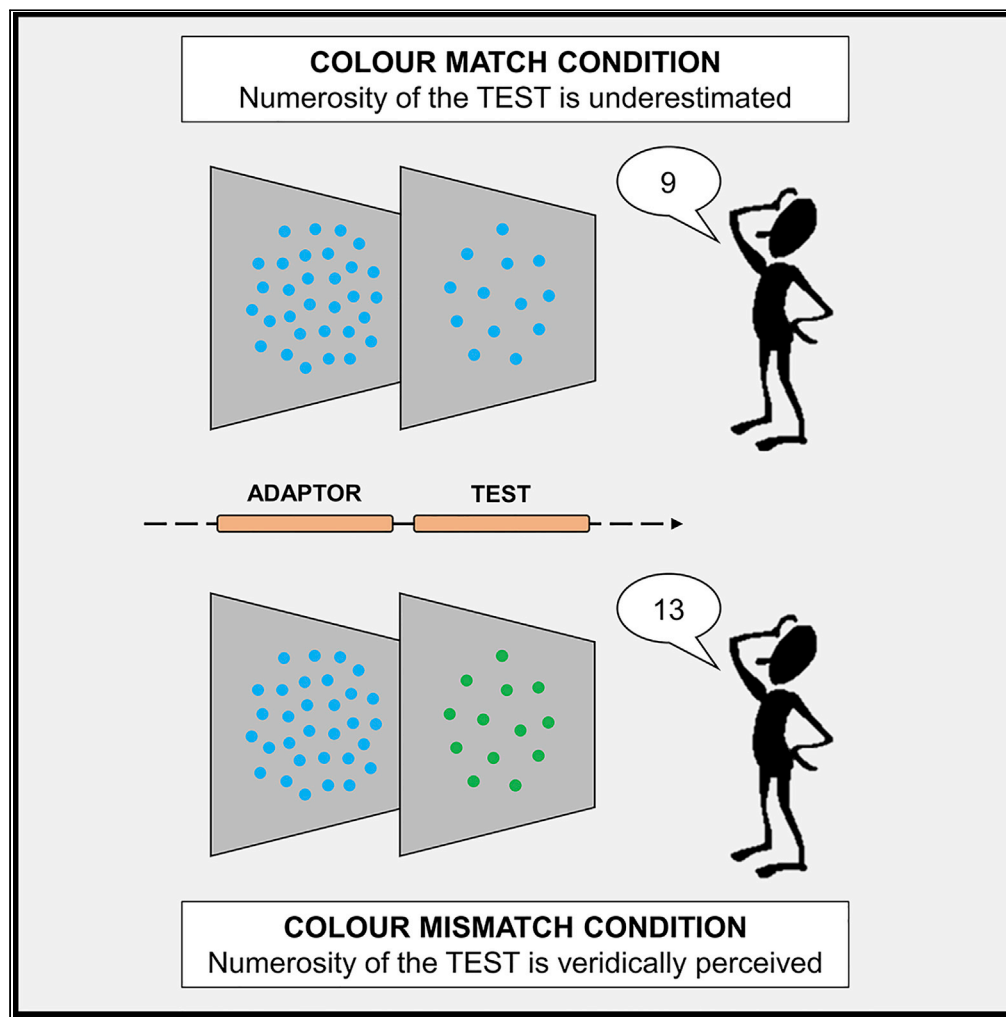**Article**

Numerosity perception is tuned to salient environmental features

Paolo Antonino Grasso, Giovanni Anobile, Roberto Arrighi, David Charles Burr, Guido Marco Cicchini

paolo.grasso@unifi.it

Highlights
Numerosity perception is highly selective to salient non-numerical features

Visual numerosity adaptation requires similar perceived test and adapter color

Auditory numerosity adaptation requires similar test and adapter pitch

Article

Numerosity perception is tuned to salient environmental features

Paolo Antonino Grasso,[1,3,*] Giovanni Anobile,[1] Roberto Arrighi,[1] David Charles Burr,[1] and Guido Marco Cicchini[2]

SUMMARY

Numerosity perception is a key ability to guide behavior. However, current models propose that number units encode an abstract representation of numerosity regardless of the non-numerical attributes of the stimuli, suggesting rather coarse environmental tuning. Here we investigated whether numerosity systems spontaneously adapt to all visible items, or to subsets segregated by salient attributes such as color or pitch. We measured perceived numerosity after participants adapted to highly numerous stimuli with color either matched to or different from the test. Matched colors caused a 25% underestimation of numerosity, while different colors had virtually no effect. This was true both for physically different colors, and for the same colors perceived as different, via a color-assimilation illusion. A similar result occurred in the acoustic domain, where adaptation magnitude was halved when the adaptor and test differed in pitch. Taken together, our results support the idea that numerosity perception is selectively tuned to salient environmental attributes.

INTRODUCTION

Numerosity is thought to be a primary visual property, like brightness, color, form, or motion. Human adults, infants, and also non-human species can make rapid and accurate estimates of the number of items without serial counting (Butterworth, 1999; Dehaene et al., 2008; Hauser et al., 2000; Whalen et al., 1999). Importantly, psychophysical and imaging (fMRI) studies have shown that this capacity is highly adaptable, supporting the idea of a dedicated brain system for the analysis of numerical elements (e.g., Aagten-Murphy and Burr, 2016; Arrighi et al., 2014; Burr and Ross, 2008; Castaldi et al., 2016; Fornaciai et al., 2016; Grasso et al., 2021a; Grasso et al., 2021b; Togoli and Arrighi, 2021; Tsouli et al., 2020).

Many models of numerosity perception propose that numerosity detectors would normalize non-numerical characteristics of the stimuli (such as size), suggesting that number sensitive units can discard the non-numerical features of the items (e.g., Dehaene and Changeux, 1993; Zorzi et al., 2005). This conceptualization is in line with imaging data reporting units in the parietal lobe that can be adapted by various types of number representation, such as both Arabic digits and dot stimuli (Piazza et al., 2007). It is also in line with reports of numerosity adaptation being insensitive to the visual attributes such as size, orientation, or contrast of single elements (Burr and Ross, 2008).

However, a system that enumerates items in a visual scene, irrespective of their category or type, would be unselective to the specificity of target objects in the visual scene. For example, it may be more important to estimate the number of red (ripe) apples, than the total number. Numerosity systems should, therefore, operate on segregated and categorized visual items, rather than on the total quantity. Halberda et al. (2006) showed that the numerosity of specifically colored items can be accurately estimated, even when spatially intermingled with other differently colored items. Indeed, adults, but also infants (Zosh et al., 2011), can enumerate simultaneously up to three different subsets of elements defined by different colors (but see also Yu et al., 2019). Similarly, the impressive demonstrations of "connectiveness" (He et al., 2009; Pomè et al., 2021) showing that connecting items by thin lines causes gross underestimation of numerosity highlights the role of "grouping" and item segregation in numerosity perception. More, when items are easily segregable in small groups, the precision of numerosity estimation increases by up to 20%, both for visual and auditory stimuli ("groupitizing"). The segregation can be aided by items having similar visual color or

[1]Department of Neuroscience, Psychology, Pharmacology and Child Health, University of Florence, Via di San Salvi 12, Padiglione 26, 50135 Florence, Italy

[2]Consiglio Nazionale delle Ricerche, Institute of Neuroscience, Via Giuseppe Moruzzi, 56124 Pisa, Italy

[3]Lead contact

*Correspondence: paolo.grasso@unifi.it

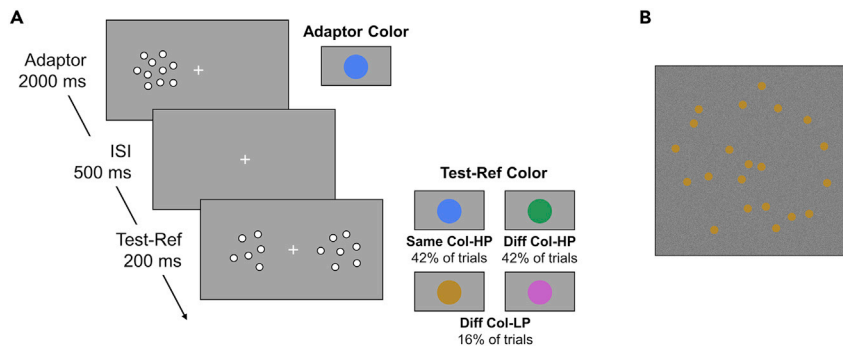https://doi.org/10.1016/j.isci.2022.104104

Figure 1. Paradigm and stimuli

(A) Schematic representation of the paradigm. The color of both test and reference was either blue, green, yellow, or purple, while the color of the adaptor was fixed (blue for seven participants, as illustrated here, and green for the other five). This produced a condition in which the color of the test and reference matched that of the adaptor with a high presentation probability (42% of trials - *Same Col-HP*), a condition in which the color of test and reference did not match that of the adaptor but had a high presentation probability (42% of trials - *Diff Col-HP*) and a condition in which the color of test and reference did not match that of the adaptor and had a low presentation probability (16% of trials - *Diff Col-LP*). Participants responded which of the two clouds of dots (test or reference) was perceived as more numerous, using the left and right arrows of the keyboard.

(B) Example of the stimuli used in the task.

auditory pitch, suggesting that sensory signals could be pre-processed by their low-level features, providing these are salient (Anobile et al., 2020, 2021b).

The site of numerosity adaptation is still debated. Several authors (e.g., Aulet and Lourenco, 2021; Dakin et al., 2011; Durgin, 2008) have suggested that number is represented at an early stage of visual processing, together with size and other low-level features. Durgin (2008) specifically claimed that adaptation to the number is also early, occurring early in processing, possibly mediated by adaptation to texture. However, evidence is mounting that number is represented at higher levels, particularly the parietal cortex. fMRI (Harvey et al., 2013; Piazza et al., 2004) and EEG (Hyde and Spelke, 2012) studies all point to a parietal- and frontal-lobe representation of number. Adaptation studies show specific changes in the parietal response to numerosity, with minimal relevant changes at earlier levels (Castaldi et al., 2016; Tsouli et al., 2021a). Psychophysical studies also point to a high level of adaptation, after the site of grouping effects (Arrighi et al., 2014; Fornaciai et al., 2016). If adaptation to numerosity occurs at relatively high levels, as most evidence suggests, then one may expect the adaptation to be conditioned by other processes relevant to numerosity perception, including object segmentation and grouping.

In the present study, we used adaptation to numerosity as a tool to investigate how highly salient perceptual features, such as color for vision and pitch for audition, interact with numerosity perception. Colors is possibly the most salient natural visual attribute enabling rapid object segregation (e.g., Found and Müller, 1996), while also providing crucial evolutionary signals related to food, sex, and fight or flight behaviors. By analogy, it has been shown that the auditory system exploits pitch for auditory stream segregation to distinguish, for example, the voices of two different persons or the sounds of different instruments (e.g., Moore and Gockel, 2002). If numerosity perception is optimized for environmentally salient features, we expect that numerosity adaptation aftereffects should be more prominent when the adapting and test stimuli share the same features. This would also suggest that numerosity adaptation occurs after feature-binding and other processes have contributed towards object recognition.

RESULTS

Visual numerosity adaptation is tuned to color

We first investigated whether the congruency in color of adaptor and test affects the magnitude of numerosity adaptation, using a numerosity discrimination paradigm where the color of test and reference either matched or differed from that of the adaptor (Figure 1). On each trial, the color of test and reference varied randomly between blue, green, yellow, and purple (see General procedure and stimuli section for further details), while the color of the adaptor was kept fixed. To disentangle between the effects of color congruency and those potentially
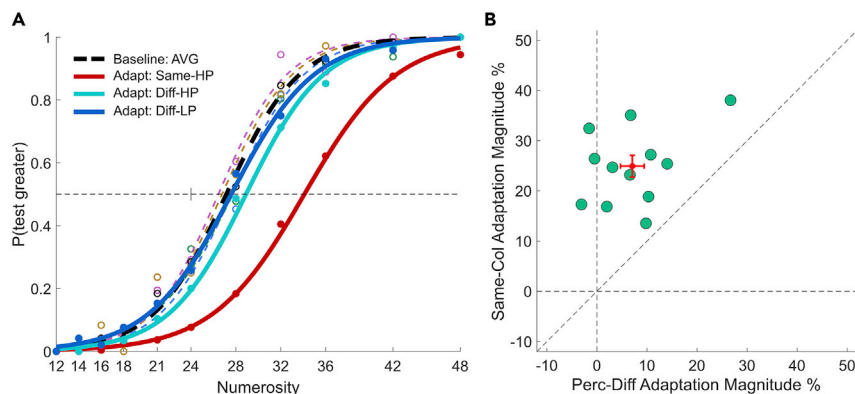
Figure 2. Visual numerosity adaptation requires same test and adaptor color

(A) Thin dotted curves depict the best-fitting curves for baseline conditions defined by the four colors (blue, green, yellow, and purple). Thick black dotted curve shows the psychometric fitting of an average baseline and solid curves the psychometric fittings of the three adaptation conditions (Same Col-HP: solid red curve; Diff Col-HP: solid cyan curve; Diff Col-LP: solid blue curve).

(B) Scatterplots depicting the normalized percentage of PSE shift in the Same-Col HP condition against the Diff-Col HP condition. Data falling above the equality line indicate a greater adaptation magnitude in the Same-Col HP condition. Red dot is the average and standard error across participants.

driven by repetition-dependent attentional color tuning, the four colors had different presentation frequencies. In most of the trials (84%) blue and green dots were presented (42% blue and 42% green), while in a minority of trials either yellow or purple dots were presented (8% yellow and 8% purple). This led to three separate conditions: color of test and reference matched that of the adaptor and had a high (42%) presentation probability (*Same Col-HP*); color of test and reference not matched to that of adaptor and had a high (42%) presentation probability (*Diff Col-HP*); color of test and reference not matched to that of the adaptor and had a low (16%) presentation probability (*Diff Col-LP*). Before the experiment, we tested that all participants were able to discriminate the color of the stimuli with 100% accuracy.

Figure 2A shows results averaged across participants, for all experimental conditions. Thin-colored dotted curves show the best-fitting psychometric functions for the four baseline conditions (blue, green, yellow, and purple) while the thick black dotted curve shows the best-fitting curve for the average baseline data. The four baseline curves are clearly superimposed, showing that basic numerosity discrimination judgments were not influenced by the color of the dots. As there was no significant main effect of color at baseline (see later in discussion), we pooled the two colors, leading to the three adaptation conditions depicted by solid curves in Figure 2A. The red solid line shows the condition where test and reference had the same color as the adaptor and were frequently presented (Same Col-HP), while cyan and blue solid lines show the conditions where test and reference were of different colors from the adaptor, presented either frequently (Diff Col-HP) or infrequently (Diff Col-LP). Adaptation of the same color produced much larger effects than adaptation to different colors, irrespective of probability of stimulus presentation, as obvious by the rightward shift of the red psychometric curve shown in Figure 2A.

Figure 2B summarizes the results for the individual participants, obtained by best-fitting psychometric functions like those of Figure 2A to the data of each participant (see Analysis procedure for details). The scatterplot plots the adaptation effect (normalized difference in adapted and baseline PSEs) for adaptors of the same color as the test, against adaptors of different colors (all high probability stimuli). All data points are above the equality line, indicating that all participants had shown stronger adaptation when the color of adaptor and test were matched than when they differed. The effect is strong, about a factor of 5, as the magnitude of numerosity adaptation in the color matched condition was around 25% while the two unmatched color conditions produced an average of 5% adaptation. Significance was confirmed by a strongly significant main effect for Condition (Same Col-HP, Diff Col-HP, Diff Col-LP) in one-way ANOVA ($F_{(2, 22)} = 38.583$; $p < 0.001$; $\eta^2 = 0.780$, $LogBF_{10} = 14.2$).

There were no differences in the magnitude of effects for participants adapted to blue or green (mixed ANOVA within factor Condition and between factor adaptor color - no main effect of color: $F_{(1, 10)} = 0.02$;
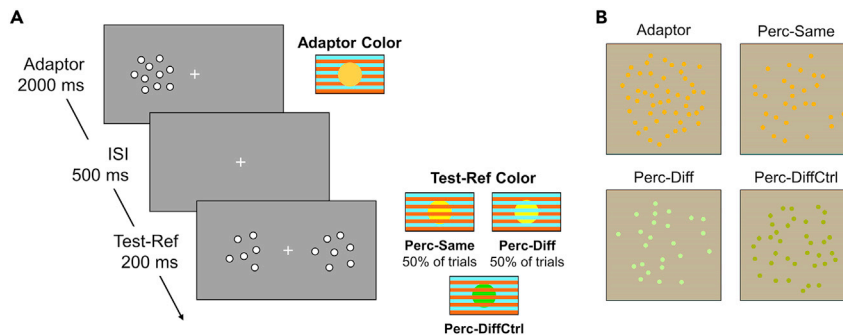
Figure 3. Paradigm and stimuli

(A) Schematic representation of the paradigm. In an intermingled presentation, test and reference stimuli were either yellow dots behind red stripes or yellow dots behind cyan stripes. The two conditions produced a perceptual illusion of dots appearing, respectively, "orange" or "green." The adaptor stimulus was always in front of both stripes and its color was perceptually matched to the color of the stimuli perceived as orange (yellow dots with superimposed red stripes), yielding one condition where the color of test and reference appeared the same as the adaptor (*Perc*-Same condition) and another where it appeared different (*Perc-Diff* condition). A further perceptually different color conditions (*Perc-DiffCtrl*) was included to control for the effects of physical color similarity between the adaptor and the foreground red stripes. Participants reported which of the two dot clouds (test or reference) was perceived as more numerous using the left and right arrows of the keyboard.

(B) Sample of the stimuli used in the task.

p = 0.88; no interaction color x Condition: $F_{(2, 20)}$ = 0.85; p = 0.44). Nor was there a significant main effect of color at baseline ($F_{(3, 33)}$ = 1.3; p = 0.29; η^2 = 0.11; LogBF$_{10}$ = −1.02), and no difference between yellow and purple tests in either baseline ($t_{(11)}$ = 0.39, p = 0.70; d = 0.11; LogBF$_{10}$ = −1.18) or adaptation ($t_{(11)}$ = −0.274, p = 0.789; d = −0.08; LogBF$_{10}$ = −1.21).

Taken together, these results show that numerosity adaptation is selective for non-numerical features (color) of the adaptor, with a selectivity that cannot be accounted for by repetition-dependent color tuning.

Visual numerosity adaptation is tuned to perceived color

We next tested whether the color-congruent selectivity of numerosity adaptation is driven by physical or perceived color similarity, to examine the neural level of interactions between color and numerical processing, given that the perceptual color experience is believed to be encoded in higher cortical regions than the encoding of stimulus chromaticity (Kim et al., 2020). We took advantage of color assimilation, a well-known color phenomenon driven by perceptual grouping of bi-color patches (King, 1988). Stimuli were yellow dots presented on a red and cyan horizontal striped background (see General procedure and stimuli section for further details). In an intermingled paradigm, dots were partially covered by either red or cyan stripes, producing the perceptual illusion of either "orange" (*Perc-Orange*) or "green" (*Perc-Green*) dots. The adaptor stimulus was always superimposed on both stripes, and its color was customized for each participant to match the color of the stimuli perceived as orange (yellow dots with superimposed red stripes), using a color matching procedure before the beginning of the experiment. Thus, the Perc-Orange stimuli corresponded to the *Perc-Same* condition (stimuli of same perceived color as adaptor), and the Perc-Green stimuli to the *Perc-Diff* condition (different perceived color from adaptor; Figure 3). A further control condition used green dots partially covered by red stripes (i.e., *Perc-DiffCtrl*), to test whether the foreground red stripes (of similar chromaticity to the orange adaptors) may be driving the adaptation. Importantly, none of the test/reference colors was physically identical to the color of the adaptor, so any chromatic congruency was purely illusory. Before taking part in the experiment, we made sure participants could discriminate perfectly between the colors used in the experiment.

Figure 4A shows the psychometric functions data averaged across participants for each experimental condition. As before, adaptation to the same perceived color produced a much larger rightward shift of the psychometric curve, showing greater underestimation of the test stimulus numerosity compared with when the perceived color of the stimuli was different (Figure 4A). Figure 4B shows that the effect held for every participant, as all data points are above the equality line. On average, the adaptation magnitude for the Perc-Same condition was 29% while for the Perc-Diff condition was 11%. Adaptation to the other control with red foreground stripes (Perc-DiffCtrl) was
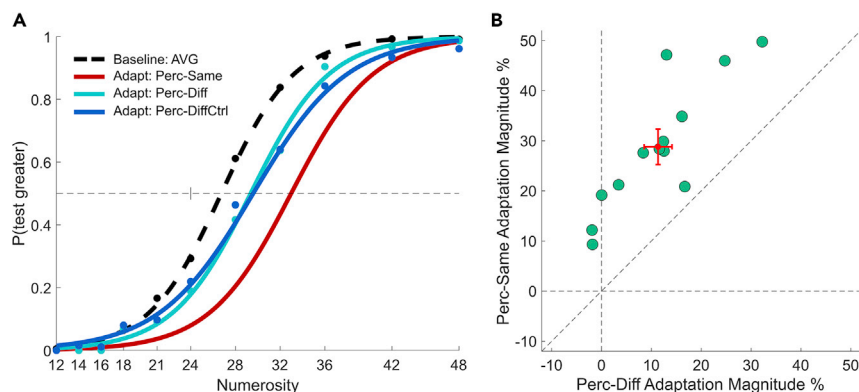
Figure 4. Visual numerosity adaptation requires same perceived color of test and adaptor

(A) Thick black dotted curve the psychometric fit of the average baseline, and solid curves the psychometric fit of the three adaptation conditions (Perc-Same: solid red curve; Perc-Diff: solid cyan curve; Perc-DiffCtrl: solid blue curve.

(B) Scatterplot of adaptation of individual participants in the Perc-Same and Perc-Diff conditions. Data falling above the equality line show greater adaptation in the Perc-Same condition. Red dot is the average and standard error across participants.

9%. The effect was confirmed by a strongly significant main effect for Condition (Perc-Same, Perc-Diff, Perc-DiffCtrl) in one-way ANOVA ($F_{(2, 24)}$ = 20.43; p < 0.001; η^2 = 0.63, LogBF$_{10}$ = 9.16). Each adaptation condition was different from zero (all ps < 0.01) but the Perc-Same condition yielded a significantly larger effect with respect to the other two (p < 0.001) which did not differ from each other.

Overall, this experiment confirmed that the adaptation-induced distortions of perceived numerosity are much more tuned to perceived rather than physical similarity of color of the adapting and adapted stimuli.

Auditory numerosity adaptation is tuned to pitch

We then tested whether feature specificity of numerosity adaptation is a general property of numerosity perception, generalizing to non-visual sensory modalities such as audition. In this experiment, participants estimated the numerosity of a sequence of tones after adapting to a fast or slow sequence of auditory pulses (Figure 5).

In a blocked design, the pitch of the test sequence was either 400 or 800 Hz, while the pitch of the adaptor was fixed at 400 Hz for five participants or 800 Hz for the other six. For all participants, the test either matched or was different from that of the adaptor (*Same-Pitch* or *Diff-Pitch*). Following the procedure of Arrighi et al., (2014), the adaptation sequence was either slow (average 2 Hz) or fast (average 8 Hz), with adaptation magnitude defined as the proportional difference between estimates after adapting to 2 or 8 Hz (see Analysis procedure for further details). To anchor estimation judgments within the correct numerical range, at the beginning of each block participants were familiarized with the presentation of the highest and the lowest numerous test sequence. Before commencing the experiment, we checked that participants could discriminate the 400 and 800 Hz pitches from each other with 100% accuracy.

Figure 6A show averaged estimates of numerosity after adaptation to fast (8 Hz) and to slow (2 Hz) sequences, for each physical numerosity, for the two adaptation conditions (Same-Pitch, Diff-Pitch). In both cases, the estimates deviated from veridicality (dashed black line), showing a regression to the mean, as to be expected (Hollingworth, 1910). However, the estimates are generally lower after adapting to high rather than low adaptors and the effect of adaptation is less when the tones are not matched than when they are.

To quantify this difference, we plotted the difference in numerosity estimates after adaptation to fast and slow adaptors, separately for the Same- and Diff-Pitch conditions (Figure 6B). Adaptation was clearly larger when adaptor and test shared the same pitch. We quantified this by fitting the data with a linear regressor anchored at zero (red and cyan lines). The slope of the regressor gives the proportion of adaptation for each condition. This was 11% for the Same-Pitch condition, and only 6% for the Diff-Pitch condition, about half as much. Single participant data confirmed this trend. Figure 6C plots adaptation magnitude for the two conditions. Almost all data points (10 out of 11) lie above the equality line, confirming that numerosity
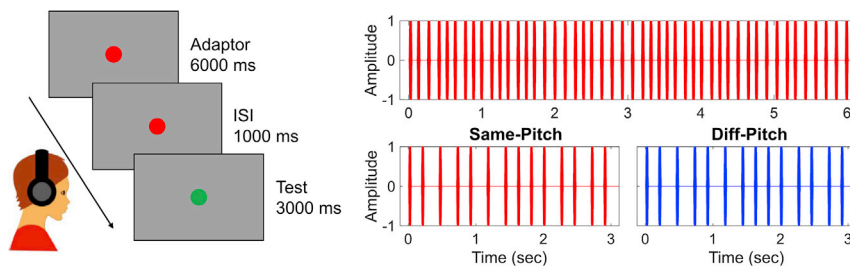
Figure 5. Paradigm and stimuli
In a blocked design, participants were presented with a 6-s adaptation sequence comprising either 48 tones (~8 Hz) or 12 tones (~2 Hz), positioned along the 6-s interval with random jitter (see STAR Methods). They then estimated the numerosity of a sequence of tones having either the same or different pitches from that of the adaptor, interleaved within each session. To help participants correctly distinguish between adaptation and test sequences, a central circle turned from red to green whenever the sequence turned from adaptation to test.

adaptation was selective for similarity of pitch. Two-tailed dependent sample t-test confirmed that the difference was significant ($t_{(10)}$ = 3.3, p = 0.008; d = 0.98; $LogBF_{10}$ = 1.95). The result was confirmed by a 2 × 9 ANOVA with within factors Type (Same-Pitch, Diff-Pitch) and Numerosity, performed on average numerosity adaptation magnitude. Results revealed a significant main effect of Type ($F_{(1, 10)}$ = 15.11; p = 0.003; η^2 = 0.1) while both Numerosity ($F_{(8,80)}$ = 1.69; p = 0.11; η^2 = 0.05) and Numerosity x Type were not significant ($F_{(1,10)}$ = 0.81; p = 0.50; η^2 = 0.03).

Precision does not change with adaptation

To test whether the selective adaptation may have been driven by differences in precision, we calculated precision for the adaptation conditions, from the coefficient of variation (threshold normalized by the mean numerosity). Figure 7 shows the results. Precision was similar across the different adaptation conditions employed, either for the visual or auditory stimuli. This was confirmed by statistical tests on the coefficients of variation. For changes in physical color: $F_{(2, 22)}$ = 0.824; p = 0.404; η^2 = 0.07, $LogBF_{10}$ = 1.08; illusory color: $F_{(2,24)}$ = 1.24; p = 0.30; η^2 = 0.09, $LogBF_{10}$ = 0.84; auditory tones: $t_{(10)}$ = 0.94, p = 0.369; d = 0.28; $LogBF_{10}$ = 0.84. Note that similar non-significant results were obtained using *Weber fractions* (thresholds normalized by apparent numerosity), or raw thresholds.

DISCUSSION

In the present work, we provide evidence that numerosity adaptation is a flexible mechanism tuned to salient environmental features. We exploited a highly salient visual attribute, color, to investigate whether numerosity perception is sensitive to salient attribute changes within separate sets of visual elements. Numerosity adaptation did not generalize across stimuli differing in color, as prolonged presentation of a numerous patch of dots did not lead to the underestimation of a subsequent patch of different colored dots. This result supports the idea that numerosity perception is closely bound to salient environmental constancies. It also suggests that adaptation occurs at a relatively high level of processing, after binding and grouping operations.

Adaptation was selective for perceived rather than physical color, as underestimation of the test stimulus also occurred when adaptor and test were physically different but perceptually identical (Figure 4), while it was robustly reduced when the color of test and reference was perceptually distinct from that of the adaptor. In this experiment, there was a residual of around 10% underestimation of the numerosity of test stimulus after adaptation to perceptually different colors (11% in the Perc-Diff and 9% in the Perc-DiffCtrl conditions), probably due to a residual similarity of the perceived color of the test dots in the control conditions to those of the adaptor (orange). Indeed, as the control experiment prescribes employing the same colors used in the "PercSame" condition we did not have enough room for employing test dots which differed completely from the warm colors used in the main condition. Indeed, the resulting colors were in the same region of the orange adaptor. In any event, even if one assumed that the two control conditions captured two essential components of the full conditions, gauging, respectively, the effect of the adaptor on yellow dots and red stripes, it is interesting to note that the predicted effect would still be about 20%. The effect in the PercSame condition on the other hand is 29%, more than root-2 larger. This suggests that the proper combination allowing assimilation into an orange illusory whole (i.e., yellow dots with foreground red stripes) is critical for the full effect to occur.
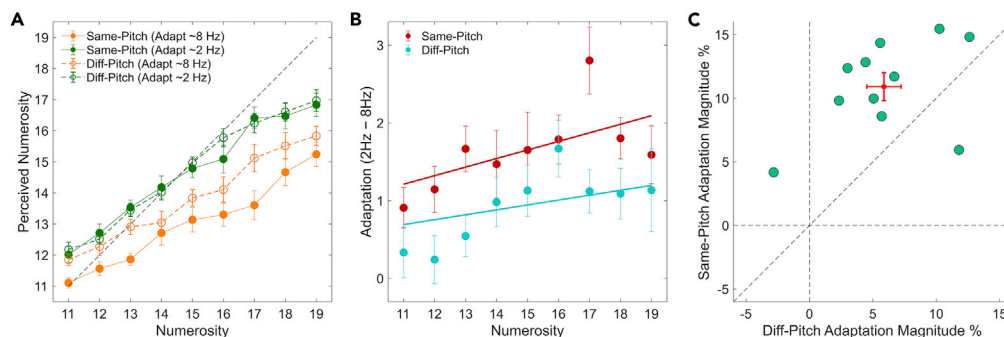
Figure 6. Auditory numerosity adaptation requires same test and adaptor pitch

(A) Average responses after adaptation to a 2 Hz sequence (green plots) or to an 8 Hz sequence (orange plots) for the Same- and the Diff-Pitch conditions. Bars represent SEM.

(B) Difference in numerosity estimation between 2 and 8 Hz adaptors, for the Same-Pitch condition (red plot) and the Diff-Pitch condition (cyan plot). Lines are linear regressions anchored to zero. Bars represent SEM.

(C) Scatterplots of average numerosity adaptation magnitude in the Same-Pitch adaptation condition against the Diff-Pitch adaptation condition. Dots represent single participant's data. Red dot is the average and standard error across participants.

The results with visual stimuli are consistent with and extend those of Halberda et al. (2006), who reported effortless enumeration of up to three subgroups defined by color (see also Liu et al., 2020): if more than three colors were employed, performance dropped. Our results indicate that as long as colors are perceptually distinct, different numerosity subsystems are likely to be engaged.

Numerosity perception transcends single sensory modalities, pointing to a generalized sense of number. This idea is reinforced by experiments showing the minimal cost in cross-modal numerosity estimation (Barth et al., 2003), and clear cross-modal adaption effects between vision, audition, and touch (Anobile et al., 2021a; Arrighi et al., 2014; Togoli and Arrighi, 2021). With a modified version of the paradigm used by Arrighi et al. (2014) we showed that adaptation was halved when the pitch was not matched. A small residual adaptation (∼6%) remained in the incongruent pitch condition. This may result from the choice of pitch frequencies (400 and 800 Hz), corresponding to the same note an octave apart, although this idea would require testing. Nevertheless, we confirmed within the auditory domain that adaptation to numerical quantities is linked to categorical differences between the adaptor and test, corroborating the idea that the numerosity system is capable of selectively discarding redundant information, while leaving salient "novelties" mostly unaffected.

Taken together, our results suggest that the numerical system could be adaptively tuned to salient characteristics of the stimuli. Although it has been reported that changes in orientation, size, and contrast do not affect the magnitude of numerosity adaptation (Burr and Ross, 2008), we speculate that these differences may not have been salient enough to group the stimuli into different categories. Colors is known to be highly salient, giving rise to what has been termed "pop-out" in the serial search literature (e.g., Carter, 1982; D'Zmura, 1991; Farmer and Taylor, 1980). Pitch is similarly salient. However, at this stage, we cannot be certain whether the results reported here are specific for color and pitch or would occur for any salient feature. In their original study, Burr and Ross (2008) tested orientation selectivity with short lines, which were not particularly salient, did not "pop out." It would be interesting to test whether more salient orientation differences lead to the selectivity of numerosity adaptation. It would also be interesting to see if other salient features, such as coherent motion, lead to selective adaptation.

Our results suggest that numerosity is encoded after categorization, and after feature binding, and that this process can occur spontaneously. Previous evidence showed that human adults can attentively encode in parallel different subsets of items defined by color, suggesting that approximate numbers can be stored as a feature of each set (Halberda et al., 2006). Here we show that categorization by color may be spontaneous and likely tuned to salient environmental constancies. Although participants were never instructed to attentively select any non-numerical features of the items, changing the color (or pitch) of the stimuli produced a robust decrease of the expected adaptation magnitude.

Perceptual adaptation is thought to be a functionally adaptive mechanism highlighting new information in the environment by reducing sensitivity to the prevailing stimulus (Benucci et al., 2013; Kohn, 2007; Webster,
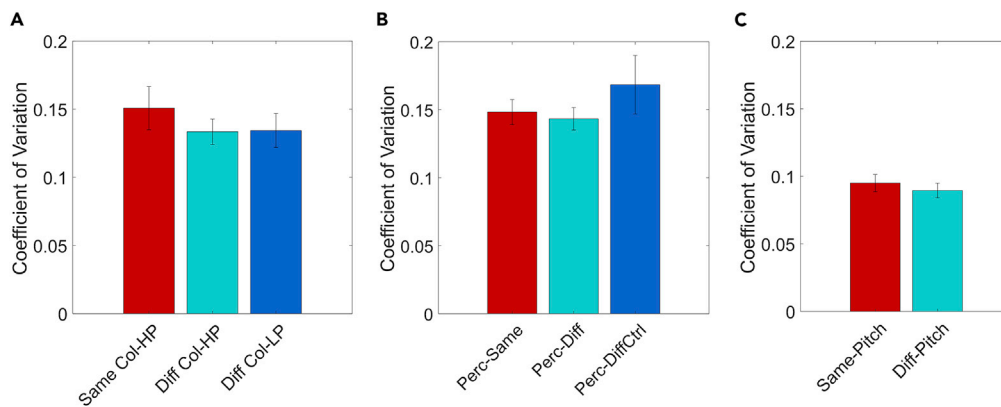
Figure 7. Tasks precision
Coefficients of Variation for the adaptation conditions of first (physical color: A), second (illusory color: B) and third (auditory tones: C) experiment. Bars represent SEM.

2015). For instance, luminance adaptation helps to increase discrimination after abrupt lightness changes through a gradual reentering of neuronal responses towards the average luminance levels of the new environment (Shapley and Enroth-Cugell, 1984). By analogy, adaptation to numerical quantities could act through desensitization of brain responses to the most recurrent numerical range while allowing an increased sensitivity for those numerical ranges laying close to the adapted. However, unlike luminance, real-world environments closely associate numerosity perception with a series of other non-numerical attributes such as color and shape, which are often fundamental to determine the belonging category of a set of stimuli. On this view, it would be crucial for the system to adapt to the sole prevailing category, rather than to all the stimuli in a scene.

The results also speak to the hierarchical level of numerosity adaptation. As the adaptation was color-specific, it would seem to occur at the level of object perception, after the object has bound with color. This would suggest that it does not occur at the very low levels of visual analysis, as has been suggested (Durgin, 2008), but is consistent with higher levels, such as the parietal cortex. That the effect was driven by perceived rather than physical color is further evidence for late analysis. This interpretation is supported by fMRI studies showing adaptation-driven changes in decoding in the parietal cortex (Castaldi et al., 2016), and by studies reporting adaptation-driven changes in neural numerosity selectivity maps within the superior parietal lobe (Harvey and Dumoulin, 2017; Tsouli et al., 2020, 2021a, 2021b), a cortical region also implicated in the categorization of objects through their color (Bramão et al., 2010). It is also consistent with previous psychophysical research showing that adaption incorporates the effects of perceptual grouping (Fornaciai et al., 2016), and with evidence showing that, unlike orientation, numerosity adaptation is robustly affected by the allocation of implicit visuospatial attention during the adaptation period (Grasso et al., 2021a, 2021b). Whatever the precise neural mechanisms involved, our results demonstrate that the visual system retains rapid abilities to shape its responses based on external needs (for recent reviews on visual plasticity see Castaldi et al., 2020; Grasso et al., 2020), and that this process is less coarse than previously thought as it is selectively tuned to salient environmental features.

To conclude, we report here that numerosity perception is highly selective to salient non-numerical features like color and pitch, and that this selectivity is mainly tuned to the perceived rather than the physical characteristics of the stimuli. Our results suggest that the brain could encode numerosity while spontaneously segregating objects into raw salient categories.

Limitations of the study

One clear limitation of the current study is the lack of a continuous measure allowing to change the match between the color/pitch of the test stimulus with respect to that of the adaptor. We here used discrete changes which revealed that numerosity adaptation is mostly absent when the stimuli are not matched. However, it would have been interesting to test whether the adaptation followed a trend as the match became closer, or whether the effect was categorical. Another limitation concerns our third experiment where we found a residual adaptation in the Diff-Pitch condition that may have resulted from the adaptor and test being the same note of the adaptor an octave apart. This can be easily tested using a different note.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Participants
- METHOD DETAILS
  - General procedure and stimuli
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Analysis procedure
  - Statistical analysis

AUTHOR CONTRIBUTIONS

P.A.G., G.A. and G.M.C. conceived the experiments. P.A.G. performed the experiments and analyzed data. All the authors contributed to interpreting data, writing and revising the article.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Aagten-Murphy, D., and Burr, D. (2016). Adaptation to numerosity requires only brief exposures, and is determined by number of events, not exposure duration. J. Vis. *16*, 22. https://doi.org/10.1167/16.10.22.

Anobile, G., Arrighi, R., Togoli, I., and Burr, D.C. (2016). A shared numerical representation for action and perception. Elife 5, e16161. https://doi.org/10.7554/elife.16161.

Anobile, G., Arrighi, R., Castaldi, E., Grassi, E., Pedonese, L., Moscoso, P.A.M., and Burr, D.C. (2018). Spatial but not temporal numerosity thresholds correlate with formal math skills in children. Dev. Psychol. *54*, 458–473. https://doi.org/10.1037/dev0000448.

Anobile, G., Castaldi, E., Moscoso, P.A.M., Burr, D.C., and Arrighi, R. (2020). ''Groupitizing'': a strategy for numerosity estimation. Sci. Rep. *10*, 13436. https://doi.org/10.1038/s41598-020-68111-1.

Anobile, G., Arrighi, R., Castaldi, E., and Burr, D.C. (2021a). A sensorimotor numerosity system.

Trends Cogn. Sci. *25*, 24–36. https://doi.org/10.1016/j.tics.2020.10.009.

Anobile, G., Castaldi, E., Moscoso, P.A.M., Arrighi, R., and Burr, D. (2021b). Groupitizing improves estimation of numerosity of auditory sequences. Front. Hum. Neurosci. *15*, 687321. https://doi.org/10.3389/FNHUM.2021.687321.

Arrighi, R., Togoli, I., and Burr, D.C. (2014). A generalized sense of number. Proc. R. Soc. B Biol. Sci. *281*, 20141791. https://doi.org/10.1098/rspb.2014.1791.

Aulet, L.S., and Lourenco, S.F. (2021). Numerosity and cumulative surface area are perceived holistically as integral dimensions. J. Exp. Psychol. Gen. *150*, 145–156. https://doi.org/10.1037/xge0000874.

Barth, H., Kanwisher, N., and Spelke, E. (2003). The construction of large number representations in adults. Cognition *86*, 201–221. https://doi.org/10.1016/S0010-0277(02)00178-6.

Benucci, A., Saleem, A.B., and Carandini, M. (2013). Adaptation maintains population homeostasis in primary visual cortex. Nat. Neurosci. *16*, 724–729. https://doi.org/10.1038/nn.3382.

Bramão, I., Faísca, L., Forkstam, C., Reis, A., and Petersson, K.M. (2010). Cortical brain regions associated with color processing: an FMRi study. Open Neuroimag. J. *4*, 164. https://doi.org/10.2174/1874440001004010164.

Burr, D., and Ross, J. (2008). A visual sense of number. Curr. Biol. *18*, 425–428. https://doi.org/10.1016/j.cub.2008.02.052.

Butterworth, B. (1999). The Mathematical Brain (London: Macmillan).

Carter, R.C. (1982). Visual search with color. J. Exp. Psychol. Hum. Percept. Perform. *8*, 127–136. https://doi.org/10.1037/0096-1523.8.1.127.

Castaldi, E., Aagten-Murphy, D., Tosetti, M., Burr, D., and Morrone, M.C. (2016). Effects of adaptation on numerosity decoding in the human

brain. Neuroimage 143, 364–377. https://doi.org/10.1016/j.neuroimage.2016.09.020.

Castaldi, E., Lunghi, C., and Morrone, M.C. (2020). Neuroplasticity in adult human visual cortex. Neurosci. Biobehav. Rev. 112, 542–552. https://doi.org/10.1016/j.neubiorev.2020.02.028.

D'Zmura, M. (1991). Color in visual search. Vis. Res. 31, 951–966. https://doi.org/10.1016/0042-6989(91)90203-H.

Dakin, S.C., Tibber, M.S., Greenwood, J.A., Kingdom, F.A.A., and Morgan, M.J. (2011). A common visual metric for approximate number and density. Proc. Natl. Acad. Sci. U S A 108, 19552–19557. https://doi.org/10.1073/pnas.1113195108.

Dehaene, S., and Changeux, J.-P. (1993). Development of elementary numerical abilities: a neuronal model. J. Cogn. Neurosci. 5, 390–407. https://doi.org/10.1162/JOCN.1993.5.4.390.

Dehaene, S., Izard, V., Spelke, E., and Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in western and Amazonian indigene cultures. Science 320, 1217–1220. https://doi.org/10.1126/science.1156540.

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

Farmer, E.W., and Taylor, R.M. (1980). Visual search through color displays: effects of target-background similarity and background uniformity. Percept. Psychophys. 27, 267–272. https://doi.org/10.3758/BF03204265.

Faul, F., Erdfelder, E., Lang, A.G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39, 175–191. https://doi.org/10.3758/BF03193146.

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

Found, A., and Müller, H.J. (1996). Searching for unknown feature targets on more than one dimension: investigating a "dimension-weighting" account. Percept. Psychophys. 58, 88–101. https://doi.org/10.3758/BF03205479.

Grasso, P.A., Gallina, J., and Bertini, C. (2020). Shaping the visual system: cortical and subcortical plasticity in the intact and the lesioned brain. Neuropsychologia 142, 107464. https://doi.org/10.1016/j.neuropsychologia.2020.107464.

Grasso, P.A., Anobile, G., and Arrighi, R. (2021a). Numerosity adaptation partly depends on the allocation of implicit numerosity-contingent visuo-spatial attention. J. Vis. 21, 12. https://doi.org/10.1167/jov.21.1.12.

Grasso, P.A., Anobile, G., Caponi, C., and Arrighi, R. (2021b). Implicit visuospatial attention shapes numerosity adaptation and perception. J. Vis. 21, 26. https://doi.org/10.1167/JOV.21.8.26.

Greenhouse, S.W., and Geisser, S. (1959). On methods in the analysis of profile data. Psychometrika 24, 95–112. 10.1007/BF02289823.

Halberda, J., Sires, S.F., and Feigenson, L. (2006). Multiple spatially overlapping sets can be enumerated in parallel. Psychol. Sci. 17, 572–576. https://doi.org/10.1111/j.1467-9280.2006.01746.x.

Harvey, B.M., and Dumoulin, S.O. (2017). A network of topographic numerosity maps in human association cortex. Nat. Hum. Behav. 1, 1–9. https://doi.org/10.1038/s41562-016-0036.

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

Hauser, M.D., Carey, S., and Hauser, L.B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. Proc. R. Soc. B Biol. Sci. 267, 829–833. https://doi.org/10.1098/rspb.2000.1078.

He, L., Zhang, J., Zhou, T., and Chen, L. (2009). Connectedness affects dot numerosity judgment: implications for configural processing. Psychon. Bull. Rev. 16, 509–517. https://doi.org/10.3758/PBR.16.3.509.

Hollingworth, H.L. (1910). The central tendency of judgment. J. Philos. Psychol. Sci. Methods 7, 461–469. https://doi.org/10.2307/2012819.

Hyde, D.C., and Spelke, E.S. (2012). Spatiotemporal dynamics of processing nonsymbolic number: an event-related potential source localization study. Hum. Brain Mapp. 33, 2189–2203. https://doi.org/10.1002/hbm.21352.

Kim, I., Hong, S.W., Shevell, S.K., and Shim, W.M. (2020). Neural representations of perceptual color experience in the human ventral visual pathway. Proc. Natl. Acad. Sci. U S A 117, 13145–13150. https://doi.org/10.1073/pnas.1911041117.

King, D.L. (1988). Assimilation is due to one perceived whole and contrast is due to two perceived wholes. New Ideas Psychol. 6, 277–288. https://doi.org/10.1016/0732-118X(88)90039-6.

Kleiner, M., Brainard, D.H., Pelli, D.G., Broussard, C., Wolf, T., and Niehorster, D. (2007). What's New in Psychtoolbox-3? Perception. https://doi.org/10.1068/v070821.

Kohn, A. (2007). Visual adaptation: physiology, mechanisms, and functional benefits. J. Neurophysiol. 97, 3155–3164. https://doi.org/10.1152/jn.00086.2007.

Liu, W., Zheng, P., Huang, S., and Cicchini, G.M. (2020). Subitizing, unlike estimation, does not process sets in parallel. Sci. Rep. 10, 15689. https://doi.org/10.1038/s41598-020-72860-4.

Moore, B.C.J., and Gockel, H. (2002). Factors influencing sequential stream segregation. Acta Acust. United Acustica 88, 320–333.

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

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. https://doi.org/10.1016/j.neuron.2006.11.022.

Pomè, A., Caponi, C., and Burr, D.C. (2021). The grouping-induced numerosity illusion is attention-dependent. Front. Hum. Neurosci. 15, 745188. https://doi.org/10.3389/fnhum.2021.745188.

Shapley, R., and Enroth-Cugell, C. (1984). Chapter 9 Visual adaptation and retinal gain controls. Prog. Retin. Res. 3, 263–346. https://doi.org/10.1016/0278-4327(84)90011-7.

Togoli, I., and Arrighi, R. (2021). Evidence for an A-modal number sense: numerosity adaptation generalizes across visual, auditory, and tactile stimuli. Front. Hum. Neurosci. 0, 448. https://doi.org/10.3389/FNHUM.2021.713565.

Tsouli, A., Cai, Y., Ackooij, M., van Hofstetter, S., Harvey, B.M., Pas, S.F. te, Smagt, M.J. van der, and Dumoulin, S.O. (2020). Neural numerosity selectivity changes after visual numerosity adaptation. J. Vis. 20, 486. https://doi.org/10.1167/JOV.20.11.486.

Tsouli, A., Cai, Y., van Ackooij, M., Hofstetter, S., Harvey, B.M., te Pas, S.F., van der Smagt, M.J., and Dumoulin, S.O. (2021a). Adaptation to visual numerosity changes neural numerosity selectivity. Neuroimage 229, 117794. https://doi.org/10.1016/J.NEUROIMAGE.2021.117794.

Tsouli, A., Harvey, B.M., Hofstetter, S., Cai, Y., van der Smagt, M.J., te Pas, S.F., and Dumoulin, S.O. (2021b). The role of neural tuning in quantity perception. Trends Cogn. Sci. 26, 11–24. https://doi.org/10.1016/J.TICS.2021.10.004.

Webster, M.A. (2015). Visual adaptation. Annu. Rev. Vis. Sci. 1, 547–567. https://doi.org/10.1146/annurev-vision-082114-035509.

Whalen, J., Gallistel, C.R., and Gelman, R. (1999). Nonverbal counting in humans: the psychophysics of number representation. Psychol. Sci. 10, 130–137. https://doi.org/10.1111/1467-9280.00120.

Yu, D., Xiao, X., Bemis, D.K., and Franconeri, S.L. (2019). Similarity grouping as feature-based selection. Psychol. Sci. 30, 376–385. https://doi.org/10.1177/0956797618822798.

Zorzi, M., Stoianov, I., and Umiltà, C. (2005). Computational modeling of numerical cognition. In The Handbook of Mathematical Cognition (Psychology Press), pp. 67–83. https://doi.org/10.4324/9780203998045-12.

Zosh, J.M., Halberda, J., and Feigenson, L. (2011). Memory for multiple visual ensembles in infancy. J. Exp. Psychol. Gen. 140, 141–158. https://doi.org/10.1037/a0022925.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw behavioural data	This work	https://datadryad.org/stash/share/QmJP4YXBqUFu1tgpFjh77eJ5EDkrljFv5sBCxdLKUrM
Software and algorithms		
Matlab R2021b	Mathworks	https://it.mathworks.com/products/matlab.html
Psychtoolbox-3	Kleiner et al., 2007	http://psychtoolbox.org/
JASP 0.14.1.0	The JASP Team	http://jasp-stats.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Paolo A. Grasso (paolo.grasso@unifi.it).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Raw behavioral data have been deposited at Dryad repository and are publicly available as of the date of publication. The DOI is listed in the key resources table. This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Participants

To calculate sample size, we performed an *a priori* power analysis using G-Power (Faul et al., 2007). Effect sizes were derived from a previous work using very similar methods and which supported the expectation for quite large effects (Grasso et al., 2021b). For the first experiment, the analysis indicated that a total sample of 12 participants would be needed to detect large effects (f = 0.40) with 80% power using a repeated measure ANOVA with an alpha level of 0.05. For second and third experiments, the analysis indicated that a total sample of 13 participants would be needed to detect large effects (d = 0.85) with 80% power using a dependent sample t-test with an alpha level of 0.05.

Twelve participants (four practiced participants) took part in the first experiment, and thirteen in the second and third experiments (three and two practiced participants for the second and third experiment respectively). Two participants from the third experiment had outlier performances (above or below twice the upper or lower quartile) and were thus excluded. There were no outliers in the first and second experiments. The final samples comprised a total of twelve participants for the first experiment (mean age: 28.4 years, sd: 4.7 years; 5 males), thirteen for the second experiment (mean age: 26.8 years, sd: 4.3 years; 5 males) and eleven for the third experiment (mean age: 25.1 years, sd: 3.8 years; 4 males). Four participants took part in all the experiments, two participants in both the first and second experiments and four in both second and third experiments. All participants provided written informed consent before taking part to the experiments. The research was approved by the local ethics committee (*"Commissione per l'Etica della Ricerca"*, University of Florence, 7th July 2020, n. 111).

METHOD DETAILS

General procedure and stimuli

All the experiments were performed in a dimly lit, low-noise environment with participants sitting 57 cm from the monitor of a laptop pc. In the first and second experiments we measured the magnitude of numerosity adaptation with a two-alternative comparison task. Stimuli comprised clouds of coloured

non-overlapping random dots (diameter: 0.4°; single dot area: 0.12 cm^2) drawn within a virtual 8° diameter circle with centre 10° left or right of a central fixation point. In the first experiment, colour of test and reference randomly varied between blue (rgb code: 80 130 240; x = 0.19, y = 0.18, L = 23.1 cd/m^2), green (rgb code: 40 152 90; x = 0.27, y = 0.46, L = 22.7 cd/m^2), yellow (rgb code: 185 140 50; x = 0.45, y = 0.44, L = 23.3 cd/m^2) or purple (rgb code: 200 100 200; x = 0.32, y = 0.21, L = 22.6 cd/m^2) while colour of the adaptor was kept fixed (seven participants were presented with a blue adaptor and five with a green adaptor), and all stimuli were presented on a textured grey background (with grey pixels having independent and random luminance between 22.0 and 23.2 cd/m^2; x = 0.33, y = 0.34). In the second experiment we used yellow dots (rgb code: 255 255 0; x = 0.42, y = 0.50, L = 41 cd/m^2) presented on a red (rgb code: 254 112 26; x = 0.55, y = 0.39, L = 32 cd/m^2) and cyan (rgb code: 113 251 254; x = 0.24, y = 0.32, L = 41 cd/m^2) horizontal striped background (strip width 0.3 mm, 2 pixels) while in the control conditions we used lime dots (rgb code: 0 255 0; x = 0.30, y = 0.60, L = 40 cd/m^2) a colour having different hue levels but same saturation and luminance of yellow (hsl yellow: 60°, 100%, 50%; hsl green: 120°, 100%, 50%).

Participants were asked to discriminate which of test and reference was more numerous using right and left arrows of the keyboard. A typical trial began with an adaptation period (2000 ms) in which an adaptor (a cloud of 48 elements) appeared on the left side of a central fixation cross. This was followed by an ISI (500 ms) which preceded the simultaneous presentation (200 ms) of test (a cloud of dots randomly varying between 12, 14, 16, 18, 21, 24, 28, 32, 36, 42 and 48 elements) and reference (a cloud of 24 dots), respectively presented left and right of central fixation. Baseline was measured with comparison trials not preceded by an adaptation period. The number of trials for each tested numerosity followed a log-normal distribution except for adaptation trials of the first experiment where the tested numerosities were equally represented across trials.

In the third experiment, we measured numerosity adaptation in audition with an estimation task. Stimuli comprised sequences of brief tones (50 ms), ramped on and off with 10 ms raised-sine ramps. They were evenly distributed throughout the period, with a random inter-tone jitter (20% of the exact inter-tone distance). Participants were asked to estimate the perceived numerosity of the test sequence and typing it on the keyboard. A typical trial began with an adaptation sequence composed of either 48 tones (∼8 Hz) or 12 tones (∼2 Hz) presented within a 6000 ms period. This was followed by an ISI (1000 ms) and a test sequence (3000 ms) which could randomly contain 10 to 20 pulses. To help participants correctly distinguish between adaptation and test sequences, a central circle turned from red to green whenever the sequence turned from adaptation to test. For each condition and each tested numerosity a total of six trials was presented. The use of an estimation rather than discrimination task was mainly motivated by the difficulties in using 2AFC for temporal sequences. Previous studies have shown that estimation works well (Anobile et al., 2016; Arrighi et al., 2014), and that thresholds using the two techniques are highly correlated (Anobile et al., 2018).

All stimuli were generated and presented with PsychToolbox 3 routines (Kleiner et al., 2007) in Matlab 2020b (The Mathworks, Inc., http://mathworks.com).

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis procedure

In the first and second experiments, an individual outlier detection procedure excluded trials with reaction times above or below three standard deviations from the average reaction times of each numerosity in each condition (baseline and adaptation). This led to discard, on average, 1.28% of trials for baseline and 1.98% of trials for adaptation in the first experiment and 1.41% of trials for baseline and 1.52% of trials for adaptation in the second experiment. Afterwards, for each participant and each condition, the proportion of trials in which the test stimulus was perceived more numerous than the reference was plotted as a function of tested values and fitted with a cumulative gaussian function. The 50% point of the function defines the point of subjective equality (PSE), the physical numerosity necessary for the test adapted stimulus to perceptually match the numerosity of the unadapted reference. Numerosity adaptation magnitudes were calculated for each condition by subtracting the PSEs at baseline from PSEs during adaptation, and normalizing by the numerosity of the reference:

$$\frac{PSE_{adaptation} - PSE_{baseline}}{N} \times 100\%$$

Task precision was measured using coefficient of variation calculated as the ratio between standard deviation of the cumulative gaussian function and numerosity of the reference.

In the third experiment, average responses to tested numerosities were computed for each experimental condition and each participant. Differences between estimation values obtained in the 2 Hz and the 8 Hz adaptation conditions were normalized by numerosity and fitted with a linear regression function anchored at zero after excluding the two extreme values (10 and 20) which were discarded to avoid edge effects. The slope of the function gave adaptation magnitude. Task precision was measured using coefficient of variation calculated as the ratio between standard deviation and physical numerosity averaged across numerosities and adaptation conditions (2 Hz and 8 Hz).

Statistical analysis

Analyses were performed using JASP 0.14.1.0. Repeated measures ANOVAs and paired sample t-tests were used to compare different experimental conditions. To compensate for violations of sphericity, Greenhouse-Geisser corrections were applied whenever appropriate (Greenhouse and Geisser, 1959) and corrected p-values (but uncorrected degrees of freedom) are reported. Post-hoc comparisons were performed using Bonferroni correction.