensory processing. In particular, these results appear consistent with earlier reports showing modulation of neuronal responses in V1 as a function of the task at hand (e.g. Li et al., 2004): even if the same stimuli are presented, the same neurons in primary visual cortex show different response profiles and tuning curves according to the specific task performed. Following

Li et al.'s (2004) interpretation, the present results then support the idea that early visual areas are adaptable processing units analyzing relevant stimulus information according to the task context.

Overall, our results show that numerosity perception is not a fixed mechanism. Rather, it has a remarkable flexibility even with simulated task contexts provided in a laboratory environment. Namely, the relevant information driving a quantitative decision is flexibly determined as a function of contextual information, likely by means of feedback from higher-level areas to earlier sensory cortices. This flexibility of numerical cognition well reflects the adaptive nature of approximate numerical abilities.

ACKNOWLEDGEMENTS

We thank Brynn Boutin for assistance in data collection, and Dr. Matthew Davidson for providing lab space and materials for testing. This study was supported by the National Science Foundation CAREER Award BCS1654089 to J. P.

Declarations of interest: none.

REFERENCES

Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2008). Do fish count? Spontaneous discrimination of quantity in female mosquitofish. *Animal Cognition*, *11*(3), 495–503. https://doi.org/10.1007/s10071-008-0140-9

Anobile, G., Cicchini, G. M., & Burr, D. C. (2016). Number As a Primary Perceptual Attribute: A Review. *Perception*, *45*(1–2), 5–31. https://doi.org/10.1177/0301006615602599

Bootsma, R. J., Marteniuk, R. G., MacKenzie, C. L., & Zaal, F. T. J. M. (1994). The speed-accuracy trade-off in manual prehension: effects of movement amplitude, object size and object width on kinematic characteristics. *Experimental Brain Research*, *98*(3), 535–541. https://doi.org/10.1007/BF00233990

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Burr, D., & Ross, J. (2008). A Visual Sense of Number. *Current Biology*, *18*(6), 425–428. https://doi.org/10.1016/j.cub.2008.02.052

Cicchini, G. M., Anobile, G., & Burr, D. C. (2016). Spontaneous perception of numerosity in humans. *Nature Communications*, *7*, 12536. https://doi.org/10.1038/ncomms12536

Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.

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*, 247–65. https://doi.org/10.1016/j.cognition.2015.05.016

Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Current Biology*, *18*(18), R855–R856. https://doi.org/10.1016/j.cub.2008.07.053

Fornaciai, M., Cicchini, G. M., & Burr, D. C. (2016). Adaptation to number operates on perceived rather than physical numerosity. *Cognition*, *151*, 63–67. https://doi.org/10.1016/j.cognition.2016.03.006

Fornaciai, M., Brannon, E. M., Woldorff, M. G., & Park, J. (2017). Numerosity processing in early visual cortex. *NeuroImage*, *157*, 429–438. https://doi.org/10.1016/j.neuroimage.2017.05.069

Fornaciai, M., & Park, J. (Under review). Dynamics of numerosity representation in early visual cortex.

Fornaciai, M., & Park, J. (2017). Distinct Neural Signatures for Very Small and Very Large Numerosities. *Frontiers in Human Neuroscience*, *11*(January), 1–14. https://doi.org/10.3389/fnhum.2017.00021

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. https://doi.org/10.1016/j.actpsy.2016.09.003

Gelman, R., & Cordes, S. (2001). Counting in Animals and Humans. *Language, Brain, and Cognitive Development: Essays in Honor of Jacques Mehler.*, *XXXIII*(2), 279–301. https://doi.org/10.1007/s13398-014-0173-7.2

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*. https://doi.org/10.1038/nrn3476

Ginsburg, N., & Nicholls, A. (1988). Perceived numerosity as a function of item size. *Perceptual and Motor Skills*, *67*(2), 656–658. https://doi.org/10.2466/pms.1988.67.2.656

Harvey, B. M., Klein, B. P., Petridou, N., & Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science*, *341*(September), 1123–1126. https://doi.org/10.1126/science.1239052

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(25), 10382–5. https://doi.org/10.1073/pnas.0812142106

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception ECVP 2007 Abstract Supplement*, 14. https://doi.org/10.1068/v070821

Lee, T. S., Yang, C. F., Romero, R. D., & Mumford, D. (2002). Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. *Nature Neuroscience*, *5*(6), 589–597. https://doi.org/10.1038/nn860

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. https://doi.org/10.1017/S0140525X16000960

Lennie, P. (1998). Single units and visual cortical organization. *Perception*, *27*(8), 889–935. https://doi.org/10.1068/p270889

Nieder, A. (2016). The neuronal code for number. *Nature Reviews. Neuroscience*, *17*(6), 366–82. https://doi.org/10.1038/nrn.2016.40

Park, J. (2017). A neural basis for the visual sense of number and its development: A steady-state visual evoked potential study in children and adults. *Developmental Cognitive Neuroscience*. https://doi.org/10.1016/j.dcn.2017.02.011

Park, J., Dewind, N. K., Woldorff, M. G., & Brannon, E. M. (2016). Rapid and Direct Encoding of Numerosity in the Visual Stream. *Cerebral Cortex*, *26*(2), 748–763. https://doi.org/10.1093/cercor/bhv017

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Pepperberg, I. M. (2006). Cognitive and communicative abilities of Grey parrots. *Applied Animal Behaviour Science*, *100*(1–2), 77–86. https://doi.org/10.1016/j.applanim.2006.04.005

Piantadosi, S. T., & Cantlon, J. F. (2017). True Numerical Cognition in the Wild. *Psychological Science*, *28*(4), 462–469. https://doi.org/10.1177/0956797616686862

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, *44*(3), 547–555. https://doi.org/10.1016/j.neuron.2004.10.014

Rahnev, D. (2017). Top-Down Control of Perceptual Decision Making by the Prefrontal Cortex. *Current Directions in Psychological Science*, 96372141770980. https://doi.org/10.1177/0963721417709807

Rahnev, D., Maniscalco, B., Graves, T., Huang, E., De Lange, F. P., & Lau, H. (2011). Attention induces conservative subjective biases in visual perception. *Nature Neuroscience*, *14*(12), 1513–1515. https://doi.org/10.1038/nn.2948

Roelfsema, P. R., & de Lange, F. P. (2016). Early Visual Cortex as a Multiscale Cognitive Blackboard. *Annual Review of Vision Science*, *2*(1), 131–151. https://doi.org/10.1146/annurev-vision-111815-114443

Rugani, R., Vallortigara, G., Priftis, K., & Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' mental number line. *Science*, *347*(6221), 534–536. https://doi.org/10.1126/science.aaa1379

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. https://doi.org/10.1016/j.cognition.2017.07.004

Tokita, M., & Ishiguchi, A. (2010). How might the discrepancy in the effects of perceptual variables on numerosity judgment be reconciled? *Attention, Perception, and Psychophysics*, *72*(7), 1839–1853. https://doi.org/10.3758/APP.72.7.1839

Tokita, M., & Ishiguchi, A. (2013). Effects of perceptual variables on numerosity comparison in 5-6-year-olds and adults. *Frontiers in Psychology*, *4*(JUL). https://doi.org/10.3389/fpsyg.2013.00431

Viswanathan, P., & Nieder, A. (2013). Neuronal correlates of a visual "sense of number" in primate parietal and prefrontal cortices. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(27), 11187–92. https://doi.org/10.1073/pnas.1308141110

Watson, A. B. (1979). Probability summation over time. *Vision Research*, *19*(5), 515–522. https://doi.org/10.1016/0042-6989(79)90136-6

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition*, *89*(1), B15–B25. https://doi.org/10.1016/S0010-0277(03)00050-7

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11. https://doi.org/10.1016/S0010-0277(99)00066-9