

The Effect of Clustering on Perceived Quantity in Humans (*Homo sapiens*) and in Chicks (*Gallus gallus*)

Marco Bertamini and Martin Guest
University of Liverpool

Giorgio Vallortigara
University of Trento

Rosa Rugani and Lucia Regolin
University of Padova

Animals can perceive the numerosity of sets of visual elements. Qualitative and quantitative similarities in different species suggest the existence of a shared system (approximate number system). Biases associated with sensory properties are informative about the underlying mechanisms. In humans, regular spacing increases perceived numerosity (regular-random numerosity illusion). This has led to a model that predicts numerosity based on occupancy (a measure that decreases when elements are close together). We used a procedure in which observers selected one of two stimuli and were given feedback with respect to whether the choice was correct. One configuration had 20 elements and the other 40, randomly placed inside a circular region. Participants had to discover the rule based on feedback. Because density and clustering covaried with numerosity, different dimensions could be used. After reaching a criterion, test trials presented two types of configurations with 30 elements. One type had a larger interelement distance than the other (high or low clustering). If observers had adopted a numerosity strategy, they would choose low clustering (if reinforced with 40) and high clustering (if reinforced with 20). A clustering or density strategy predicts the opposite. Human adults used a numerosity strategy. Chicks were tested using a similar procedure. There were two behavioral measures: first approach response and final circumnavigation (walking behind the screen). The prediction based on numerosity was confirmed by the first approach data. For chicks, one clear pattern from both responses was a preference for the configurations with higher clustering.

Keywords: numerosity, occupancy, clustering

Supplemental materials: <http://dx.doi.org/10.1037/com0000114.supp>

Often in a visual scene, there are distinct elements that form a set, such as people in a room, fruits on a tree, or spots on fabric. For large sets, the exact number can be found through counting, but the approximate number can be estimated (Feigenson, Dehaene, & Spelke, 2004). The ability to perform nonsymbolic number estimation is present in nonhuman species (Adachi, 2014; Brannon & Terrace, 1998; Rugani, Cavazzana, Vallortigara, & Regolin, 2013; Rugani, Fontanari, Simoni, Regolin, & Vallortigara, 2009, Vallortigara, 2017a), preverbal infants (Cordes & Brannon, 2009; de Hevia, Izard, Coubart, Spelke, & Streri, 2014), and in adult humans whenever the use of language is prevented (Cantlon & Brannon, 2007; Cordes, Gelman, Gallistel, & Whalen,

2001). Two common characteristics of nonsymbolic numerical estimation are the distance effect (numerical tasks become easier as the difference between the numbers increases) and the size effect (numerical tasks become more difficult as the magnitude of numbers increases) for both humans (Moyer & Landauer, 1967) and nonhuman animals (Cantlon & Brannon, 2007; Scarf, Hayne, & Colombo, 2011). A qualitatively and quantitatively similar performance in different species, including human adults, suggests a shared nonverbal number estimation process (Cantlon & Brannon, 2007; Dehaene, 2011). These results have led to the introduction of the terms number sense and approximate number system (Burr & Ross, 2008; Dehaene, 1992, 2003, 2011; Halberda & Feigenson, 2008).

There is evidence of the representation of numerosity in the parietal lobe. For example, Piazza, Izard, Pinel, Le Bihan, and Dehaene (2004) found that the bilateral intraparietal sulci respond selectively to a change in number. In humans and monkeys, the intraparietal sulcus is implicated in the representation of numerals and numerosities (Eger et al., 2009; Pearson, Hayden, Raghavachari, & Platt, 2009; Piazza et al., 2004). Nieder and Merten (2007) analyzed single-cell recordings and found neurons that respond maximally to a preferred number, and with decreasing strength to neighboring numbers. A classic feature of perception is

This article was published Online First April 30, 2018.

Marco Bertamini and Martin Guest, Department of Psychological Sciences, University of Liverpool; Giorgio Vallortigara, Centre for Mind/Brain Sciences, University of Trento; Rosa Rugani and Lucia Regolin, Department of General Psychology, University of Padova.

Correspondence concerning this article should be addressed to Marco Bertamini, Department of Psychological Sciences, Institute of Psychology, Health & Society, Eleanor Rathbone Building, University of Liverpool, Liverpool L69 7ZA. E-mail: m.bertamini@liverpool.ac.uk

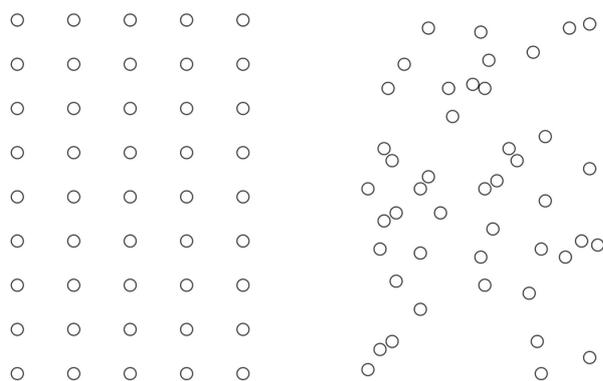
that perceived change is proportional to the strength of the stimulus (Weber–Fechner law), both in human and animals (Cantlon & Brannon, 2006; Cantlon & Brannon, 2007; Jordan & Brannon, 2006). Single cells that respond to numerosity with a logarithmically compressed scaling of numerical information, thus obeying the Weber–Fechner law, have been described in a region of the avian brain, the nidopallium caudolaterale (Ditz & Nieder, 2015). Anatomical correspondences of this region with regions of the mammalian brain are currently not well established (Vallortigara, 2017b).

Perception of numerosity is affected by a number of biases in relation to visual properties of the image. In particular, there are known effects of the size of the elements (Ginsburg & Nicholls, 1988; Hurewitz, Gelman, & Schnitzer, 2006; Tokita & Ishiguchi, 2010), the regularity of the pattern (Frith & Frith, 1972), the spacing of elements (Ginsburg, 1976, 1991), and the area of the configuration (Dakin, Tibber, Greenwood, Kingdom, & Morgan, 2011; Hurewitz et al., 2006; Tokita & Ishiguchi, 2010).

Spacing of the Elements and Clustering

In this study, we are interested in the issue of spacing. This effect was originally described as the random-regular numerosity illusion (Cousins & Ginsburg, 1983; Ginsburg, 1980). When the elements are more evenly separated, the configuration appears more numerous (Figure 1). It is likely that regularity has an effect on perceived numerosity because regularly spaced elements are not approaching each other (Vos, Van Oeffelen, Tibosch, & Allik, 1988). In other words, spatial proximity (and therefore clustering) rather than regularity is the important factor.

This effect of clustering on numerosity has led to the development of the occupancy model (Allik & Tuulmets, 1991). The key idea is that each element has a region of influence; if other elements come within this region and therefore the two regions overlap, their effectiveness is reduced. Therefore, elements forming clusters contribute less than the same elements in isolation. The occupancy model uses the total area as a predictor of perceived numerosity; total area is lower the more intersections exist be-



Regular-random numerosity illusion

Figure 1. Most people judge the first configuration (regular) as more numerous than the second (random). The spatial distribution of elements matches the example in the study by Ginsburg (1980).

tween regions of influence (Allik & Tuulmets, 1991; Burgess & Barlow, 1983).

The occupancy model has received support from recent studies. For example, Valsecchi, Toscani, and Gegenfurtner (2013) have shown that perceived numerosity depends on clustering and on eccentricity. Their implementation of clustering was based on constraining the center-to-center distance of the elements. Greater clustering reduced perceived numerosity (measured as a shift in the point of subjective equality). They also found reduced perceived numerosity in the periphery; on this basis, they argued that crowding plays a role (on “crowding,” see e.g., Cicchini, Anobile, & Burr, 2016; Pelli & Tillman, 2008).

Bertamini, Zito, Scott-Samuel, and Hulleman (2016) used a procedure in which observers were forced to choose the pattern that appeared more numerous between two patterns with the same number of elements. In addition to the task of selecting on the basis of numerosity, they used tasks in which observers were asked to judge perceived clustering and perceived dispersion. The positions of the elements within a circle were not constrained, but the measures of clustering were used to predict which of the two images was selected in each task. Bertamini et al. found that people could judge numerosity, dispersion, and clustering and that they used different information to do so (low correlations between responses in different tasks). They also confirmed that estimates of numerosity were affected by clustering (total area of the regions of influence). Although this had been demonstrated before, this was the first time clustering was shown to matter in unconstrained configurations where variability in clustering was purely random.

Density

Density affects perceived numerosity, and density can be manipulated in terms of interstimulus distance. In comparative studies, an explanation for density effects comes from optimal foraging models. Children overestimate sets that are more spaced (Piaget, 1952), but when Uller, Urquhart, Lewis, and Berntsen (2013) presented 10-month-old infants with sets of edible items (cookies), one densely arranged and the other sparsely arranged, infants chose the denser set. Stevens, Wood, and Hauser (2007) tested cotton-top tamarins (*Saguinus oedipus*) and marmosets (*Callithrix jacchus*). They confirmed a preference for densely arranged items relative to the same amount of food sparsely arranged for cotton-top tamarins.

Based on the foraging theory, organisms maximize caloric intake while reducing energy expenditure (Stephens & Krebs, 1986). It may be more efficient to choose dense patches of food, depending on the cost of moving from one patch to another. Consistently with this hypothesis, there is a preference for more densely baited food sites in species such as orangutans (MacDonald & Agnes, 1999) and pigeons (Mitchell, Calton, Threlkeld, & Schachtman, 1996).

Parrish, James, and Beran (2017) studied the density bias in capuchin monkeys (*Cebus apella*) and rhesus monkeys (*Macaca mulatta*). There was a density bias in capuchin monkeys and to a lesser extent in rhesus monkeys. Interestingly, Parrish et al. analyzed how this density bias was related to an illusion of numerosity known as the Solitaire illusion (Frith & Frith, 1972). They did not find evidence for a link.

The name of the Solitaire illusion derives from a board game, and the configuration is shown in Figure 2. The illusion is robust and depends on grouping. The effect is, therefore, consistent with the density bias, but it seems paradoxically the opposite of the regular-random illusion in that the white dots are more spaced out and the black dots more clustered. However, as we have already mentioned, area affects numerosity, and the area occupied by the black elements is larger. This is true if we do not take the white dots as a single group but as the sum of four small groups, each covering a small area.

Early work on the role of extent on numerosity was conducted by Ponzo using the, now famous, Ponzo illusion (Ponzo, 1928). We have, therefore, placed the two illusions side-by-side in Figure 2. If the region with black elements is perceived as larger, this may be the driving factor in the Solitaire illusion. To study density and clustering on their own, it is, therefore, essential to control overall area.

Although there is a compelling body of evidence for the existence of a sense of number, it is necessary to understand how perceptual systems achieve this estimation. In this context, it is useful to study cues that bias numerosity estimation. We have mentioned some cues earlier in the introduction, such as size, spacing, and density. The regular-random numerosity illusion and the Solitaire illusion are also examples of properties of the configuration that bias perceived numerosity.

Numerosity as a Salient Dimension

If approximate numerosity is perceived for any large sets of visual elements, do individuals spontaneously use differences in numerosity to discriminate and classify stimuli? Cantlon and Brannon (2007) asked the question whether individuals spontaneously represent the numerical attributes of their environments. They tested rhesus monkeys and compared the influence of number with that of shape, color, and area on a matching task with more than one correct answer: There was a numerical match and a non-numerical match (based on color, area or shape). The results were clear: All monkeys based their decisions on the difference in numerosity. A similar conclusion applies to human infants. In a recent review, Ferrigno and Cantlon (2017) concluded that non-

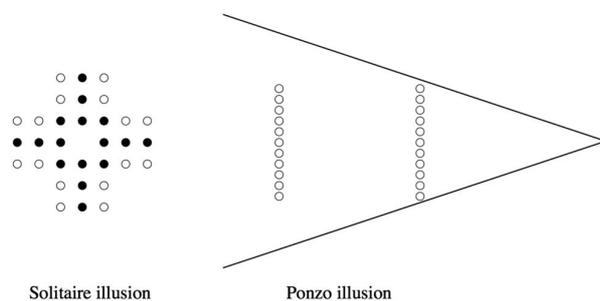


Figure 2. On the left, most people judge the black elements as more numerous than the white elements. The spatial distribution of elements matches the example in the study by Frith and Frith (1972). On the right, most people judge the elements nearer the apex as more numerous. The distribution of elements is based on the example in the study by Ponzo (1928), also reproduced in Vicario (2011).

verbal numerical reasoning about physical objects is the first type of numerical cognition that emerges in human development. Moreover, human adults are more sensitive to numerosity than to density or area. Cicchini et al. (2016) found that when stimuli varied in density, area, and numerosity, observers reacted with greater sensitivity to changes in numerosity relative to the other dimensions. Not only can humans and other animals estimate numbers, but they also seem to use the number as an important dimension in the environment.

The Approximate Number System and Normalization

In a polemic against the mainstream view, Gebuis, Kadosh, and Gevers (2016) have argued that numerosity is estimated by a sensory integration system that compares stimuli by integrating different sensory cues related to changes in number. However, proponents of the approximate number system are aware that irrelevant properties of the stimulus affect numerosity. They argue that the necessary normalization stage may not be perfect (Dehaene & Changeux, 1993) and that interference may also exist after numerosity has been computed, at the decision or response stage (Ingilis & Gilmore, 2013).

Estimation of numerosity has to rely on spatial properties of the patterns. More regular patterns, with less clustering of the elements, are overestimated relative to less regular patterns. These clustering effects may highlight the way the estimation process works, and they are a byproduct of computations carried out by an ancient and shared system. Based on this hypothesis, the role of clustering should be robust across individuals, groups, and even different species.

An alternative view is that the approximate number system achieves normalization and is unaffected by correlations between numerosity and other cues. There are effects of irrelevant cues, such as clustering, but they are separate from numerosity estimation per se. Based on this view, the role of irrelevant cues on numerosity judgments should not be robust across individuals, groups, and different species. It is therefore necessary to study the effect of regularity and clustering in people and in other species to compare and analyze the similarities and differences.

The Procedure

In this study, we have used with humans a procedure that has been developed in the study of animal behavior. In the first phase, individuals are reinforced when they select one of a pair of stimuli. In the case of numerosity, some individuals are reinforced when they select the higher numerosity, and other individuals are reinforced when they select the lower numerosity. For nonhuman animals, the reinforcement is usually food. For human observers, the reinforcement was a “correct” message presented after the choice was performed. In the second phase of the procedure, individuals have to choose between a pair of novel stimuli. Here the numerosity was the same in the pair of stimuli but there was a difference in the type of configuration. We used configurations with more or less clustering, by means of a minimum interdot distance constraint similar to that used by Valsecchi et al. (2013). Therefore, if individuals could estimate numerosity independently from clustering, there would not be any preference, but if they perceived a difference in numerosity due to clustering, they would be guided by that variable.

A similar approach has already been used to study numerosity with humans and monkeys. Although we were not aware of that study when designing our own, the logic of our experiment is similar to that of Experiment 2 in Beran (2006). In that study, he compared humans to four rhesus monkeys and used dot configurations that varied in regularity. However, following the type of stimuli used in the original work on the regular-random illusion (Ginsburg, 1980; Figure 1), the regular configuration was a perfect square matrix. There was evidence that both humans and monkeys experienced the regular-random illusion, but with some individual differences. Different methodology provided different results, including some cases where monkeys had a preference for the less regular configuration (Experiment 1).

More recently, Agrillo, Parrish, and Beran (2014) tested capuchin monkeys, rhesus monkeys, and chimpanzees using the Solitaire illusion. Overall, there was no or weak evidence of the configuration affecting perceived numerosity in these species using this illusion. Parrish, Agrillo, Perdue, and Beran (2016) further studied the Solitaire illusion in children and capuchin monkeys. There was some evidence that the monkeys perceived the illusion, although there were large individual differences. Interestingly, similar large individual differences were also present in the young children. Parrish et al. conclude that perhaps experience plays a role in shaping the emergence of the Solitaire illusion.

Because of the importance of experience, chicks provide a unique opportunity to test numerosity perception in animals with minimal experience (a few days after hatching) and no social influence (the chicks are kept in isolation).

Experiment 1: Human Adults

We tested human adults using the modified reinforcement task. Participants, naïve to the purpose of the study, were presented with two images of red dots. They were told to start by selecting either stimulus and that they would receive feedback as to whether it was correct. Images were presented sequentially so that observers could look at the center of the screen, and to control amount of time per image. Participants were randomly assigned to the condition in which the 20 dots stimulus was the correct choice or the condition in which the 40 dots stimulus was the correct choice.

The feedback is expected to lead individuals to use a rule. After they reached a criterion, the training phase was followed by an experimental phase. New stimuli were interleaved with the original stimuli. All stimuli in the new pairs had 30 dots, but the two images differed in the minimum interelement distance (MD). In one, the minimum distance was smaller than the original, and in the other it was greater than the original. Therefore, the test stimuli did not differ in density (over the whole area), but they differed in amount of clustering Figure 3. Based on the regular-random numerosity illusion, one would predict that stimuli with higher values of MD are perceived as more numerous. This prediction is related to occupancy, which increases with higher values of MD.

It is important to analyze which rules participants could use. An obvious choice is numerosity; it is easy to perceive a difference and estimate which set has more dots (20 or 40). However, numerosity was confounded with density as dots were always within the same circular area. There is a debate in the literature on whether judgments of numerosity and density are based on the same mechanism (Anobile, Turi, Cicchini, & Burr, 2015; Tibber,

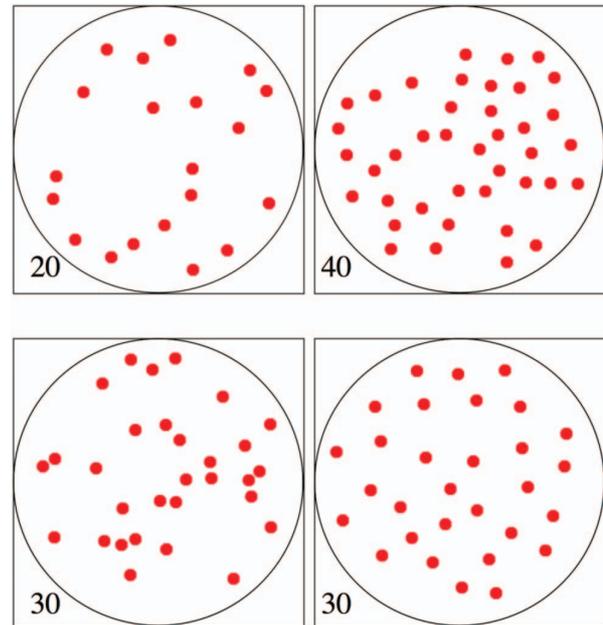


Figure 3. The first pair of stimuli is an example of stimuli used during training. Numerosity was always 20 in one pattern and 40 in the other. The dot had a radius of 10 pixels. The minimum interelement distance was 3.6 times the radius (36 pixels). The second pair of stimuli are examples used in the experimental phase. Numerosity was always 30 dots. However, one pattern has a minimum interelement distance of 2.1 radii or 21 pixels (left example) and the other a minimum of 5.1 radii or 51 pixels (right example). In the case of 2.1 radii, elements could be close but without touching. See the online article for the color version of this figure.

Greenwood, & Dakin, 2012). Density increased the clustering of elements. Therefore, effective strategies were available based on numerosity or clustering. For clustering, the prediction for the choice between the test stimuli is the opposite than what is predicted by occupancy (Figure 4 and Supplementary analysis). There are, therefore, two different predictions: one based on occupancy (a cue to numerosity) and one based on clustering.

Experiment 1 used human adults, and the correct response was reinforced by feedback. Experiment 2 used domestic chicks (*Gallus gallus*) within few days of hatching. The reinforcement was food. We compare adult humans with domestic chicks for two reasons. We chose the domestic chick as an animal model because it masters a variety of numerical comparisons, such as 5 versus 10 or 6 versus 9 (Rugani, Vallortigara, & Regolin, 2013, Rugani, Vallortigara, & Regolin, 2014). Moreover, because chicks are a precocial species, they are characterized by quite sophisticated motoric capacities just days after hatching, thus allowing for a control of the role played by experience. In both cases, we were interested in whether individuals will use numerosity to select the answer and whether they will quickly learn this rule.

Method

Participants. A total of 48 psychology students took part in the study (30 female and 18 male). They were naïve with respect

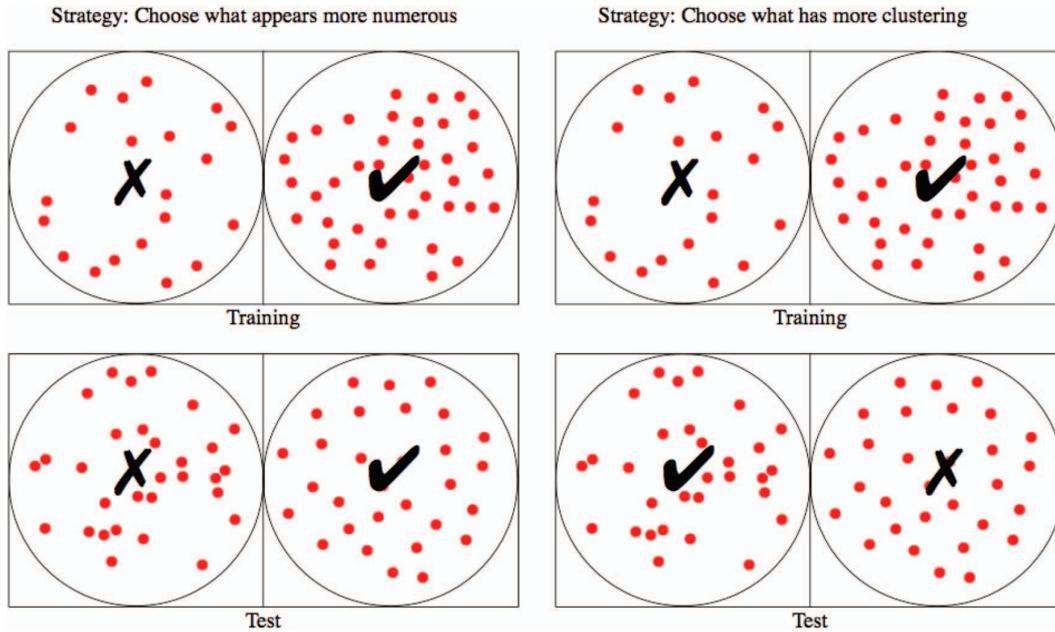


Figure 4. A strategy based on estimated numerosity predicts that when reinforced for $N = 40$, for test stimuli there will be a preference for the one with more spacing (minimum interelement distance = 51). A strategy based on clustering, or element proximity, predicts that when reinforced for $N = 40$, for test stimuli there will be a preference for the one with less spacing (minimum interelement distance = 21). For a quantification of these variables see Supplementary analysis. See the online article for the color version of this figure.

to the purpose of the study. Participants were alternatively (odd or even) assigned to two conditions: one in which low numerosity was the correct answer ($N = 24$) and one in which high numerosity was the correct answer ($N = 24$). The study had local ethics approval by the ethics committee (IPHS-1516-SMc-199) and was conducted in accordance with the Declaration of Helsinki (Revised 2008).

Materials. The procedure, including stimulus generation, was controlled by a Python program using the PsychoPy libraries (v1.80; Peirce, 2007, www.psychopy.org). The computer was a HP EliteDesk 800 G1 (Hewlett-Packard, Palo Alto, USA), with a Dell 17-in. CRT monitor and 60Hz refresh rate. Each red dot was a circle with a 10-pixel radius. Dots were placed inside a circular region of a 200-pixel radius and presented on a white background.

Procedure. Observers were seated at approximately 80 cm away from the monitor in a quiet room. On each trial, there was a fixation cross for 2500 ms, followed by two patterns. Each pattern was on the screen for 200 ms, and there was a 200-ms blank screen between them (Figure 5).

Participants were informed that there was a right and a wrong answer, but they would have to use trial and error to work out the correct response. They could choose the first pattern by pressing the “a” key on the keyboard or the second pattern by pressing the “l” key. Correct answers were followed by the word “correct” displayed on the screen. There was no feedback for incorrect answers. The experiment had two phases.

Training Phase: Each participant completed a minimum of 20 trials, and after that, the phase stopped based on accuracy. The

criterion to complete the training was 75% accuracy. The experiment would then move on to the next phase.

Experimental phase: To the participant, there was no indication of any change in the procedure, but in the experimental phase, new test stimuli were interleaved with the training stimuli. The length of the experimental phase was fixed at 36 trials. Of these, 24 were identical to the training phase (pairs of stimuli with 20 and 40 dots) and included feedback. The other 12 trials had a pair of stimuli with 30 dots. The difference between the two patterns was the MD. One pattern has a minimum inter element distance of 2.1 radii (21 pixels), and the other of 5.1 radii (51 pixels). For 2.1 radii, elements could be close but not touching. No feedback was provided on these trials. The computer stored responses for all trials.

At the end of the experiment, participants were asked to select what best described the strategy they used. Choices were: Greater number of dots, Fewer dots, Dots were more dense, Dots were less dense, and Other. We refer to the first two choices as numerosity strategy and the next two as density strategy (in the sense of lower interelement distance). This was followed by a debrief: Participants described, in their own words, what strategy they had used.

Results

Every individual tested reached the criterion. The number of trials necessary to reach it varied between 20 and 99. There were always 12 test pairs, and therefore the main analysis is on responses to these stimuli. If participants choose randomly, we expect equal number of responses to the MD = 21 (clustered) and

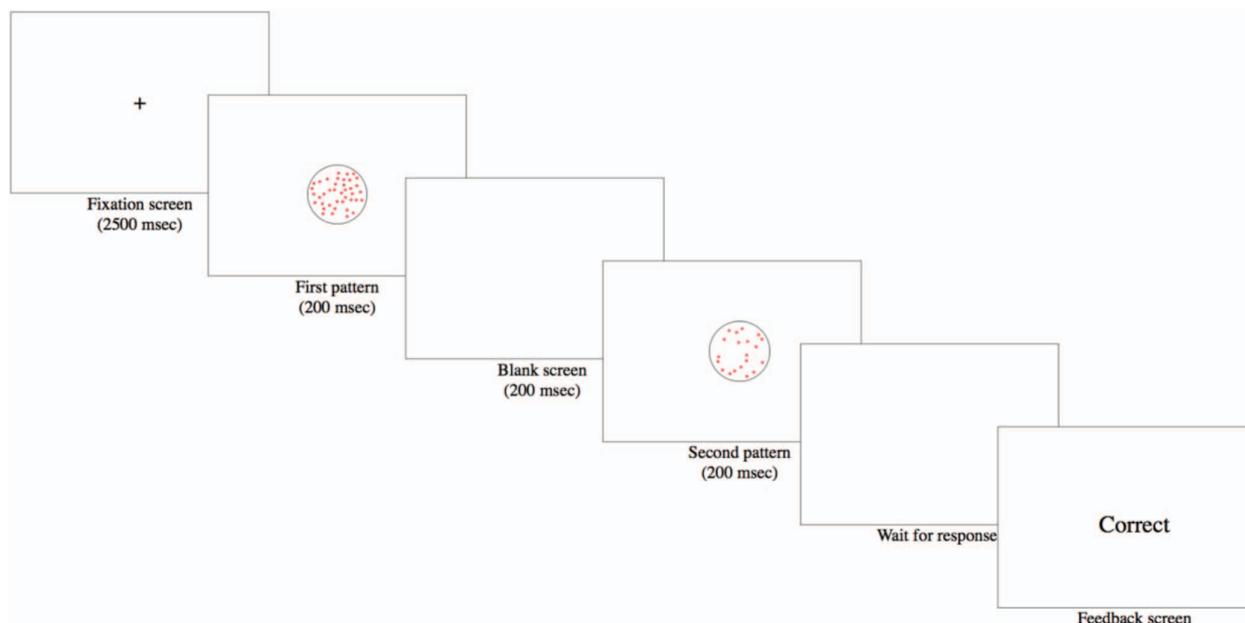


Figure 5. An illustration of the structure of a trial. The fixation was followed by two patterns, shown sequentially for 200 ms each, and by a blank screen. After a response was produced there was a feedback. See the online article for the color version of this figure.

MD = 51 (dispersed) stimuli. We tested two possible biases with two independent *t* tests. Clustered and dispersed stimuli were chosen 52% and 48% of the times, respectively, and this difference was not significant, $t(47) = 0.417, p = .678, \text{Cohen's } D = 0.06$. With respect to selecting the type of clustering depending on training, the prediction based on the illusion was supported by the data, 69%: $t(47) = 6.370, p < .001, \text{Cohen's } D = 0.92$.

Results support the hypothesis that most people responded to the test stimuli on the basis of which pattern appeared more or less numerous (Figure 6). The test stimuli always had 30 elements, but perceived numerosity was influenced by spacing. This result therefore supports the occupancy model of perceived numerosity.

Although the aggregate data suggest strong support for the idea that numerosity was the salient aspect of the stimuli and that

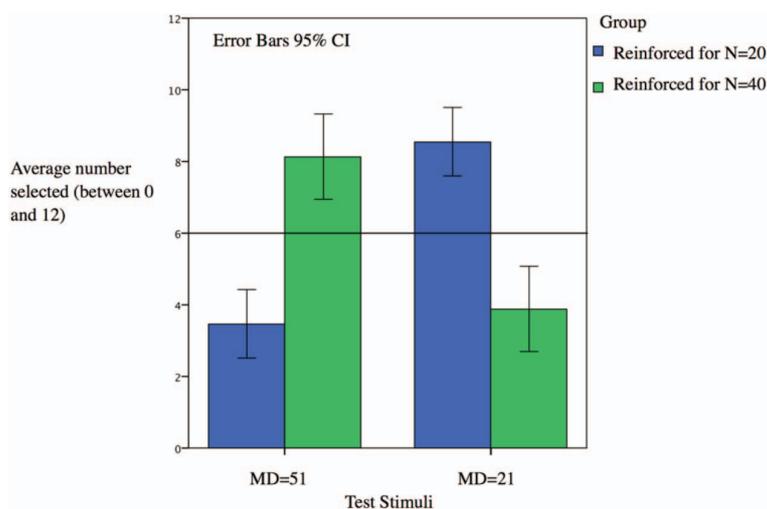


Figure 6. The average number of times that each type of test stimulus was selected. For each participant, this number can be between 0 and 12. Blue and green bars are the two groups with different reinforcement, and therefore the data are from different individuals. See the online article for the color version of this figure.

observers used a strategy based on numerosity, there were some exceptions. Figure 7 shows the data for all individuals.

Participants had to report whether they used one of three possible strategies: numerosity, density, or other. Out of 48, 34 said that they had used numerosity, 11 said they had used density, and three said other. The graphs of Figure 8, however, do not show any direct relationship between this answer and the results. This is surprising but we should consider the possible ambiguity of the terms. Some people may have used density as a word to refer to numerosity. The interview gave us a chance to understand the strategy. Ignoring those who displayed no strategy (six choices either way out of 12), there were eight participants who responded in a direction inconsistent with the majority and inconsistent with a strategy based on occupancy. However, interviews did not highlight any pattern common to all eight. One mentioned number of dots, four said they started using number but then switched to density, one mentioned dot spacing, and two did not make any comments. The spacing of the dots was mentioned by other participants, although again they may have been referring to density in general (elements in relation to area) rather than clustering.

Overall, Experiment 1 confirmed that human adults quickly develop a rule to select images based on difference in numerosity. A clear majority also reported this as part of the poststudy debrief. For the test stimuli, again a clear majority selected the more or less spaced stimulus as predicted by the occupancy model of numerosity.

Experiment 2: Chicks

Experiment 2 used a procedure as similar as possible to that of Experiment 1. We were interested in the similarity in the pattern of behavior across species. If the occupancy model is a good model to predict perceived numerosity in general, its prediction should be met for both studies. However, to the best of our knowledge, the

occupancy model has never been tested in species other than humans.

Method

Participants. Of 32 domestic chicks (*Gallus gallus*, Ross 608 Aviagen, Huntsville, USA) that began the study, 24 passed the minimum criteria for engagement with the apparatus (see the following text) and were used in the experiment. Sample size was calculated on the basis of previous researches on numerical discrimination in which a similar paradigm has been used (Rugani, McCrink, de Hevia, Vallortigara, & Regolin, 2016; Rugani et al., 2014). Chicks were obtained weekly, several hours after hatching, from a local commercial hatchery (Agricola Berica, Vicenza, Italy). The experiment complied with all applicable national and European laws concerning the use of animals in research and was approved by the Italian Ministry of Health (permit number: 1992). All procedures used in this study were examined and approved (permit number: 100845) by the Ethical Committee of the University of Padua (Comitato Etico di Ateneo per la Sperimentazione Animale—C.E.A.S.A.), as well as by the Italian National Institute of Health.

Procedure. On arrival, the chicks were housed individually in standard metal cages ($28 \times 32 \times 40$ cm). The rearing room was monitored for temperature (28°C – 31°C) and humidity (68%) and was illuminated by fluorescent lamps (36 W) 45 cm above the floor of the cage. Water and food, located in transparent glass jars (5 cm in diameter and 5 cm high) were available ad libitum. Chicks were offered mealworms (*Tenebrio molitor* larvae) twice a day, to familiarize them with this food, as that was used as reinforcement in all experimental phases. Chicks were kept in these rearing conditions from Monday morning (11 a.m.) to Wednesday morning (8 a.m.) before the food was removed (water was available). After a couple of hours (10 a.m.), chicks underwent individual

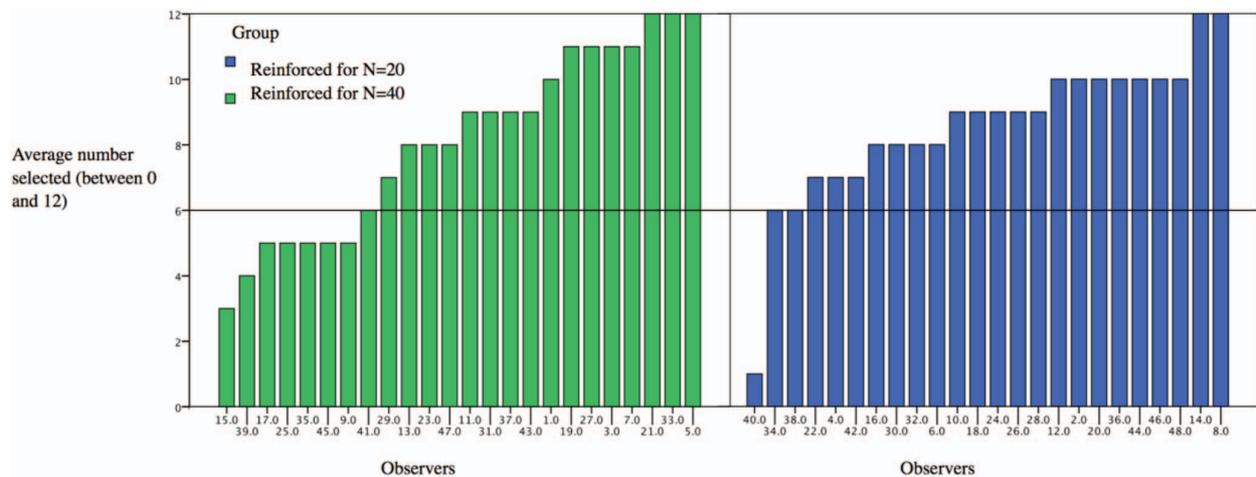


Figure 7. The graph shows how often the test stimulus predicted by a choice based on numerosity was selected by each observer. For each participant, this number can be between 0 and 12. The predicted response is the selection of the clustered stimulus for the group reinforced with $N = 20$ and the dispersed stimulus for the other group. Anything above 6 fits the prediction. Anything below suggests a different strategy. Blue and green bars are the two groups with different reinforcement and therefore the data are from different individuals. See the online article for the color version of this figure.

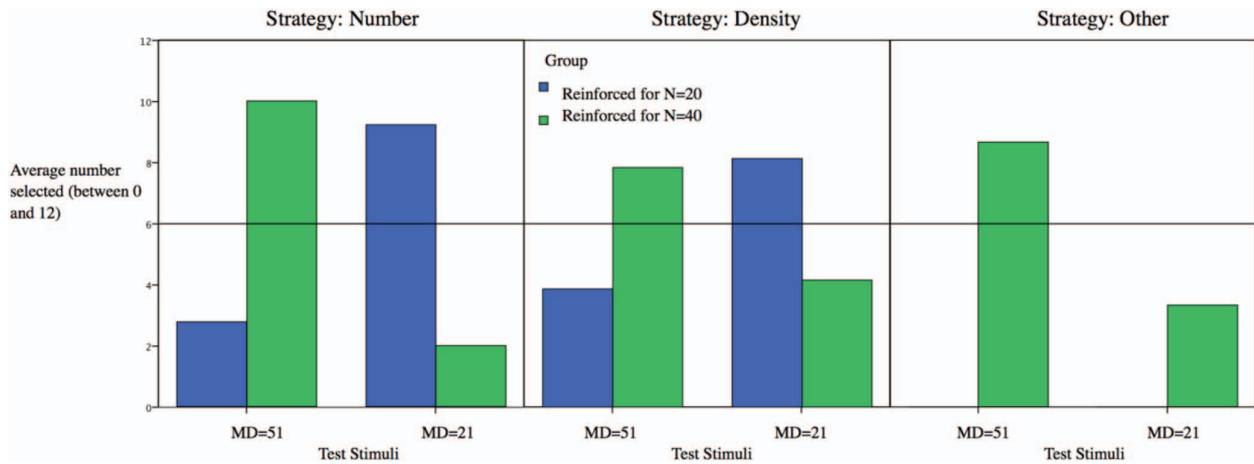


Figure 8. The same information in Figure 7 is replotted by dividing participants in three groups based on the strategy they reported at the end of the study. See the online article for the color version of this figure.

shaping and afterward were placed back in their home cages. They underwent individual training 2 hr later. After another 2 hr interval (in their home cages), they underwent testing individually. After the end of the behavioral observations, chicks were caged in social groups (three birds), with food and water available, and a few hours later all the chicks were donated to local farmers.

Apparatus. The apparatus was in a separate room (experimental room) in the same building as the rearing room. Temperature and humidity were controlled (25 °C and 70%, respectively) and four 58-W lamps provided lighting (placed on the ceiling, 194 cm above the floor of the apparatus). The experimental apparatus consisted of a diamond-shaped arena (height 20 cm) made of uniformly white plastic panels (for the dimensions of the arena, see Figure 9). The floor consisted of a white plastic basement. A starting area, delimited by a transparent removable partition (10 × 20 cm), was located in the proximity of one vertex of the arena, 10 cm from it. The transparent partition confined subjects for a few seconds before the beginning of each trial. Chicks, while in the starting area, had the opportunity to see the inner apparatus and the

stimuli. During familiarization, we used a single screen (16 × 8 cm) located in the center of the arena and 30 cm away from the starting position. During training and testing, we used two identical white plastic screens, located 30 cm away from the transparent partition and 6 cm away from the lateral walls. Screens had sides (3 cm) to prevent chicks from looking behind the screen where food reinforcement (*Tenebrio molitor* mealworm) was hidden. At the end of each trial, chicks were moved in a separate box (20 × 40 × 40 cm) to prevent them from seeing the experimenter moving the screens and changing the stimuli.

Stimuli. The stimuli were generated by the same procedure as Experiment 1 and then printed on white paper. Each red dot was a circle of a radius of 10 pixels. Stimuli were placed in front of the white plastic screens in such a way that they were visible from the starting area.

Shaping. On the morning of the third day, chicks individually underwent familiarization. In this phase, we used a single screen and stimuli depicting the positive numerosity (i.e., the numerosity that will become associated with food through training). The chick



Figure 9. A schematic representation of the experimental apparatus (left) and two photos (right). The experimental arena had a diamond shape. The starting area and the two screens are positioned in the arena as they were during training and testing sessions. A video illustrating the procedure is available at <https://osf.io/7eavy/>. See the online article for the color version of this figure.

was at first placed within the apparatus, in the space between the starting area and the screen, and it was free to move around and to get acquainted with the novel environment for a couple of minutes. To get the birds used to feeding in the new environment, five mealworms were offered to the chicks before shaping. The food reinforcement was progressively moved behind the screen, requiring the bird to go behind the screen to retrieve and eat the food. This procedure was repeated until the chick circumnavigated the screen three consecutive times. Overall, the shaping phase could last from 10 to 20 min. Chicks that showed little interest in the food reinforcement (i.e., poor mealworm-following behavior), chicks that were too anxious in the new environment, and chicks that were inattentive to the stimuli were discarded. This occurred in about 25% of cases, and such chicks are not included in the study.

Training. Training took place 2 hr after the end of shaping. For 12 subjects, the numerosity associated with food was 20 (each stimulus depicted 20 dots), for the other 12 subjects, the numerosity associated with food was 40. On entering the area, the chick was placed in a starting area behind the transparent partition, which allowed the subject to see the arena and the stimuli for a few seconds before the beginning of each trial. In each trial, a positive and a neutral stimulus were positioned on the left or the right panel. The left–right (L–R) position of the numerosity associated with food was changed from trial to trial according to a semirandom sequence (L–R–L–R–L–L–R–R–L–R– L–R–L–R–L–L–R–R–L–R; Fellows, 1967). The chick remained in the starting area for about 5 s, then the transparent partition was removed, and the chick was free to search for food. Only the first screen circumnavigated was taken into consideration. If it was the one depicting the stimulus associated with food, the response was considered correct. Only by emitting a correct response could chicks gain the food reward (mealworm). Therefore, the only correction on the behavior of the chicks was the possibility to gain the reward. To complete the training phase each chick had to reach a learning criterion: choosing the stimulus associated with food at least 17 times out of 20 valid trials (Rugani, Regolin, & Vallortigara, 2008). On reaching the criterion, training was considered successful, and the chick was placed back in its cage until testing. Overall, depending on the chick's behavior, the training phase could last from 60 to 120 min.

Testing. At the beginning of a testing trial, the chick was placed in the starting area behind the transparent partition, from where it could see the two stimuli. The test phase consisted of a testing session, composed of 30 trials: 20 in-test training trials (tr) and 10 testing trials (T) mixed together (tr,tr,T,tr,tr,T,tr,T,tr,tr,T,tr,T,tr,tr,T,tr,T,tr,T,tr,T,tr,T; Rugani et al., 2016). For the in-test training trials, the previously reinforced pairs of stimuli were used, and chicks received the reward when they emitted a correct response. In the testing trials, we used a new set of stimuli depicting novel numbers of dots. All testing stimuli had 30 dots, with two different interelement distances. Testing trials were never rewarded, therefore lessening any effect of learning during testing.

The animal remained in the starting area for about 5 s, and then the transparent partition was removed. The trial terminated when the bird had placed its head and about three-fourth of its body behind a screen, beyond the side edges (Regolin, Rugani, Pagni, & Vallortigara, 2005; Rosa Salva, Regolin, & Vallortigara, 2012; Rugani et al., 2014). This experiment included an aspect concerning the perception of a visual illusion and one concerning numer-

ical discrimination. Two behavioral responses were considered to analyze both aspects. One measure concerned the first stimulus approached by the chicks. It measures the immediate response to the stimuli and has been previously used in the study of visual illusions in day-old domestic chicks (Rosa Salva, Rugani, Cavazzana, Regolin, & Vallortigara, 2013). The area in the proximity (10 cm) of the two screens was the approached area. This choice was recorded whenever a chick entered the area with its entire body and looked at the stimulus for an interval longer than 3 s. The approach to the screen could sometimes be corrected, and the chick could approach one screen and then move to a different one.

A second behavioral measure concerned the first screen circumnavigated by the chick. This measure has been previously used in a numerical discrimination task (Rugani et al., 2016) in which chicks were trained to discriminate between two numerical stimuli and then to generalize to new stimuli. A choice was recorded when the head of the chick entered the area behind a screen (beyond the side edges). A trial was concluded as soon as the choice had been made.

The chicks were observed from a monitor connected to a video camera so as not to disturb them. All trials were video-recorded, allowing behavior to be scored offline. Online and offline scoring was found to be highly consistent with one other (for approach measure, Cronbach's $\alpha = 1$ and for circumnavigation measures, Cronbach's $\alpha = .99$).

Results

We analyzed the results the same way as in Experiment 1: We tested two possible biases with two independent *t* tests. If participants choose randomly, we expect equal number of responses to the MD = 21 (clustered) and MD = 51 (dispersed) stimuli. The smaller sample size means that we have less power, and this is a limitation.

First, we analyzed performance on the reinforced trials of the testing phase. Chicks trained on 20, but not those trained on 40, approached the stimulus depicting the reinforced stimuli, chicks trained on 20: $M = 68.33$; $t(11) = 7.607$, $p < .001$, Cohen's $D = 2.19$; chicks trained on 40: $M = 54.17$; $ES = 3.79$; $t(11) = 1.101$, $p = .294$, Cohen's $D = 0.31$. With respect to the bias of selecting either the low or high clustering image, we found a preference for high clustering (58.9% vs. 41.1%), $t(23) = 2.516$, $p = .019$, Cohen's $D = 0.51$. With respect to selecting the type of clustering depending on training, there was evidence in support of the prediction based on the illusion, 59.3%: $t(23) = 2.668$, $p = .014$, Cohen's $D = 0.54$. Based on this criterion, therefore, chicks did behave like humans. This pattern is shown in Figure 10, which should be compared with Figure 6 (Experiment 1).

The second analysis was based on the data for circumnavigation behavior. We analyzed performance on the reinforced trials of the testing phase. Chicks continued to circumnavigate the stimulus reinforced at training: (Chicks trained on 20: $M = 79.17$; $t(11) = 12.741$, $p < .001$, Cohen's $D = 3.67$; Chicks trained on 40: $M = 61$; $t(11) = 3.138$, $p = .009$, Cohen's $D = 0.91$). With respect to the bias of selecting either the low or high clustering image, results confirmed a preference for high clustering (55.5% vs. 44.1%), $t(23) = 2.228$, $p = .036$, Cohen's $D = 0.45$. The results are shown in Figure 11. Chicks selected high clustering more often than low clustering (independently of training). With respect to selecting the type of clustering depending on training, the prediction based on

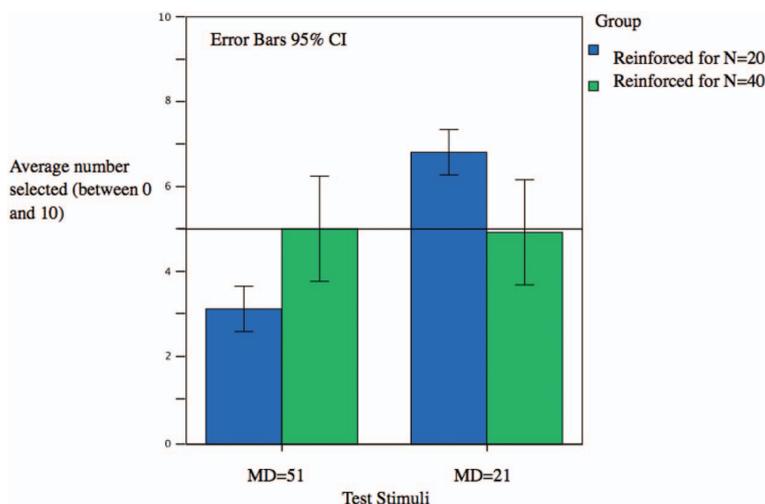


Figure 10. The average number of times that each type of test stimulus was selected based on an analysis of the first approach to the screen. For each participant, this number can be between 0 and 10. Blue and green bars are the two groups with different reinforcement, and therefore the data are from different individuals. See the online article for the color version of this figure.

the illusion was not supported by the data, 50.8%, $t(23) = 0.289$, $p = .775$, Cohen's $D = 0.06$).

The preference for the more clustered pattern (where the MD was less) is clear both in Figure 10 and in Figure 11. This was absent in Experiment 1 for humans. Figure 11 also shows the lack of effect of reinforcement, compared with Figure 10, although the small trend is in the same direction.

Figure 12 shows individual differences for both measures. Overall the higher blue bars relative to green testifies to the preference for the stimuli with more clustering. For the second response (circumnavigation), chicks never systematically chose just one strategy in relation to reinforcement. In particular, there is little

evidence that the chicks were divided in two groups, one using numerosity and one using a different cue, like clustering, and therefore producing two opposite pattern of results.

The overall preference for higher clustering stimuli is interesting. Aside from its theoretical interpretation, it provides evidence that chicks did attend and could discriminate the two types of test stimuli.

General Discussion

Studies on estimation of numerosity have shown that humans and other animals can discriminate the numerosity of sets of visual elements. There is evidence that individuals of many species can

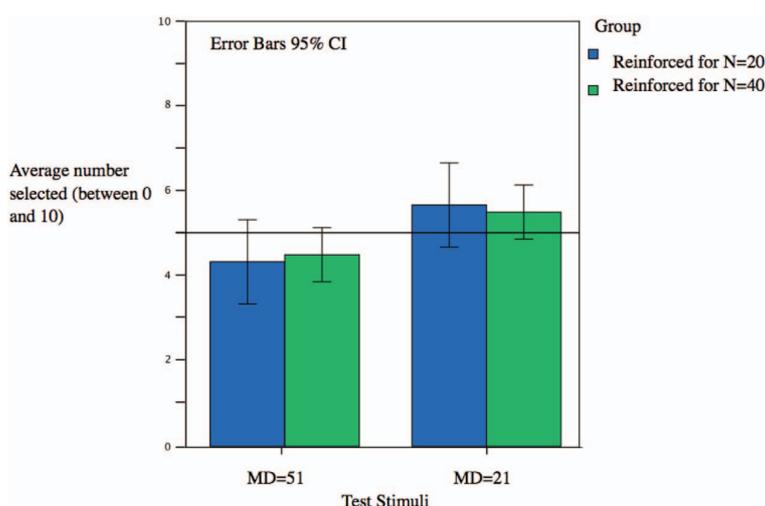


Figure 11. The average number of times that each type of test stimulus was selected in terms of circumnavigation. For each participant this number can be between 0 and 10. See the online article for the color version of this figure.

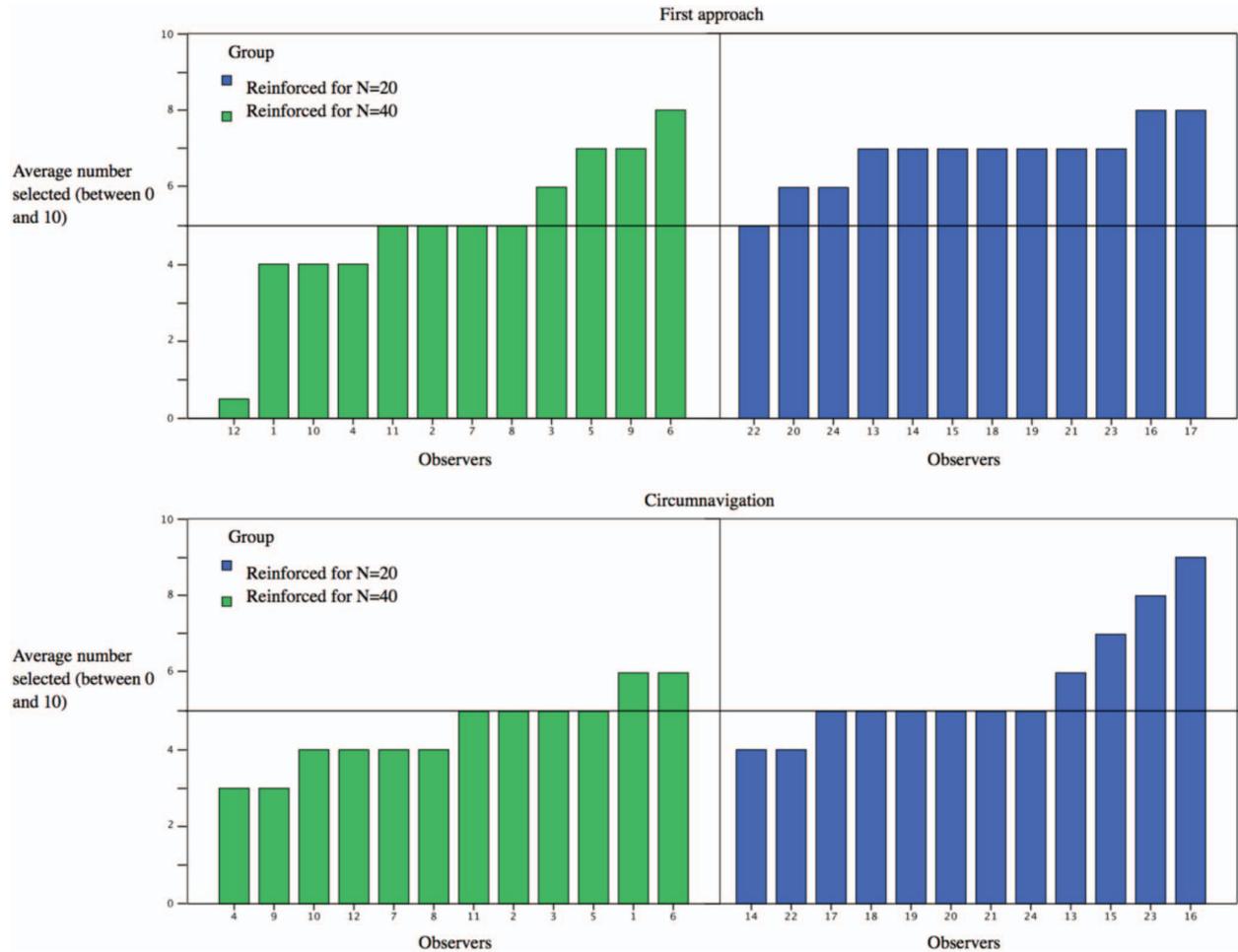


Figure 12. How often the test stimulus predicted by a choice based on numerosity was selected by each observer. Anything above 5 fits the prediction. The two graphs are for the two measures (first approach and circumnavigation). See the online article for the color version of this figure.

compare numbers in simple arithmetic problems. This is true for monkeys (Flombaum, Junge, & Hauser, 2005), chicks (Rugani et al., 2009; Rugani et al., 2008), amphibians (Stancher, Rugani, Regolin, & Vallortigara, 2015), and fish (Agrillo, Piffer, & Bisazza, 2010; Potrich, Sovrano, Stancher, & Vallortigara, 2015). The approximate number system is, therefore, believed to be an ancient system in evolutionary terms (Cantlon, Platt, & Brannon, 2009; Vallortigara, 2014; Versace & Vallortigara, 2015).

We know of several properties of the estimation mechanism, for example, that it does not rely on counting and that the perceived difference is related to the ratio between the two numerosities (Anobile, Cicchini, & Burr, 2014). Estimation can be biased by irrelevant dimensions, such as size (Dakin et al., 2011; Hurewitz et al., 2006) and spatial distribution of the elements (Allik & Tuulmets, 1991; Bertamini et al., 2016). Here we have focused on the role of spacing of the elements and the occupancy model that predicts this effect (Allik & Tuulmets, 1991).

For humans, regular spacing increases perceived numerosity (the regular-random numerosity illusion; Ginsburg, 1980), but this

effect has never been tested in other species, with the exception of Beran (2006). Here we compared humans with domestic chicks. Our stimuli were random configurations that varied in terms of clustering, and we used a procedure as similar as possible for both groups. Chicks are ideal animals to study the generality of the effect because they can be tested after minimal experience of the environment (a few days after hatching) and without social influence (the chicks were kept in isolation).

During training, the configurations had 20 or 40 elements, and density and clustering covaried with numerosity. Either the dimension numerosity or the dimension clustering could be used to perform the task. Test trials presented configurations with 30 elements, but one had a larger interelement distance than the other. Most human observers adopted a strategy based on numerosity and chose the low-clustering test image (when reinforced with 40 elements) or the high-clustering test image (when reinforced with 20 elements). If observers had adopted a strategy based on clustering (e.g., average interelement distance), they would have done

the opposite. In the verbal description of the strategy, they mentioned both numerosity and density, but not clustering.

This clear-cut effect of numerosity and spacing for humans was present in chicks, but results were affected by the behavioral measure. First, chicks could be trained to discriminate configurations with 20 or 40 elements. This confirms that they can discriminate relatively large numerosities. Data for the first approach toward the screen showed an overall bias to select the configuration with more clustering. We discuss this in the next paragraph. However, the first approach behavior was also influenced by training, in a way similar to human adults. There was, therefore, evidence that chicks estimate numerosity using a mechanism similar to humans (and according to the random-regularity illusion). This was not observed when considering the screen the chicks walked behind to find the food. In this measurement, chicks only displayed an overall bias to select the configuration with more clustering. It could be that the more immediate approach response was more affected by perceptual variables, whereas the more considered walking around the panel was more affected by continuous inspection and change of views and distance with respect to the two stimuli.

The finding that chicks prefer overall configurations with high clustering was interesting and can be related to what has been observed before when density was manipulated. Although this was not tested with chicks, new world monkeys (Stevens et al., 2007) and human infants (Uller et al., 2013) show a preference for items that are clustered. This could be explained in terms of foraging theory if we make the assumption that clusters can lower energy expenditure.

More clustering was achieved in our stimuli by a weaker constraint on location. That is, elements could be located close to each other as long as there was no overlap. The configuration with less clustering by contrast was one in which elements had a more constrained and therefore regular distribution (Figure 3). Another possibility, therefore, is that chicks have a preference for more natural distribution of random elements, consistent with a distribution of food in the environment. More constrained and regular distributions may be associated more with textures and less with independent elements, such as seed, on the ground.

With respect to the issue of the role of clustering on perceived numerosity, there is evidence in the literature that animals may sometimes not display the same behavior as humans in tasks that relate to numerosity. As discussed in the introduction, Agrillo et al. (2014) compared humans and monkeys using an illusion and concluded that there may be differences in the visual mechanisms underlying the Solitaire illusion between human and nonhuman primates, but also continuity. It has also been noted, however, how difference in procedures may make comparisons between species difficult (Beran, 2006). By contrast, many aspects of the approximate number system are found across a large range of species. It is therefore possible that the effect of clustering on estimation is intrinsically part of how large numerosities are compared.

In summary, we found evidence that numerosity is a salient property of a configuration. Humans use it to develop a response strategy. The same strategy was present in chicks at least as a first, immediate approach to one of the two stimuli. The main difference between humans and chicks was unrelated to training because overall chicks had a strong preference for images with more

clustering. This bias drove both the first approach and the final circumnavigation behavior.

References

- Adachi, I. (2014). Spontaneous spatial mapping of learned sequence in chimpanzees: Evidence for a SNARC-like effect. *PLoS ONE*, *9*, e90373. <http://dx.doi.org/10.1371/journal.pone.0090373>
- Agrillo, C., Parrish, A. E., & Beran, M. J. (2014). Do primates see the solitaire illusion differently? A comparative assessment of humans (*Homo sapiens*), chimpanzees (*Pan troglodytes*), rhesus monkeys (*Macaca mulatta*), and capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, *128*, 402–413. <http://dx.doi.org/10.1037/a0037499>
- Agrillo, C., Piffer, L., & Bisazza, A. (2010). Large number discrimination by mosquitofish. *PLoS ONE*, *5*, e15232. <http://dx.doi.org/10.1371/journal.pone.0015232>
- Allik, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception and Psychophysics*, *49*, 303–314. <http://dx.doi.org/10.3758/BF03205986>
- Anobile, G., Cicchini, G. M., & Burr, D. C. (2014). Separate mechanisms for perception of numerosity and density. *Psychological Science*, *25*, 265–270. <http://dx.doi.org/10.1177/0956797613501520>
- Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. C. (2015). Mechanisms for perception of numerosity or texture-density are governed by crowding-like effects. *Journal of Vision*, *15*, 4. <http://dx.doi.org/10.1167/15.5.4>
- Beran, M. J. (2006). Quantity perception by adult humans (*Homo sapiens*), chimpanzees (*Pan troglodytes*), and rhesus macaques (*Macaca mulatta*) as a function of stimulus organization. *International Journal of Comparative Psychology*, *19*, 386–397.
- Bertamini, M., Zito, M., Scott-Samuel, N. E., & Hulleman, J. (2016). Spatial clustering and its effect on perceived clustering, numerosity, and dispersion. *Attention, Perception and Psychophysics*, *78*, 1460–1471. <http://dx.doi.org/10.3758/s13414-016-1100-0>
- Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science*, *282*, 746–749. <http://dx.doi.org/10.1126/science.282.5389.746>
- Burgess, A., & Barlow, H. B. (1983). The precision of numerosity discrimination in arrays of random dots. *Vision Research*, *23*, 811–820. [http://dx.doi.org/10.1016/0042-6989\(83\)90204-3](http://dx.doi.org/10.1016/0042-6989(83)90204-3)
- Burr, D., & Ross, J. (2008). A visual sense of number. *Current Biology*, *18*, 425–428. <http://dx.doi.org/10.1016/j.cub.2008.02.052>
- Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, *17*, 401–406. <http://dx.doi.org/10.1111/j.1467-9280.2006.01719.x>
- Cantlon, J. F., & Brannon, E. M. (2007). How much does number matter to a monkey (*Macaca mulatta*)? *Journal of Experimental Psychology: Animal Behavior Processes*, *33*, 32–41. <http://dx.doi.org/10.1037/0097-7403.33.1.32>
- Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, *13*, 83–91. <http://dx.doi.org/10.1016/j.tics.2008.11.007>
- Cicchini, G. M., Anobile, G., & Burr, D. C. (2016). Spontaneous perception of numerosity in humans. *Nature Communications*, *7*, 12536. <http://dx.doi.org/10.1038/ncomms12536>
- Cordes, S., & Brannon, E. M. (2009). Crossing the divide: Infants discriminate small from large numerosities. *Developmental Psychology*, *45*, 1583–1594. <http://dx.doi.org/10.1037/a0015666>
- 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 and Review*, *8*, 698–707. <http://dx.doi.org/10.3758/BF03196206>
- Cousins, J. B., & Ginsburg, N. (1983). Subjective correlation and the regular-random numerosity illusion. *Journal of General Psychology*, *108*, 3–10. <http://dx.doi.org/10.1080/00221309.1983.9711472>

- Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 19552–19557. <http://dx.doi.org/10.1073/pnas.1113195108>
- Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42. [http://dx.doi.org/10.1016/0010-0277\(92\)90049-N](http://dx.doi.org/10.1016/0010-0277(92)90049-N)
- Dehaene, S. (2003). The neural basis of the Weber-Fechner law: A logarithmic mental number line. *Trends in Cognitive Sciences*, *7*, 145–147. [http://dx.doi.org/10.1016/S1364-6613\(03\)00055-X](http://dx.doi.org/10.1016/S1364-6613(03)00055-X)
- Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. New York, NY: Oxford University Press.
- Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience*, *5*, 390–407. <http://dx.doi.org/10.1162/jocn.1993.5.4.390>
- de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., & Streri, A. (2014). Representations of space, time, and number in neonates. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 4809–4813. <http://dx.doi.org/10.1073/pnas.1323628111>
- Ditz, H. M., & Nieder, A. (2015). Neurons selective to the number of visual items in the corvid songbird endbrain. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 7827–7832. <http://dx.doi.org/10.1073/pnas.1504245112>
- Eger, E., Michel, V., Thirion, B., Amadon, A., Dehaene, S., & Kleinschmidt, A. (2009). Deciphering cortical number coding from human brain activity patterns. *Current Biology*, *19*, 1608–1615. <http://dx.doi.org/10.1016/j.cub.2009.08.047>
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314. <http://dx.doi.org/10.1016/j.tics.2004.05.002>
- Ferrigno, S., & Cantlon, J. F. (2017). Evolutionary constraints on the emergence of human mathematical concepts. In J. Kaas (Ed.), *Evolution of nervous systems 2e* (Vol. 3, pp. 511–521). Oxford: Elsevier. <http://dx.doi.org/10.1016/B978-0-12-804042-3.00099-3>
- Fellows, B. J. (1967). Chance stimulus sequences for discrimination tasks. *Psychological Bulletin*, *67*, 87–92.
- Flombaum, J. I., Junge, J. A., & Hauser, M. D. (2005). Rhesus monkeys (*Macaca mulatta*) spontaneously compute addition operations over large numbers. *Cognition*, *97*, 315–325. <http://dx.doi.org/10.1016/j.cognition.2004.09.004>
- Frith, C. D., & Frith, U. (1972). The solitaire illusion: An illusion of numerosity. *Perception and Psychophysics*, *11*, 409–410. <http://dx.doi.org/10.3758/BF0320627>
- Gebuis, T., Kadosh, R. C., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review. *Acta Psychologica*, *171*, 17–35. <http://dx.doi.org/10.1016/j.actpsy.2016.09.003>
- Ginsburg, N. (1976). Effect of item arrangement on perceived numerosity: Randomness vs regularity. *Perceptual and Motor Skills*, *43*, 663–668. <http://dx.doi.org/10.2466/pms.1976.43.2.663>
- Ginsburg, N. (1980). The regular-random numerosity illusion: Rectangular patterns. *Journal of General Psychology*, *103*, 211–216. <http://dx.doi.org/10.1080/00221309.1980.9921000>
- Ginsburg, N. (1991). Numerosity estimation as a function of stimulus organization. *Perception*, *20*, 681–686. <http://dx.doi.org/10.1068/p200681>
- Ginsburg, N., & Nicholls, A. (1988). Perceived numerosity as a function of item size. *Perceptual and Motor Skills*, *67*, 656–658. <http://dx.doi.org/10.2466/pms.1988.67.2.656>
- 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. <http://dx.doi.org/10.1037/a0012682>
- Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 19599–19604. <http://dx.doi.org/10.1073/pnas.0609485103>
- Inglis, M., & Gilmore, C. (2013). Sampling from the mental number line: How are approximate number system representations formed? *Cognition*, *129*, 63–69. <http://dx.doi.org/10.1016/j.cognition.2013.06.003>
- Jordan, K. E., & Brannon, E. M. (2006). Weber’s Law influences numerical representations in rhesus macaques (*Macaca mulatta*). *Animal Cognition*, *9*, 159–172. <http://dx.doi.org/10.1007/s10071-006-0017-8>
- MacDonald, S. E., & Agnes, M. M. (1999). Orangutan (*Pongo pygmaeus abelii*) spatial memory and behavior in a foraging task. *Journal of Comparative Psychology*, *113*, 213–217. <http://dx.doi.org/10.1037/0735-7036.113.2.213>
- Mitchell, K. G., Calton, J. L., Threlkeld, R. C., & Schachtman, T. R. (1996). Attenuation and reacquisition of foraging behavior in a patchy environment. *Behavioural Processes*, *36*, 239–252. [http://dx.doi.org/10.1016/0376-6357\(95\)00034-8](http://dx.doi.org/10.1016/0376-6357(95)00034-8)
- Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*, 1519–1520. <http://dx.doi.org/10.1038/2151519a0>
- Nieder, A., & Merten, K. (2007). A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *The Journal of Neuroscience*, *27*, 5986–5993. <http://dx.doi.org/10.1523/JNEUROSCI.1056-07.2007>
- Parrish, A. E., Agrillo, C., Perdue, B. M., & Beran, M. J. (2016). The elusive illusion: Do children (*Homo sapiens*) and capuchin monkeys (*Cebus apella*) see the Solitaire illusion? *Journal of Experimental Child Psychology*, *142*, 83–95. <http://dx.doi.org/10.1016/j.jecp.2015.09.021>
- Parrish, A. E., James, B. T., & Beran, M. J. (2017). Exploring whether nonhuman primates show a bias to overestimate dense quantities. *Journal of Comparative Psychology*, *131*, 59–68. <http://dx.doi.org/10.1037/com0000058>
- Pearson, J. M., Hayden, B. Y., Raghavachari, S., & Platt, M. L. (2009). Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. *Current Biology*, *19*, 1532–1537. <http://dx.doi.org/10.1016/j.cub.2009.07.048>
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13. <http://dx.doi.org/10.1016/j.jneumeth.2006.11.017>
- Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, *11*, 1129–1135. <http://dx.doi.org/10.1038/nn.2187>
- Piaget, J. (1952). *The origins of intelligence in children*. New York, NY: Norton & Co. <http://dx.doi.org/10.1037/11494-000>
- 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. <http://dx.doi.org/10.1016/j.neuron.2004.10.014>
- Ponzo, M. (1928). Urteilstauschungen über Mengen [Misperceptions of quantities]. *Archiv für die Gesamte Psychologie*, *65*, 129–162.
- Potrich, D., Sovrano, V. A., Stancher, G., & Vallortigara, G. (2015). Quantity discrimination by zebrafish (*Danio rerio*). *Journal of Comparative Psychology*, *129*, 388–393. <http://dx.doi.org/10.1037/com0000012>
- Regolin, L., Rugani, R., Pagni, P., & Vallortigara, G. (2005). Delayed search for a social and a non-social goal object by the young domestic chick (*Gallus gallus*). *Animal Behaviour*, *70*, 855–864. <http://dx.doi.org/10.1016/j.anbehav.2005.01.014>
- Rosa Salva, O., Regolin, L., & Vallortigara, G. (2012). Inversion of contrast polarity abolishes spontaneous preferences for face-like stimuli in newborn chicks. *Behavioural Brain Research*, *228*, 133–143. <http://dx.doi.org/10.1016/j.bbr.2011.11.025>
- Rosa Salva, O., Rugani, R., Cavazzana, A., Regolin, L., & Vallortigara, G. (2013). Perception of the Ebbinghaus illusion in four-day-old domestic

- chicks (*Gallus gallus*). *Animal Cognition*, 16, 895–906. <http://dx.doi.org/10.1007/s10071-013-0622-2>
- Rugani, R., Cavazzana, A., Vallortigara, G., & Regolin, L. (2013). One, two, three, four, or is there something more? Numerical discrimination in day-old domestic chicks. *Animal Cognition*, 16, 557–564. <http://dx.doi.org/10.1007/s10071-012-0593-8>
- Rugani, R., Fontanari, L., Simoni, E., Regolin, L., & Vallortigara, G. (2009). Arithmetic in newborn chicks. *Proceedings Biological Sciences*, 276, 2451–2460. <http://dx.doi.org/10.1098/rspb.2009.0044>
- Rugani, R., McCrink, K., de Hevia, M. D., Vallortigara, G., & Regolin, L. (2016). Ratio abstraction over discrete magnitudes by newly hatched domestic chicks (*Gallus gallus*). *Scientific Reports*, 6, 30114. <http://dx.doi.org/10.1038/srep30114>
- Rugani, R., Regolin, L., & Vallortigara, G. (2008). Discrimination of small numerosities in young chicks. *Journal of Experimental Psychology: Animal Behavior Processes*, 34, 388–399. <http://dx.doi.org/10.1037/0097-7403.34.3.388>
- Rugani, R., Vallortigara, G., & Regolin, L. (2013). Numerical abstraction in young domestic chicks (*Gallus gallus*). *PLoS ONE*, 8, e65262. <http://dx.doi.org/10.1371/journal.pone.0065262>
- Rugani, R., Vallortigara, G., & Regolin, L. (2014). From small to large: Numerical discrimination by young domestic chicks (*Gallus gallus*). *Journal of Comparative Psychology*, 128, 163–171. <http://dx.doi.org/10.1037/a0034513>
- Scarf, D., Hayne, H., & Colombo, M. (2011). Pigeons on par with primates in numerical competence. *Science*, 334, 1664. <http://dx.doi.org/10.1126/science.1213357>
- Stancher, G., Rugani, R., Regolin, L., & Vallortigara, G. (2015). Numerical discrimination by frogs (*Bombina orientalis*). *Animal Cognition*, 18, 219–229. <http://dx.doi.org/10.1007/s10071-014-0791-7>
- Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ: Princeton University Press.
- Stevens, J. R., Wood, J. N., & Hauser, M. D. (2007). When quantity trumps number: Discrimination experiments in cotton-top tamarins (*Saguinus oedipus*) and common marmosets (*Callithrix jacchus*). *Animal Cognition*, 10, 429–437. <http://dx.doi.org/10.1007/s10071-007-0081-8>
- Tibber, M. S., Greenwood, J. A., & Dakin, S. C. (2012). Number and density discrimination rely on a common metric: Similar psychophysical effects of size, contrast, and divided attention. *Journal of Vision*, 12, 8. <http://dx.doi.org/10.1167/12.6.8>
- 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, 1839–1853. <http://dx.doi.org/10.3758/APP.72.7.1839>
- Uller, C., Urquhart, C., Lewis, J., & Berntsen, M. (2013). Ten-month-old infants' reaching choices for "more": The relationship between inter-stimulus distance and number. *Frontiers in Psychology*, 4, 84. <http://dx.doi.org/10.3389/fpsyg.2013.00084>
- Vallortigara, G. (2014). Foundations of number and space representations in non-human species. In D. C. Geary, D. B. Beach, & K. Mann Koepeke (Eds.), *Evolutionary origins and early development of number processing* (pp. 35–66). New York, NY: Elsevier.
- Vallortigara, G. (2017a). An animal's sense of number. In J. W. Adams, P. Barmby, & A. Mesoudi (Eds.), *The nature and development of mathematics: Cross disciplinary perspective on cognition, learning and culture* (pp. 43–65). New York, NY: Routledge.
- Vallortigara, G. (2017b). Comparative cognition of number and space: The case of geometry and of the mental number line. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 373, 20170120. <http://dx.doi.org/10.1098/rstb.2017.0120>
- Valsecchi, M., Toscani, M., & Gegenfurtner, K. R. (2013). Perceived numerosity is reduced in peripheral vision. *Journal of Vision*, 13, 7. <http://dx.doi.org/10.1167/13.13.7>
- Versace, E., & Vallortigara, G. (2015). Origins of knowledge: Insights from prosocial species. *Frontiers in Behavioral Neuroscience*, 9, 338. <http://dx.doi.org/10.3389/fnbeh.2015.00338>
- Vicario, G. B. (2011). *Illusioni ottico-geometriche: Una rassegna di problemi*. Venezia, Italy: Istituto Veneto di Scienze Lettere e Arti.
- Vos, P. G., van Oeffelen, M. P., Tibbosch, H. J., & Allik, J. (1988). Interactions between area and numerosity. *Psychological Research*, 50, 148–154. <http://dx.doi.org/10.1007/BF00310175>

Received July 5, 2017

Revision received January 6, 2018

Accepted January 15, 2018 ■