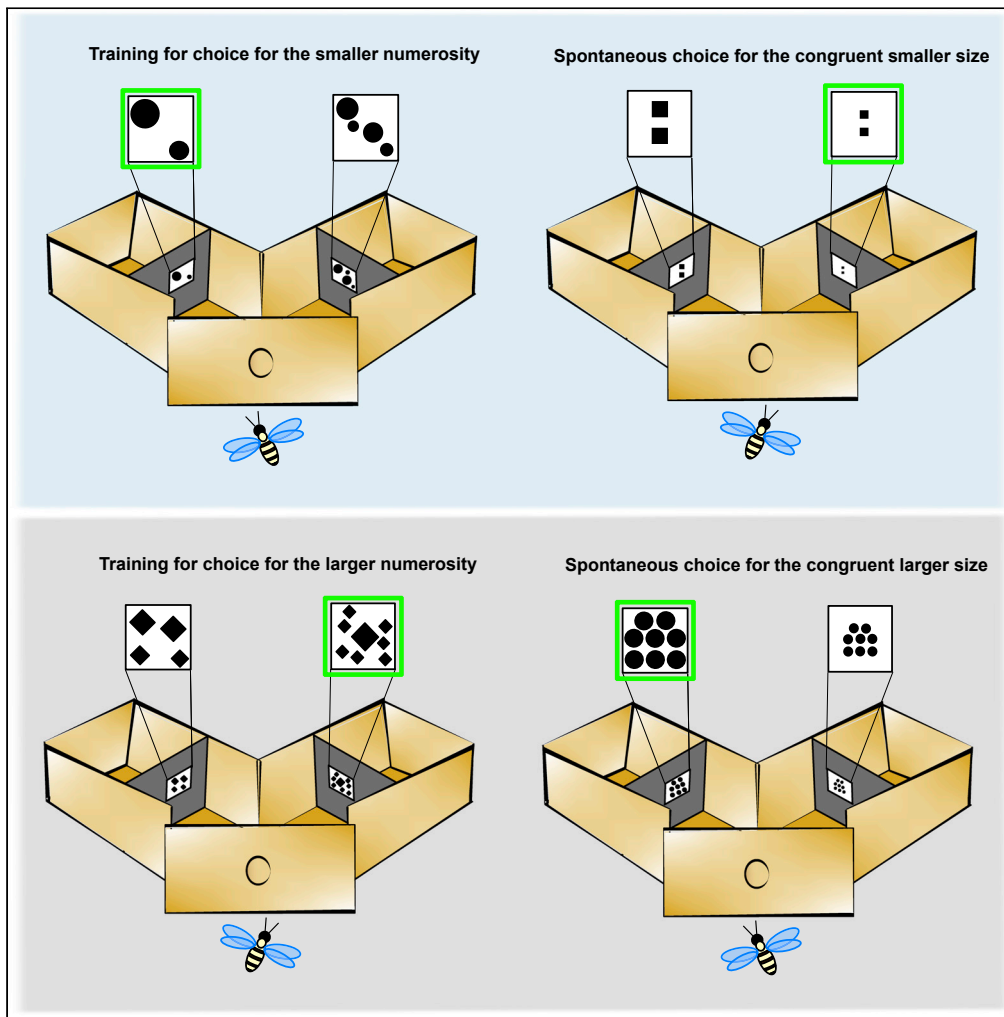


Article

Transfer from Number to Size Reveals Abstract Coding of Magnitude in Honeybees



Maria Bortot,
Gionata Stancher,
Giorgio
Vallortigara

maria.bortot@unitn.it (M.B.)
giorgio.vallortigara@unitn.it
(G.V.)

HIGHLIGHTS

Honeybees exhibit a cross-dimensional transfer from discrete to continuous dimensions

Honeybees trained to choose the smaller numerosity, choose the congruent smaller size

Honeybees trained to choose the larger numerosity, choose the congruent larger size

This supports the hypothesis of a cognitive universality of a coding for magnitudes

Bortot et al., iScience 23,
101122
May 22, 2020 © 2020 The
Authors.
[https://doi.org/10.1016/
j.isci.2020.101122](https://doi.org/10.1016/j.isci.2020.101122)

Article

Transfer from Number to Size Reveals Abstract Coding of Magnitude in Honeybees

Maria Bortot,^{1,3,*} Gionata Stancher,² and Giorgio Vallortigara^{1,*}

SUMMARY

Number discrimination has been documented in honeybees. It is not known, however, whether it reflects, as in vertebrates, the operating of an underlying general magnitude system that estimates quantities irrespective of dimensions (e.g., number, space, time) and format (discrete, continuous). We trained bees to discriminate between different numerical comparisons having either a 0.5 (2 versus 4; 4 versus 8) or 0.67 ratio (2 versus 3; 4 versus 6). Bees were then tested for spontaneous choice using comparisons with identical numbers but different sizes. Irrespective of the ratio of stimuli, bees trained to select the smaller numerical quantity chose the congruent smaller size; bees trained to select the larger numerical quantity chose the congruent larger size. This finding provides the evidence for a cross-dimensional transfer between discrete (numerical) and continuous (spatial) dimensions in an invertebrate species and supports the hypothesis of a cognitive universality of a coding for general magnitude.

INTRODUCTION

Honeybees (*Apis mellifera*) have been shown to be able to process the numerical attributes of visual stimuli (Chittka and Geiger, 1995; Dacke and Srinivasan, 2008; Bortot et al., 2019; Howard et al., 2019), including the zero as a quantity (Howard et al., 2018). Given their distant phylogenetic origins, it is unclear, however, to what extent bees (invertebrates) and vertebrates share similarities in number cognition.

Humans and others non-human vertebrates make use of a nonverbal, nonsymbolic representation of number, the so-called Approximate Number System (ANS). The ANS obeys Weber's law—it is thus mainly limited by the ratio between the numerical values being compared—and is thought to be supported by an evolutionarily ancient mechanism for representing quantity in an analog fashion. Gallistel (1989) first argued that discrete countable quantity (i.e., number) and continuous quantity (e.g., space and time) must be represented by a common mental currency to enable animals to perform arithmetic operations across domains (as in the case of the rate of return to a food patch that can be computed only if organisms represent time and number in a single currency). According to this hypothesis, quantity representations in the various domains (i.e., number, space and time) would be processed by a «common magnitude system», which represents these dimensions via the same unit of magnitude (Gallistel, 1989). Evidence that the temporal, spatial, and numerical features of a stimulus can interact with one another has been provided for vertebrates such as monkeys (Merritt et al., 2010) and birds (De Corte et al., 2017) and for prelinguistic human babies (Di Giorgio et al., 2019).

Interestingly, honeybees have been shown to exhibit the numerical distance effect (i.e., the fact that the ability to discriminate between numbers improves as the numerical distance increases, (e.g., 0 versus 4 is easier than 0 versus 1, Howard et al., 2018). The numerical distance effect is one of the signatures of the ANS and suggests the existence of an analog magnitude system in honeybees that would allow the processing of different numbers. Moreover, honeybees have been shown to be able to process the relative size of visual stimuli (Avarguès-Weber et al., 2014), but it is not known whether in bees a common set of coding mechanisms underlies quantity estimation in different domains.

Here we investigated whether honeybees could make a transfer from discrete (number) to continuous (size) magnitudes. Bees were trained to discriminate between different numerical comparisons having either a 0.5 ratio (2 versus 4 and 4 versus 8) or 0.67 ratio (2 versus 3 and 4 versus 6). Half of the subjects

¹Centre for Mind/Brain Sciences, University of Trento, 38068 Rovereto, Italy

²Rovereto Civic Museum Foundation, 38068 Rovereto, Italy

³Lead Contact

*Correspondence: maria.bortot@unitn.it (M.B.), giorgio.vallortigara@unitn.it (G.V.)

<https://doi.org/10.1016/j.isci.2020.101122>



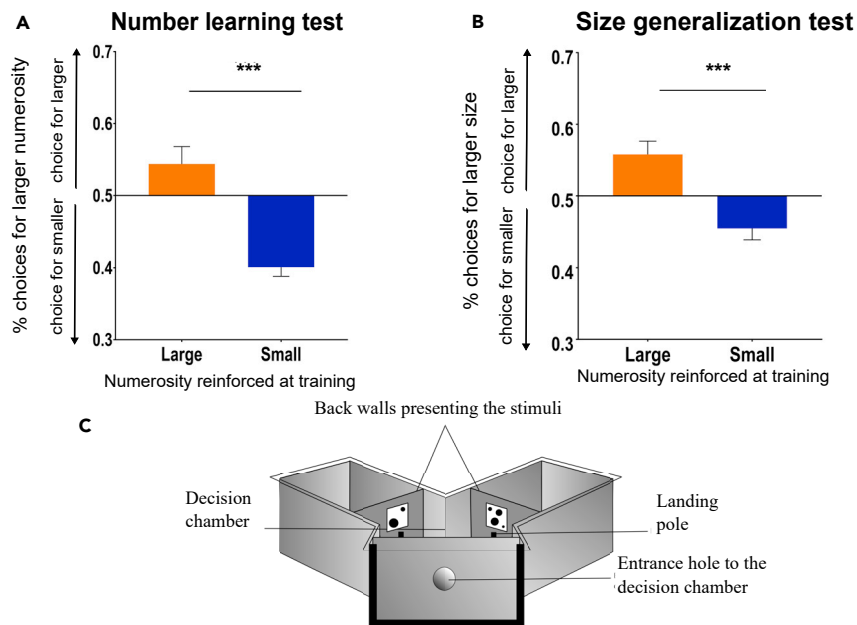


Figure 1. Results of the Test Phase and Representation of the Apparatus

(A) In the *number learning* test, honeybees trained to discriminate the smaller or the larger numerosity showed correct spontaneous choices in the absence of reward (mean \pm SEM; *** p < 0.001, Analysis of variance (ANOVA)).

(B) In the *size transfer* test, bees previously trained to select the larger numerosity showed a preference for the larger size, conversely bees previously trained to select the smaller numerosity showed a preference for the smaller size (mean \pm SEM; *** p < 0.001, Analysis of variance (ANOVA)).

(C) Schematic representation of the Y-maze used to train bees to discriminate numerosity and to test them for transfer from numerical to spatial (size) dimensions.

learnt to choose the smaller quantity and the other half the larger quantity. Then at test, bees were presented with stimuli of different size but identical numerosity under extinction condition (i.e., in the absence of reward). If bees possess a common mechanism to process different magnitudes, then animals trained to choose the smaller/larger quantity in the number comparisons were expected to choose the congruent smaller/larger size in the size comparison. Moreover, choice of the congruent size would not be affected by the ratio of the stimuli (i.e., ratios that proved to be discriminable for numbers should prove discriminable for sizes as well).

RESULTS

The results are shown in [Figures 1A](#) and [1B](#) as proportions of choices for the larger magnitude. An analysis of variance revealed a significant main effect of the type of training (smaller versus larger numerosity as positive; $F_{(1, 48)} = 45.8$, $p < 0.001$, $\omega^2 = 0.399$) but not of the type of test (*number learning* test versus *size generalization* test; $F_{(1, 48)} = 3.5$, $p = 0.066$) and the ratio (0.5 versus 0.67; $F_{(1, 48)} = 0.3$, $p > 0.05$).

No significant interactions were observed (ratio \times numerical comparisons: $F_{(2, 48)} = 1.9$, $p > 0.05$; ratio \times type of training: $F_{(1, 48)} = 1.7$, $p > 0.05$; ratio \times type of test: $F_{(1, 48)} = 0.5$, $p > 0.05$; type of training \times type of test: $F_{(1, 48)} = 1.2$, $p > 0.05$; ratio \times type of training \times type of test: $F_{(1, 48)} = 0.09$, $p > 0.05$; ratio \times type of training \times numerical comparisons: $F_{(2, 48)} = 1.2$, $p > 0.05$; ratio \times type of test \times numerical comparisons: $F_{(2, 48)} = 1.9$, $p > 0.05$; ratio \times type of training \times type of test \times numerical comparisons: $F_{(2, 48)} = 0.03$, $p > 0.05$).

As can be seen in [Figure 1](#), bees trained to select the larger/smaller numerosity ([Figure 1A](#); choice for correct, either larger or smaller, numerosity: $57.16\% \pm 0.01$, mean% \pm SEM; two-tailed one-sample t test: $t_{(31)} = 5.02$, $p < 0.001$, $d = 0.89$) chose the congruent larger/smaller size ([Figure 1B](#); choice for congruent, either larger or smaller, size: $55.14\% \pm 0.01$, mean% \pm SEM; two-tailed one-sample t test: $t_{(31)} = 4.26$, $p < 0.001$, $d = 0.75$).

DISCUSSION

Results of *number learning* test confirmed previous studies (Bortot et al., 2019; Howard et al., 2018, 2019) showing that bees can discriminate numerosities with 0.5 and 0.67 ratios when continuous physical variables were controlled for. Moreover, we found that honeybees can make a transfer from discrete (number) to continuous (size) magnitudes. This provides the evidence for a common code for magnitudes in an invertebrate species.

The hypothesis of the existence of a prelinguistic framework to process different prothetic dimensions (i.e., dimensions that can be "more" or "less" than) was first proposed by Gallistel (Gallistel, 1989) and then developed by Walsh (Walsh, 2003). Research in humans and other vertebrates has revealed that the temporal, spatial, and numerical features of a stimulus can interact with one another (Merritt et al., 2010; Gallistel and Gelman, 2000; Rugani et al., 2015; Lourenco and Longo, 2010) and evidence of similar activation in the parietal cortex in humans and non-human primates in quantity discrimination seems to support the hypothesis of an encoding by a common magnitude (Piazza et al., 2007). Our results show that bees generalize from a numerical dimension to a spatial (size) dimension, suggesting that a general magnitude encoding can be shared among vertebrates and invertebrates.

Our experiment demonstrates that, despite their small brains and the absence of a cortex, honeybees generalize from discrete (number) to continuous (size) magnitudes. It could be that there are regions in the honeybee's brain that play a similar function to the mammalian parietal cortex or to the avian nidopallium caudolateral (Ditz and Nieder, 2020). The central neuronal structure of the bee's brain, the mushroom bodies, a very complex high-order integration center, seems to be crucial in mediating the ability of bees to perform conceptual learning and extraction of relational rules (Giurfa, 2013). Whether mushroom bodies are supporting the cognitive mechanism of processing magnitudes is unknown.

Future studies should investigate whether the ability of insects to generalize between number and space is widespread to other magnitudes such as, for instance, time or brightness.

Limitations of the Study

This study investigated the association between spatial and numerical dimensions. However, the relationship between other magnitudes should be investigated in order to strongly support the existence of a common processing mechanism of magnitude in invertebrate species.

METHODS

All methods can be found in the accompanying [Transparent Methods supplemental file](#).

RESOURCE AVAILABILITY

Lead Contact

Maria Bortot, maria.bortot@unitn.it.

Materials Availability

This study did not generate new unique Materials.

Data and Code Availability

Original data have been deposited to Mendeley Data: <https://doi.org/10.17632/dgwbncbs58.1>.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.isci.2020.101122>.

ACKNOWLEDGMENTS

We thank Alvis Kalarikkan for help with data collection. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 833504, ERC Advanced Grant SPANUMBRA to G.V.).

AUTHOR CONTRIBUTIONS

M.B. and G.V. designed research; M.B. performed the experiments; M.B. analyzed the data; G.S. contributed materials, animals, and space; M.B. and G.V. wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 31, 2020

Revised: March 12, 2020

Accepted: April 29, 2020

Published: May 22, 2020

REFERENCES

- Avarguès-Weber, A., d Amaro, D., Metzler, M., and Dyer, A.G. (2014). Conceptualization of relative size by honeybees. *Front. Behav. Neurosci.* *8*, 80.
- Bortot, M., Agrillo, C., Avarguès-Weber, A., Bisazza, A., MilettoPetrazzini, M.E., and Giurfa, M. (2019). Honeybees use absolute rather than relative numerosity in number discrimination. *Biol. Lett.* *15*, 20190138.
- Chittka, L., and Geiger, K. (1995). Can honey bees count landmarks? *Anim. Behav.* *49*, 159–164.
- Dacke, M., and Srinivasan, M.V. (2008). Evidence for counting in insects. *Anim. Cogn.* *11*, 683–689.
- De Corte, B.J., Navarro, V.M., and Wasserman, E.A. (2017). Non-cortical magnitude coding of space and time by pigeons. *Curr. Biol.* *27*, R1264–R1265.
- Di Giorgio, E., Lunghi, M., Rugani, R., Regolin, L., DallaBarba, B., Vallortigara, G., and Simion, F. (2019). A mental number line in human newborns. *Dev. Sci.* *22*, e12801.
- Ditz, H.M., and Nieder, A. (2020). Format-dependent and format-independent representation of sequential and simultaneous numerosity in the crow endbrain. *Nat. Commun.* *11*, 1–10.
- Gallistel, C.R., and Gelman, R. (2000). Non-verbal numerical cognition: from reals to integers. *Trends Cogn. Sci.* *4*, 59–65.
- Gallistel, C.R. (1989). Animal cognition: the representation of space, time and number. *Annu. Rev. Psychol.* *40*, 155–189.
- Giurfa, M. (2013). Cognition with few neurons: higher-order learning in insects. *Trends Neurosci.* *36*, 285–294.
- Howard, S.R., Avarguès-Weber, A., Garcia, J.E., Greentree, A.D., and Dyer, A.G. (2019). Numerical cognition in honeybees enables addition and subtraction. *Sci. Adv.* *5*, eaav0961.
- Howard, S.R., Avarguès-Weber, A., Garcia, J.E., Greentree, A.D., and Dyer, A.G. (2018). Numerical ordering of zero in honey bees. *Science* *360*, 1124–1126.
- Lourenco, S.F., and Longo, M.R. (2010). General magnitude representation in human infants. *Psychol. Sci.* *21*, 873–881.
- Merritt, D.J., Casasanto, D., and Brannon, E.M. (2010). Do monkeys think in metaphors? Representations of space and time in monkeys and humans. *Cognition* *117*, 191–202.
- Piazza, M., Pinel, P., Le Bihan, D., and Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron* *53*, 293–305.
- Rugani, R., Vallortigara, G., Pifftis, K., and Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' mental number line. *Science* *347*, 534–536.
- Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* *7*, 483–488.

iScience, Volume 23

Supplemental Information

Transfer from Number to Size Reveals

Abstract Coding of Magnitude in Honeybees

Maria Bortot, Gionata Stancher, and Giorgio Vallortigara

SUPPLEMENTAL ITEMS

FIGURES

Figure S1. Numerical contrast 2 vs. 3, Related to Figure 1a.

The 90 stimuli pairs displaying the 2 vs. 3 numerical contrast. Stimuli are divided by shape categories: circles, diamond and squares. Within each shape, in one quarter of the pairs the cumulative surface area was matched to 100% whereas in the second quarter it was not controlled. In the third and the fourth quarter the contour length was matched to 100% and not controlled, respectively. Moreover, within each shape, half of the stimuli was controlled for the convex hull and the other half for the density of the elements.

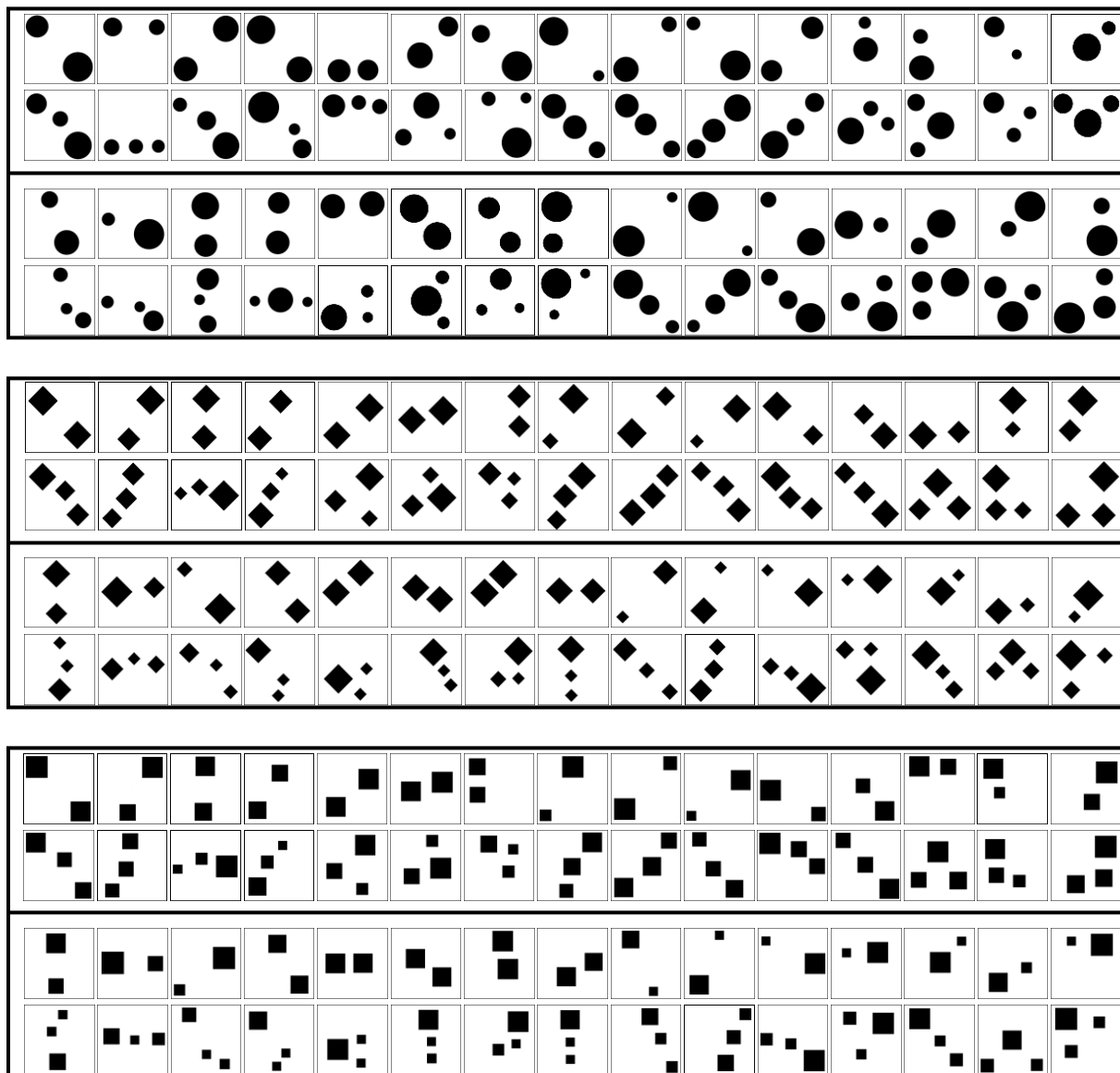


Figure S2. Numerical contrast 2 vs. 4, Related to Figure 1a.

The 90 stimuli pairs displaying the 2 vs. 4 numerical contrast. Stimuli are divided by shape categories: circles, diamond and squares. Within each shape, in one quarter of the pairs the cumulative surface area was matched to 100% whereas in the second quarter it was not controlled. In the third and the fourth quarter the contour length was matched to 100% and not controlled, respectively. Moreover, within each shape, half of the stimuli was controlled for the convex hull and the other half for the density of the elements.

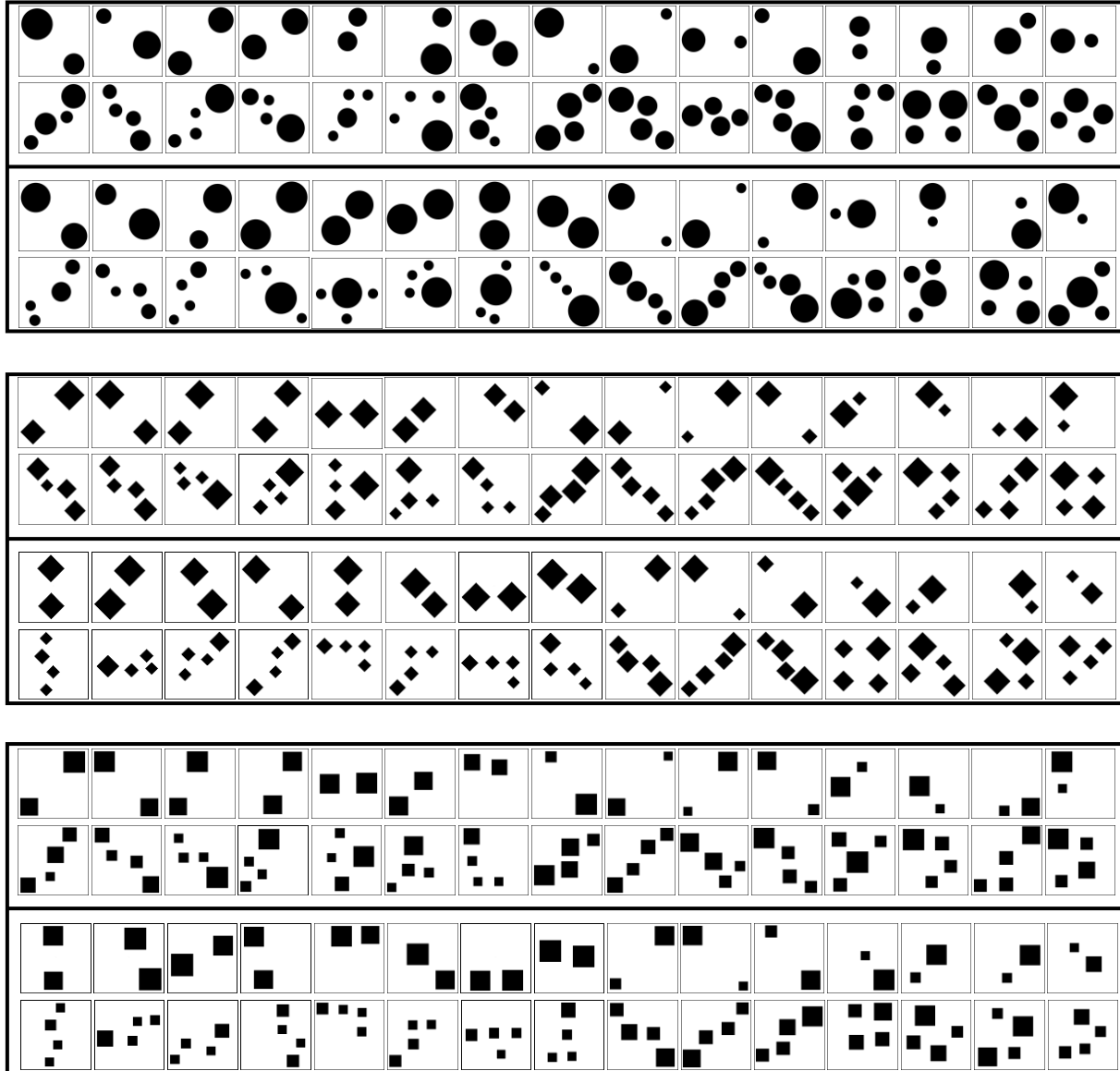


Figure S3. Numerical contrast 4 vs. 6, Related to Figure 1a.

The 90 stimuli pairs displaying the 4 vs. 6 numerical contrast. Stimuli are divided by shape categories: circles, diamond and squares. Within each shape, in one quarter of the stimuli the cumulative surface area was matched to 100% whereas in the second quarter it was not controlled. In the third and the fourth quarter the contour length was matched to 100% and not controlled, respectively. Moreover, within each shape, half of the stimuli was controlled for the convex hull and the other half for the density of the elements.

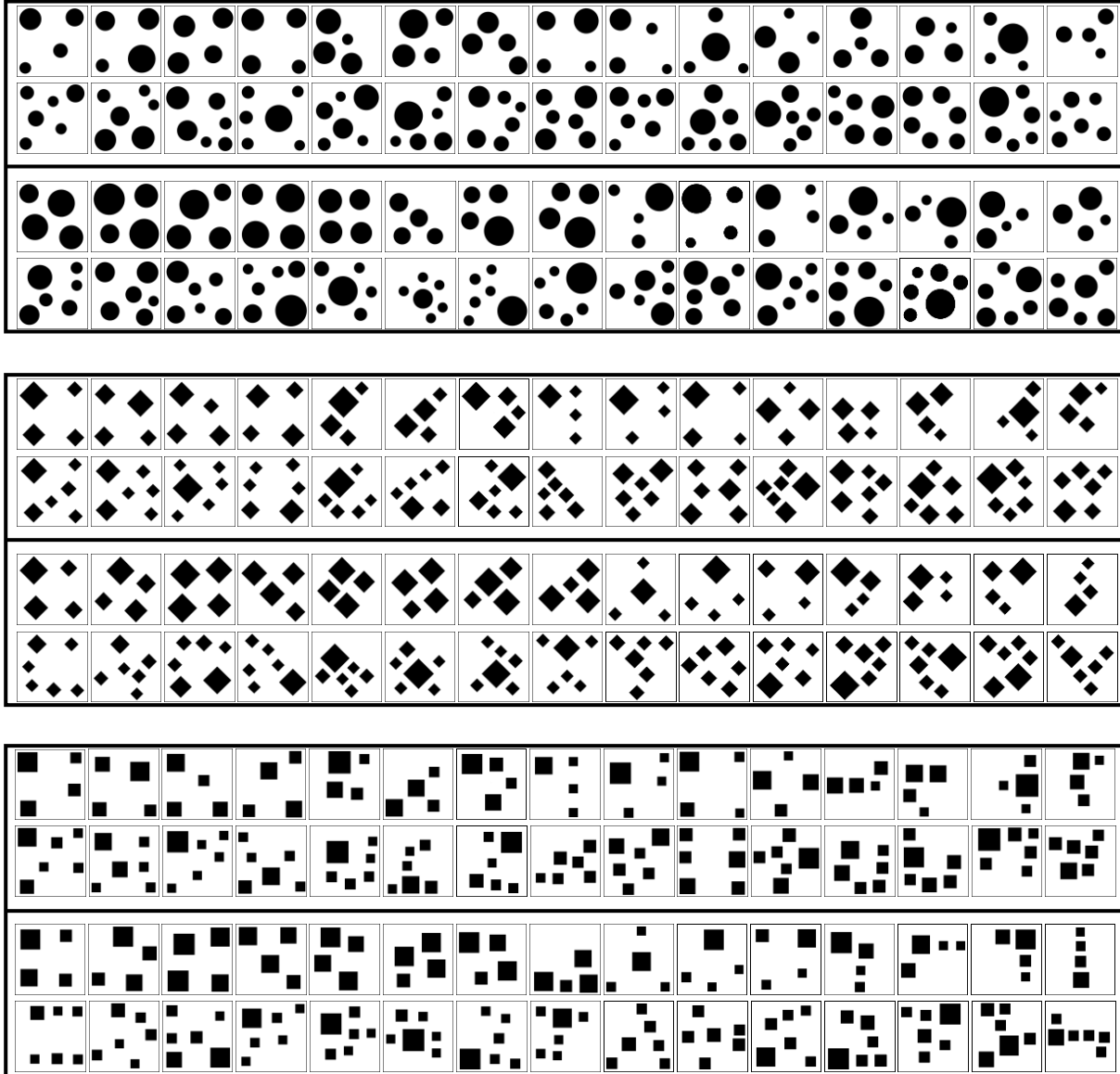


Figure S4. Numerical contrast 4 vs. 8, Related to Figure 1a.

The 90 stimuli pairs displaying the 4 vs. 8 numerical contrast. Stimuli are divided by shape categories: circles, diamond and squares. Within each shape, in one quarter of the pairs the cumulative surface area was matched to 100% whereas in the second quarter it was not controlled. In the third and the fourth quarter the contour length was matched to 100% and not controlled, respectively. Moreover, within each shape, half of the stimuli was controlled for the convex hull and the other half for the density of the elements.

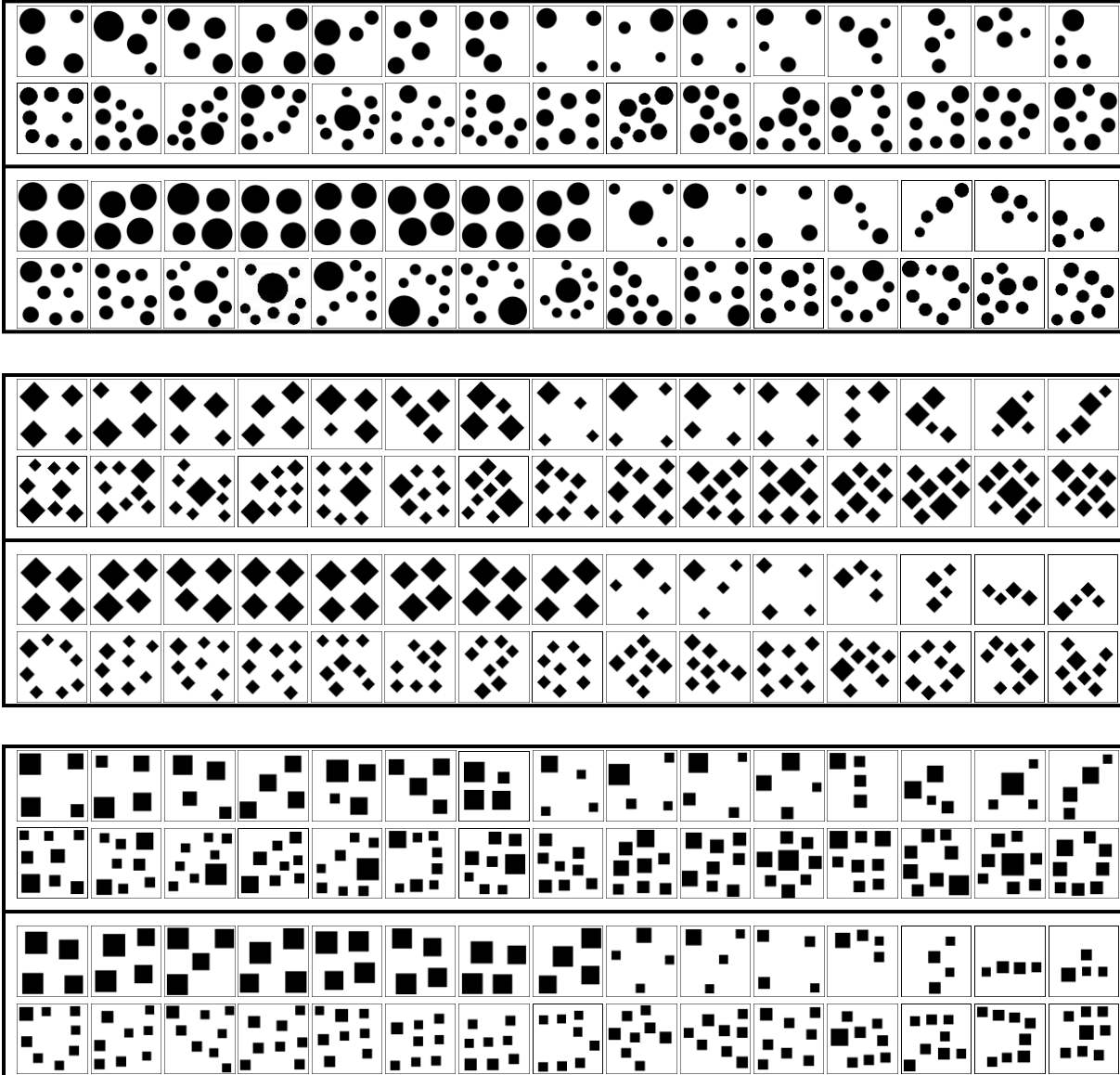


Figure S5. Size contrast, Related to Figure 1b.

The stimuli pairs used during the *size transfer* test. The stimuli consisted of two pairs of novel shapes (i.e., the shape that was not presented in the training phase) with sizes that differed by a ratio of either 0.5 or 0.67, depending on the numerical training previously completed by each subject. Within each pair, the two arrays had the same number and disposition of elements.

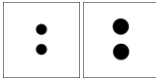
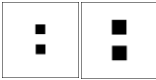
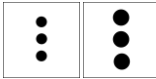
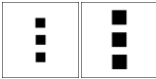
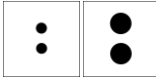
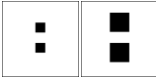
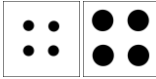
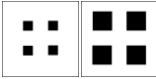
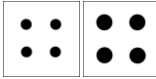
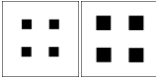
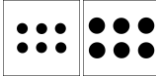
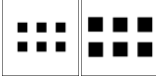
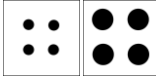
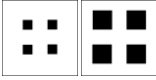
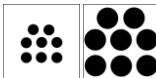
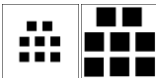
Numerical contrast	Correct numerosity during training	Stimuli <i>size transfer</i> test	
		Circles	Squares
2 vs. 3	2 elements		
2 vs. 3	3 elements		
2 vs. 4	2 elements		
2 vs. 4	4 elements		
4 vs. 6	4 elements		
4 vs. 6	6 elements		
4 vs. 8	4 elements		
4 vs. 8	8 elements		

Figure S6. Original scoring, Related to Figure 1a,b.

The original scoring of the test phase of 30 subjects (***) $P < 0.001$, Analysis of variance (ANOVA)).

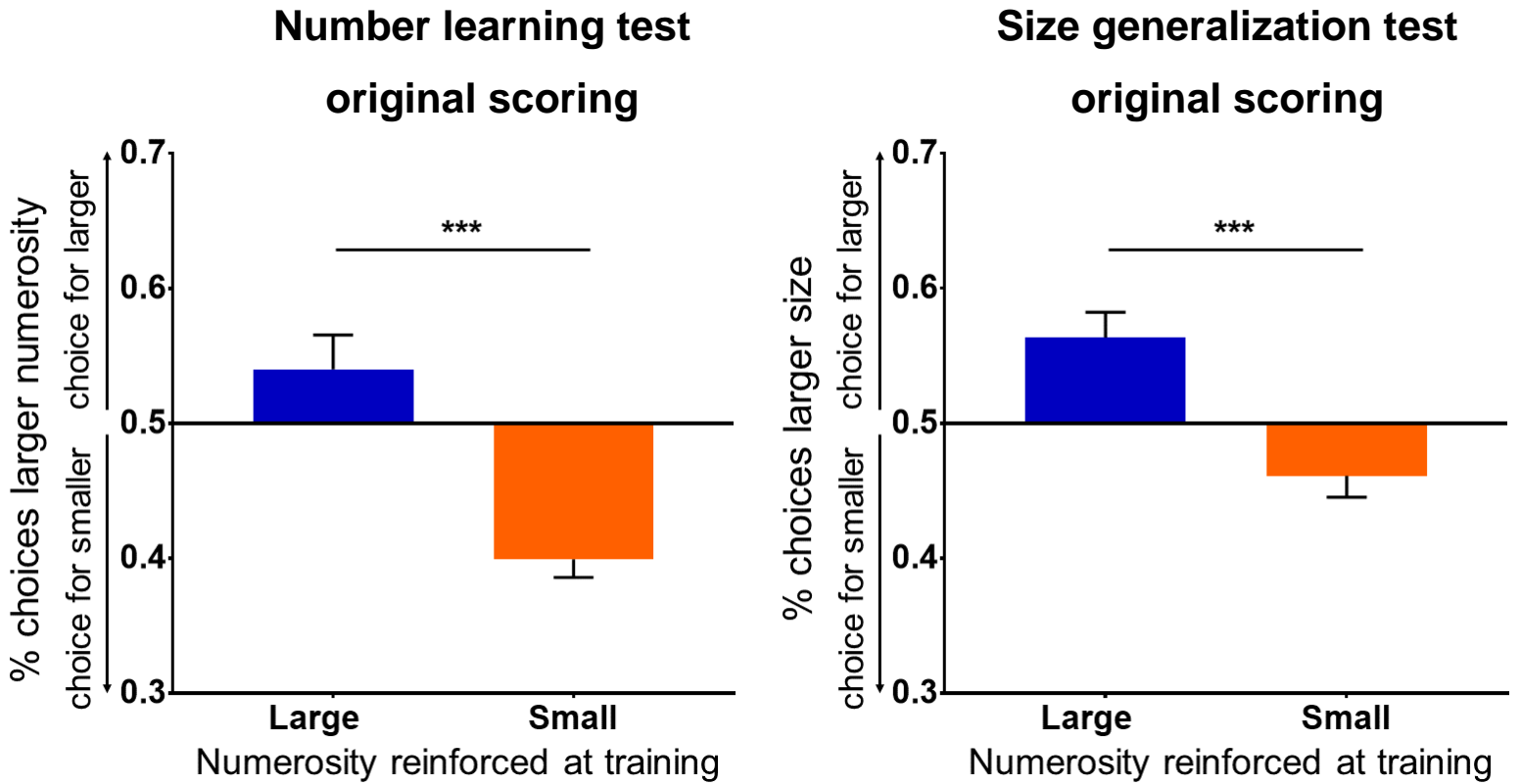
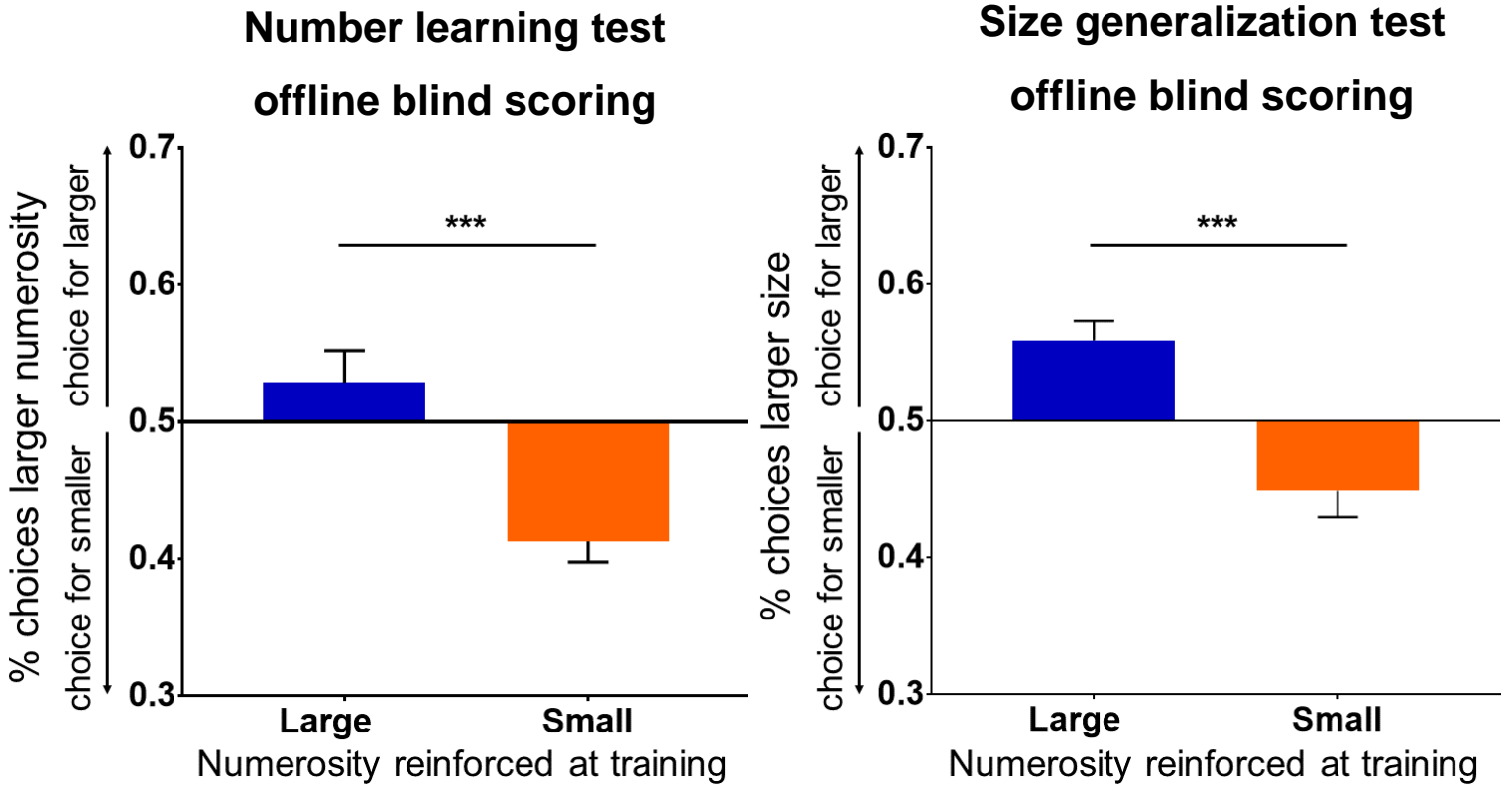


Figure S7. Offline blind scoring, Related to Figure 1a,b.

The offline blind scoring of the test phase of 30 subjects (***) $P < 0.001$, Analysis of variance (ANOVA))



TRANSPARENT METHODS

Experiments were performed during the Summer 2019 at SperimentArea, a field station run by the local Natural History Museum, in Rovereto (North of Italy). Thirty-two free-flying honeybees (*Apis mellifera*) were trained singly to fly into a wooden Y-maze (Fig. 1c). This sample size is common in experiments on free-flying honeybees' visual learning abilities (Bortot et al., 2019; Howard et al., 2019; Howard et al., 2018), due to the nature of the free-flight condition that do not allow a control by the experimenter on the decision of the single bee to come back freely to the apparatus. One half of the bees were trained with a 0.5 ratio (N=16) and the other half with a 0.67 ratio (N=16). In the 0.5 ratio one half of the subjects was tested with a 2 vs. 4 comparison and the other half with a 4 vs. 8 comparison; in the 0.67 ratio one half of the subjects was tested with a 2 vs. 3 comparison and the other half with a 4 vs. 6 comparison.

The stimuli consisted of black elements, either squares, diamonds or dots on a white squared-shape background (8 cm x 8 cm) located at 15 cm distance from the decision chamber (Fig. 1c). The stimuli size ranged from 1.12 cm to 3.56 cm (diameter of dots) and from 1 cm to 2.5 cm (side of squares and diamonds). A total of 30 couples of stimuli were used for each shape (i.e., squares, diamonds and dots). The spatial disposition and the size of the elements were varied among trials to prevent the use of non-numerical cues. In order to control for the continuous variables that covary with numerosity (e.g., area, contour length, density), we adopted a procedure previously used in other studies on numerical abilities of bees (Bortot et al., 2019). Within each shape, in one quarter of the stimuli (N=7) the cumulative surface area was matched to 100%, whereas in the second quarter (N=8) was not controlled (i.e., the ratio between the cumulative surface area within

each pair was congruent with the numerical ratio: 0.5 in 2 vs. 4 and 4 vs. 8; 0.67 in 2 vs. 3 and 4 vs. 6). In addition, half of the stimuli was controlled for the convex hull and the other half for the density of the elements. Furthermore, in the third and fourth quarter of the stimuli, the cumulative contour length was matched to 100% (N=8) and not controlled (N=7), respectively, following the same logic. Again, within each shape, half of the stimuli were controlled for the convex hull and the other half for the density of the elements. The control of these variables was performed for each shape (Fig. S1, S2, S3, S4).

During the training phase, half of the bees (N=16) were presented with squares and diamonds, whereas the other half (N=16) was presented with diamonds and dots. Thus, in the training phase bees were presented with 60 couple of stimuli in random order differing in shape, spatial disposition, size of the element and combinations of controlled continuous variable, and only the numerical information was kept constant. The stimuli used in the *number learning* test, were taken from the training sample of stimuli with the area matched to 100%. In the *size transfer* test, stimuli consisted of two pairs of novel shapes (i.e., the shape that was not presented in the training phase) having sizes that differed by a ratio of either 0.5 or 0.67, depending on the numerical training previously completed by each subject. Within each pair, the two arrays had the same number and disposition of elements. In particular, the number of elements presented was equal to the numerosity reinforced during the training phase (e.g., bees trained to select 2 elements over 4 elements during the training phase, were then presented with a 2 vs. 2 comparison where one group of 2 elements had the double size of the other group of 2 elements) (Fig. S5).

The experimental procedure comprised a pre-training phase followed by a training and tests phase. All the phases were completed by all subjects in 1 or 2 consecutive days. During the pre-training phase, each bee was individually habituated to fly inside the apparatus and to collect food

by landing on two grey poles placed in both arms, in the absence of visual stimuli. In the training phase, four different numerical comparisons (ratio 0.5: 2 vs. 4, 4 vs. 8; ratio 0.67: 2 vs. 3, 4 vs. 6) were presented to each independent group, separately. Within each group, half of the subjects was trained to select the smaller numerosity in the comparisons (either 2 or 4), whereas the other half was trained to choose the larger numerosity in the comparison (either 3, 4, 6 or 8) in order to get the food reward. During this phase, an appetitive-aversive conditioning paradigm was used: the correct numerosity was always associated with the food (0.88 M of sucrose solution) whereas the incorrect numerosity was always associated with a bitter 60 mM quinine solution, used as punishment (Avarguès-Weber et al., 2018). The use of this appetitive-aversive conditioning has been shown to improve the ability of bees to discriminate between numerosities (Howard et al., 2019). Each subject had to complete 60 consecutive trials of training. The stimuli were presented in a pseudo-random sequence (i.e., the correct/incorrect stimulus was never presented for more than two consecutive times on the same side).

Once completed the training phase, honeybees started the test phase. During this phase, two non-reinforced tests were presented: a *number learning* test and a *size transfer* test. Each test was presented twice to counterbalance the position of the correct array and avoid side preferences. The tests lasted 1 minute during which the number of choices (i.e., direct contact made with a body part, either the antennae or legs, on one of the two grey poles placed in front of each stimulus) made by the subjects were counted online. In the *number learning* test, bees were presented with the same numerical comparisons and shapes used during the training but in the absence of any reward. In the *size transfer* test bees were exposed to the novel stimuli displaying only the size information (even in this case without any reward).

In the test phase, the percentage of choices for the larger numerosity and larger size was calculated for each subject and analyzed, giving rise one single value per bee to exclude pseudo-replication. The data were checked for normality (Shapiro-Wilk normality test: $W = 0.98$, $P > 0.05$) and homoscedasticity (Levene's test: $P > 0.05$) and then analyzed with parametric statistical tests. An analysis of variance was performed with ratio (0.5 and 0.67), type of training (smaller vs. larger as positive) and type of test (*number learning* test vs. *size generalization* test) as factors. The effect of the numerical comparisons, as factor nested in ratio, was analyzed with a nested factorial Anova. The omega-squared effect size of any significant results in the Anova analysis was reported. The proportion of choices for the correct numerosity during the *number learning* test and the proportion of choices for the congruent size during the *size generalization* test were calculated for each subject and analyzed with a two-tailed one-sample *t*-test. The Cohen's *d* effect size of any significant result was also reported.

Analyses Blind Video Coding

We performed an offline scoring in blind condition for the videos of the test phase and compared them with the original analysis (due to corruption of two videos, 30 subjects were used in both conditions). The results are shown in the Figure S6 (original scorings) and Figure S7 (offline blind scoring) below together with statistical analyses and confirmed the original analyses.

For each group – *offline blind scoring* and *original* scoring - we calculate the percentage of choices for the larger numerosity during the *number learning* test and for the larger size during the *size generalization* test for each subject and analyzed. The data were checked for normality (*offline blind scoring*: Shapiro-Wilk normality test: $W: 0.99$, $P > 0.05$; *original scoring*: Shapiro-Wilk normality test: $W: 0.98$, $P > 0.05$) and homoscedasticity (*offline blind scoring*: Levene's test: $P >$

0.05; *original scoring*: Levene's test: $P > 0.05$) and then analyzed with parametric statistical tests. An analysis of variance was performed with ratio (0.5 and 0.67), type of training (smaller vs. larger as positive) and type of test (*number learning test vs. size generalization test*) as factors. The effect of the numerical comparisons, as factors nested in ratio, was analyzed with a nested factorial Anova. The omega-squared effect size of any significant results was reported.

Results

Offline blind scoring

An analysis of variance revealed a significant main effect of the type of training (smaller vs. larger numerosity as positive; $F_{(1, 44)} = 35.9, P < 0.001, \omega^2 = 0.371$) but not of the type of test (*number learning test vs. size generalization test*; $F_{(1, 44)} = 3.5, P = 0.069$) and of the ratio (0.5 vs. 0.67; $F_{(1, 44)} = 0.19, P > 0.05$).

No significant interactions were observed (ratio x numerical comparisons: $F_{(2, 44)} = 1.7, P > 0.05$; ratio x type of training: $F_{(1, 44)} = 2.2, P > 0.05$; ratio x type of test: $F_{(1, 44)} = 1.06, P > 0.05$; type of training x type of test: $F_{(1, 44)} = 0.09, P > 0.05$; ratio x type of training x type of test: $F_{(1, 44)} = 1.2, P > 0.05$; ratio x type of training x numerical comparisons: $F_{(2, 44)} = 0.04, P > 0.05$; ratio x type of test x numerical comparisons: $F_{(2, 44)} = 0.1, P > 0.05$; ratio x type of training x type of test x numerical comparisons: $F_{(2, 44)} = 0.5, P > 0.05$).

Original scoring

An analysis of variance revealed a significant main effect of the type of training (smaller vs. larger numerosity as positive; $F_{(1, 44)} = 41.6, P < 0.001, \omega^2 = 0.389$) and of the type of test (*number*

learning test vs. size generalization test; $F_{(1, 44)} = 5.1, P = 0.029, \omega^2 = 0.039$) but not of the ratio (0.5 vs. 0.67; $F_{(1, 44)} < 0.0005, P > 0.05$).

No significant interactions were observed (ratio x numerical comparisons: $F_{(2, 44)} = 2.2, P > 0.05$; ratio x type of training: $F_{(1, 44)} = 2.3, P > 0.05$; ratio x type of test: $F_{(1, 44)} = 0.4, P > 0.05$; type of training x type of test: $F_{(1, 44)} = 1.1, P > 0.05$; ratio x type of training x type of test: $F_{(1, 44)} = 0.6, P > 0.05$; ratio x type of training x numerical comparisons: $F_{(2, 44)} = 0.8, P > 0.05$; ratio x type of test x numerical comparisons: $F_{(2, 44)} = 1.09, P > 0.05$; ratio x type of training x type of test x numerical comparisons: $F_{(2, 44)} = 0.06, P > 0.05$).

References

Avarguès-Weber, A., de Brito Sanchez, M.G., Giurfa, M. and Dyer, A.G., 2010. Aversive reinforcement improves visual discrimination learning in free-flying honeybees. *PLoS One*, 5(10).

Bortot, M., Agrillo, C., Avarguès-Weber, A., Bisazza, A., Miletto Petrazzini, M.E. and Giurfa, M., 2019. Honeybees use absolute rather than relative numerosity in number discrimination. *Biology Letters*, 15(6), p.20190138.

Howard, S.R., Avarguès-Weber, A., Garcia, J.E., Greentree, A.D. and Dyer, A.G., 2019. Numerical cognition in honeybees enables addition and subtraction. *Science advances*, 5(2), p.eaav0961.

Howard, S.R., Avarguès-Weber, A., Garcia, J.E., Greentree, A.D. and Dyer, A.G., 2018. Numerical ordering of zero in honey bees. *Science*, 360(6393), pp.1124-1126.

Howard, S.R., Avarguès-Weber, A., Garcia, J.E., Greentree, A.D. and Dyer, A.G., 2019. Surpassing the subitizing threshold: appetitive–aversive conditioning improves discrimination of numerosities in honeybees. *Journal of Experimental Biology*, 222(19), p.jeb205658.