Different reaction-times for subitizing, estimation, and texture

Antonella Pomè
Department of Neuroscience, Psychology, Pharmacology, and Child Health, University of Florence, Florence, Italy

Giovanni Anobile
Department of Developmental Neuroscience, Stella Maris Scientific Institute, Calambrone, Pisa, Italy

Guido Marco Cicchini
Institute of Neuroscience, National Research Council, Pisa, Italy

David Charles Burr
Department of Neuroscience, Psychology, Pharmacology, and Child Health, University of Florence, Florence, Italy
Institute of Neuroscience, National Research Council, Pisa, Italy
School of Psychology, University of Western Australia, Perth, Australia

Humans can estimate and encode numerosity over a large range, from very few items to several hundreds. Two distinct mechanisms have been proposed: subitizing, for numbers up to four and estimation for larger numerosities. We have recently extended this idea by suggesting that for very densely packed arrays, when items are less segregable, a third "texture" mechanism comes into play. In this study, we provide further evidence for the existence of a third regime for numerosity. Reaction times were very low in the subitizing range, rising rapidly for numerosities greater than four. However, for tightly packed displays of very high numerosities, reaction times became faster. These results reinforce the idea of three regimes in the processing of numerosity, subitizing, estimation, and texture.

Introduction

Kaufman and Lord (1949) coined the term *subitizing* (from the Latin *subitus* meaning "immediately") to refer to the capacity to enumerate accurately and "immediately" small quantities of items. The primary evidence for subitizing was the characteristic form of the reaction time curve: For up to four items, reaction times are almost constant, increasing by about 40–100 ms for every dot, while for large numbers they increase by 250–350 ms per dot, leading to a clear change in the curve slope (Atkinson, Campbell, & Francis, 1976; Fernberger (1921); Jevons, 1871; Kaufman & Lord, 1949; Mandler & Shebo, 1982; Oberly, 1924). Many studies have reinforced these findings by showing that both reaction times and precision differ from small to large sets of items. For example, with a number-naming task over the range of one to eight (grain of one) or 10 to 80 (grain of 10), Revkin, Piazza, Izard, Cohen, and Dehaene (2008) found that precision was higher and reaction times faster in the range one to four than 10 to 40. Also Choo and Franconeri (2014) showed that comparing two versus three elements was much faster and more accurate than comparing 20 versus 30. Furthermore, individual subitizing capacity and numerosity comparison thresholds were not correlated (Piazza, Fumarola, Chinello, & Melcher, 2011; Revkin et al., 2008). These distinct patterns indicate the involvement of different processes.

Not all studies confirm the idea that subitizing and estimation are driven by completely separate processes. For example, Sengupta, Bapiraju, and Melcher (2017) showed that a single flexible network can allow different number ranges to emerge through a self-organization of the same network. Similarly, Balakrishnan and Ashby (1992) have questioned the existence of a single mechanism for subitizing, showing a lack in discontinuity in reaction time data inside the subitizing range. Depriving attentional resources, sensory precision within the subitizing range (greater than five) becomes indistinguishable from the one for higher numerosity, suggesting two separate but over-

Citation: Pomè, A., Anobile, G., Cicchini, G. M., & Burr, D. C. (2019). Different reaction-times for subitizing, estimation, and texture. *Journal of Vision*, *19*(6):14, 1–9, https://doi.org/10.1167/19.6.14.

https://doi.org/10.1167/19.6.14 Received January 15, 2019; published June 13, 2019 ISSN 1534-7362

lapping processes (Anobile, Turi, Cicchini, & Burr, 2012; Burr, Turi, & Anobile, 2010).

Despite the mounting evidence for the existence of a *number sense*, many have questioned whether observers estimate numerosity per se, or rather infer it from other visual properties, such as density and area (for reviews see Anobile, Cicchini, & Burr, 2016; Burr, Anobile, & Arrighi, 2017). This research has led to the suggestion that while moderate ranges of numerosities are processed directly, higher numerosities may be estimated by a different mechanism, linked to the perception of *texture density*—the apparent density of a dense pattern comprising unsegregable structures. The evidence for different mechanisms comes largely from the fact that numerosity discrimination thresholds tend to obey Weber's law, increasing with perceived numerosity; there is a point at which Weber fractions cease to be constant but decrease with the square root of numerosity (Anobile, Cicchini, & Burr, 2014; Anobile, Turi, Cicchini, & Burr, 2015). Two psychophysical regimes suggest two separate mechanisms: one for estimating numerosity at moderate densities, the other for estimating the density of textures at higher densities. The transition from estimation mechanisms (following Weber's law) to texture-like mechanisms (following a square-root law) depends on several factors, including eccentricity and transitioning earlier in the periphery than in central vision, according to crowding-like rules (Anobile et al., 2015).

The idea of separate systems for numerosity and texture density may reconcile seemingly conflicting evidence for or against an approximate number system (Gebuis, Kadosh, & Gevers, 2016). Several authors have suggested that number could be derived as the product of area and density (Dakin, Tibber, Greenwood, Kingdom, & Morgan, 2011; Durgin, 2008; Morgan, Raphael, Tibber, & Dakin, 2014; Tibber, Greenwood, & Dakin, 2012), while other research suggests that number rather than density is sensed (Cicchini, Anobile, & Burr, 2016, 2019). One possible explanation for the discrepancy is that both processes operate, direct estimation for sparse displays and texture mechanisms for dense displays. Indeed, the spontaneous emergence of numerosity gives way to density-like mechanisms at higher densities (Anobile et al., 2014; Cicchini et al., 2016, 2019). Similarly, interactions between area and density on number judgments, often reported in the literature (Dakin et al., 2011; Morgan et al., 2014; Tibber et al., 2012), are much reduced in sparse displays (figure 8 in Anobile, Cicchini, & Burr, 2016). Other evidence for separate numerosity and texture systems comes from the connected-dot numerosity illusion: Connecting adjacent dots within a cloud of dots with thin lines to produce "dumbbells" reduces drastically the apparent numerosity of the stimuli. However, the connectivity effect is greatly reduced at high densities, consistent with the suggestion of separate mechanisms at high dot densities (Anobile, Cicchini, Pomè, & Burr, 2017).

Zimmermann and Fink (2016) measured number adaptation for low and high numerosities and showed that both the magnitude of adaptation and the spatial spread differed considerably in the two conditions, providing further evidence that processing of low and high numbers could involve different mechanisms. The results also suggest that processing of low and high numbers could involve different receptive field sizes, with larger receptive fields for low than high number processing (Zimmermann & Fink, 2016). Electroencephalogram (EEG) studies demonstrated distinctive neural signatures in event-related potential (ERP) signals for very low numerosities (up to 4), intermediate numerosity, and very high numerosities (from 100 to 400), suggesting again three regimes for numerosity encoding and that when objects become too close to be segregated, different mechanisms come into play (Fornaciai & Park, 2017; Park, DeWind, Woldorff, & Brannon, 2016).

Classical studies describing the dichotomy between subitizing and estimation are based mainly on reaction times while those proposing separate mechanisms between estimation and texture-density employed sensory thresholds. The current study aims to investigate reaction times for the two different regimes of number and density perception. We hypothesized that if reaction times vary in the same way that thresholds do, they should follow the same trend, and begin to decrease for high numerosities.

Methods

Participants

Sixteen subjects (nine male, $M = 28$ years old; $SD = 2.50$) with normal or corrected-to-normal vision participated. Eight (five males, $M = 28$ years old; $SD = 1.70$) completed the central viewing condition; the other nine (one shared, four males, $M = 28$ years old; $SD = 2.4$) were tested in the peripheral condition. All participants gave written informed consent. Experimental procedures were approved by the local ethics committee (Comitato Etico Pediatrico Regionale—Azienda Ospedaliero-Universitaria Meyer—Firenze) and are in line with the Declaration of Helsinki.

Stimuli

Stimuli were generated with the Psychophysics Toolbox (Brainard, 1997) and presented at a viewing

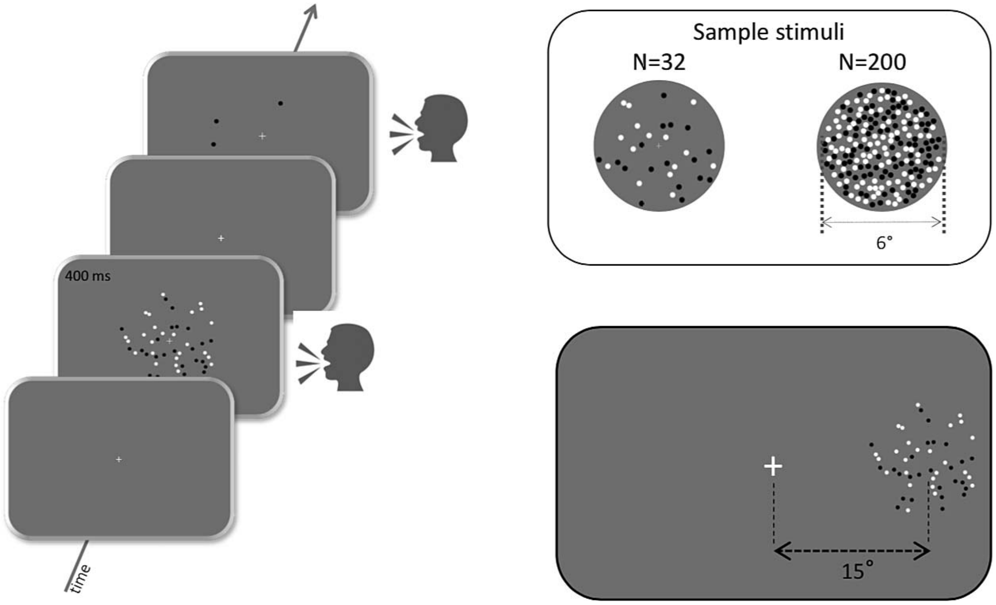

Figure 1. Illustration of the stimulus sequence for the first experiment. Each trial started with a cloud of dots presented for 400 ms either in central viewing or at 15° of eccentricity (lower side of the figure). The subjects were asked to voice the numerosity of the patch. Upper side: Sample stimuli ($n = 32$ and $n = 200$).

distance of 57 cm on a 23-in. LCD Acer monitor (resolution = 1,920 × 1,080 pixels; refresh rate = 60 Hz), run by a Macintosh laptop (MacBookAir, Apple, Cupertino, CA). Half of the dots were white, the other half black, to avoid luminance being a cue for numerosity. Each dot had a diameter of 10 pixels (0.25°) and were always separated from each other by at least 0.25°. Dots were randomly displayed within a virtual circular patch with a diameter of 6°.

Procedure and data analysis

We measured accuracy and precision of numerosity estimation for a cloud of nonoverlapping dots confined to a circular region displayed at the middle of the screen (Figure 1).

All trials started with a central fixation cross presented on a gray screen, on which subjects maintained fixation throughout the trial. On initiation by the experimenter, a cloud of dots was presented for 400 ms, and subjects called out how many dots they had seen. The computer detected the onset of the vocal response, from which reaction times were computed, and the experimenter recorded the responses on a keyboard. Reaction times were measured by voice onset, and averaged for each condition. Vocal responses were recorded using a Psychtoolbox function on MATLAB (MathWorks, Natick, MA), which recorded audio data from the internal microphone of the computer. Subjects were instructed to call out swiftly and clearly the number, which nearly always yielded neat soundtracks to estimate reaction times from (checked manually by experimenter). Sound thresholds for detection of responses were set to about one-tenth of the typical vocal intensity, which excluded the rare environmental sounds in the experimental room. In the event of coughing, unclear utterances, or particularly loud noises, the experimenter tagged the response, which was excluded from analyses. Trials with response times outside ±2 *SD*s from the mean of each subject were considered outliers and also excluded from the analysis (a total of 354 trials, 9.7% in total).

Subjects were instructed to respond as quickly and accurately as possible. All participants started with a training phase of 10 trials in which they were shown a subset of the stimuli used in the actual experiment (numerosities of roughly one half, one quarter, one fifth, and one eighth of the maximal stimulus in the range) and were given feedback of the actual numerosity of the stimulus. In the main experiment, 31 numerosities were used, roughly equispaced on a logarithmic scale ($N = 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 13, 15, 17, 19, 22, 26, 29, 34, 39, 45, 52, 60, 69, 79, 91, 105, 121, 139, 160, 184, 212$). Three 62-trial sessions were run (each numerosity was presented six times). Importantly, subjects were never instructed about the actual range used to avoid edge effects (Jazayeri & Shadlen, 2010; Poulton, 1973; Teghtsoonian & Teghtsoonian, 1978), which might contaminate precision measures.

To test the prediction that the switching from estimation to density range depends on eccentricity, we tested additional subjects on a different range of numerosities from the previous experiment, asking

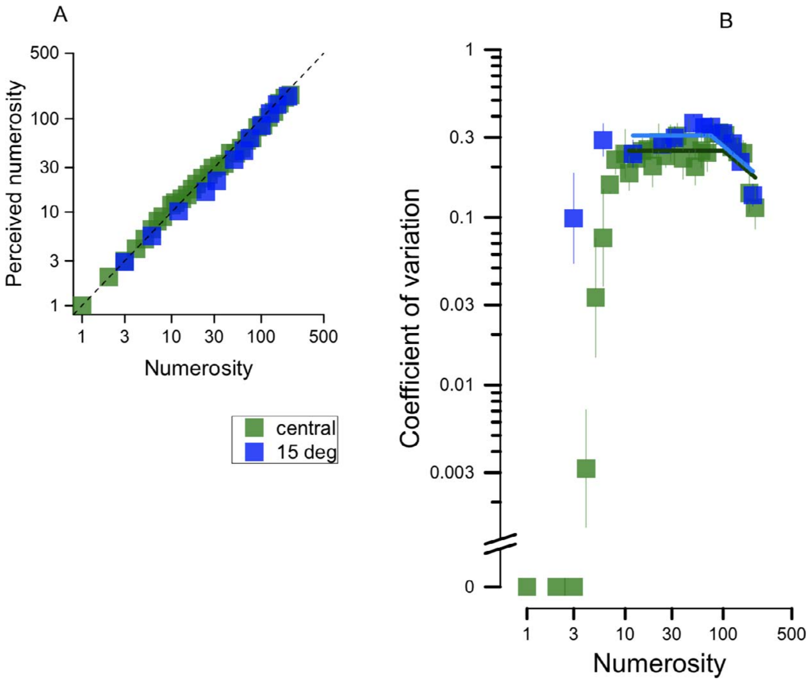

Figure 2. Accuracy, precision, and reaction times for verbal numerosity estimation. (A) Average perceived numerosity for central (green squares) and 15° peripheral viewing (blue squares). Eccentricity had no effect on average accuracy. (B) Means of CV (standard deviation divided by numerosity) as a function of numerosity, for the two conditions (central-green and 15° eccentricity-blue). Numerosity ranged from one to 212 for the central presentation condition and from three to 200 for the 15° condition. Continuous lines are two-limb linear functions (slope 0 and –0.5 on log coordinates) that best fit the data.

them to report the estimated number of dots confined within circular regions alternated randomly on either side of a central fixation cross at 15° of eccentricity. We tested twelve different numerosities ($N = 3$, 6, 12, 24, 32, 50, 64, 75, 100, 125, 150, 200). Three 72-trial sessions were run for each condition.

Data were analyzed separately for each subject and numerosity, and then averaged over subjects. Responses were pooled for each condition and numerosity, from which the mean and standard deviation were estimated. The mean reflects systematic biases in judgments, while the standard deviation provides an estimate of response precision which, normalized by the numerosity for each condition, yields the coefficient of variation (CV).

Results

We measured numerosity estimation thresholds and reaction times by asking subjects to call out as quickly and accurately as possible the numerosity of a briefly presented cloud of dots (see Figure 1 and the Stimuli and procedure section). We tested a large range of numerosities at two stimulus eccentricities (0° and 15°), in two separate conditions. Figure 2A shows the average perceived numerosity for each physical numerosity, averaged over subjects. In general, these estimates were quite accurate (bias-free) for both conditions, following the physical numerosity tested (dashed diagonal), and showing no sign of edge effects (Jazayeri & Shadlen, 2010; Poulton, 1973; Teghtsoonian & Teghtsoonian, 1978), indicating that training the subjects with a smaller range of stimuli was sufficient to obtain unbiased estimates. Figure 2B shows average numerosity estimation precision, expressed as the CV (standard deviation normalized by numerosity), as a function of numerosity, separately for the two eccentricity conditions. As we previously demonstrated (Anobile et al., 2014), the CV remains stable (following Weber's law) over the low numerosity range, then starts to decrease. To estimate where thresholds switched from one regime to another, we fitted the data with a two-limb piecewise linear function, with the slope of the first limb fixed at 0 and that of the second fixed at –0.5 (on logarithmic coordinates), as our previous studies showed that the decrease followed an approximate square-root law (Anobile et al., 2014; Anobile et al., 2015). The goodness of fit had a coefficient of determination (R^2) of 0.41 and 0.63 for the data of central and peripheral viewing, respectively.

Previous work (Anobile et al., 2015) has shown that the break in the two-limb function determining the point where Weber's law gives way to the square-root

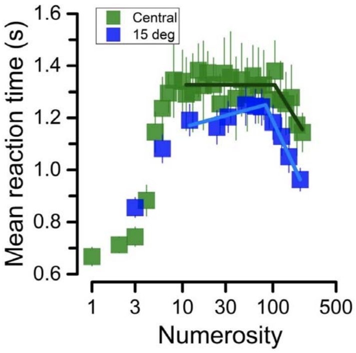

Figure 3. Reaction times for verbal numerosity estimation. Mean of reaction times (in seconds) as a function of numerosity, for the two conditions (central-green and 15° eccentricity-blue). Continuous lines are two-limb linear functions (both slopes free to vary on semi-log coordinates) that best fit the data.

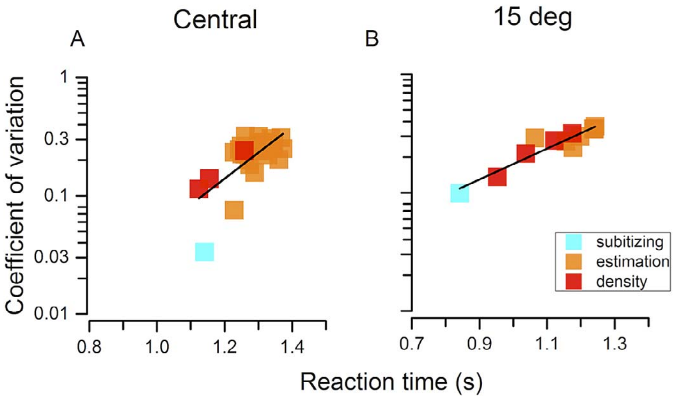

Figure 4. Correlations between reaction times and coefficient of variation. (A–B) Correlations of mean reaction times with CVs separately for the two different conditions (central and 15° of eccentricity). Different colors refer to different numerosity ranges, calculated from the point in which the slopes fall. Black continuous lines are linear functions that best fit the data.

law depends on eccentricity, occurring at lower numerosities in the periphery than in central vision. In the current study, the change in psychophysical regime occurs at 101 elements in the central condition (green square) and at 75 elements in the 15° eccentricity condition. This extends the previous study by showing the effect of eccentricity on numerosity judgments also occurs when subjects are required to estimate numerosities, rather than discriminate them with a two-alternative forced-choice task. Also replicating our previous work, the effect of eccentricity is different for numerosities either side of the knee point. For lower numerosities (i.e., $9 < N < 68$), the CVs are not only flat but also dependent on eccentricity, with better performance for central viewing conditions (0.24 vs. 0.3, Kolmogorov-Smirnov test = 0.29, $p = 0.007$). For numerosities above the turning points (i.e., $n > 105$) the slope in the two viewing conditions are statistically indistinguishable (0.206 vs. 0.200, KS-test = 0.14, $p = 0.85$).

Figure 3 plots reaction times for the same data averaged over all subjects as a function of numerosity, for both central and eccentric viewing. Mean reaction times increase with dot number, from about 700 ms in the subitizing range (one to four dots), to around 1300 ms over the range of 10–100 dots (for central viewing), first sharply then more gradually. They then decline sharply for higher numerosities. To verify that there is a significant modulation of reaction times outside of the subitizing range, we fitted a two-limbed linear function and pitted it against a simple linear regression. For eccentric presentation, we found that a two-limbed function predicted the data much better ($R^2 = 0.90$ vs. 0.45). To verify that this is statistically significant, we bootstrapped the data 10,000 times and compared the residuals of the two fits. Since the two-limbed function has two more degrees of freedom, we fixed two parameters (the peak of the curve and the rising part) to make a fair comparison with the linear model. The analysis revealed that even when the two-limbed function was run with only two free parameters, it still yielded a better R^2 in virtually all cases ($p < 0.0001$). Also, in the case of central viewing, the R^2 of the two-limbed function much better predicted the data ($R^2 = 0.45$ vs. 0.25). Bootstrap resampling demonstrated that also this difference in fitting performance is statistically significant ($p = 0.044$).

The fits were similar for the two eccentricities, except that the change in slope at 15° eccentricity occurred at a lower numerosity. The reaction time results parallel those of the CV, in showing a steep decrease as the numerosity become dense. When fitting average data, the points where the curve sharply changes slope are similar for CVs and reaction times: 101 and 105 respectively for CVs and reaction times in the central viewing condition; 75 and 82 elements for CVs and reaction times in the peripheral viewing.

Bootstrap *t* tests confirmed that the difference of the knee points on reaction time data was significantly different between the two presentation conditions (114 ± 15 central vs. 84 ± 11 peripheral, sign test $p = 0.049$). On the other hand, the slopes of the curve fits, both before and after the knee point, were statistically indistinguishable (before: -0.01 ± 0.04 central vs. 0.09 ± 0.07 peripheral, sign test $p = 0.10$; after: -0.63 ± 0.25 central vs. -0.76 ± 0.14 peripheral, bootstrap $p > 0.4$).

Given the similarity in the pattern of results for precision and reaction times, we examined the correlation between the two measures across the whole dataset (Figure 4A and B). In order to obtain a more robust fit, we excluded the first four numerosities of the central condition, where estimation was nearly perfect, and quantification depends crucially on the occurrence of rare errors. The correlation between the two was positive for both the conditions ($r = 0.67$ and $r = 0.94$ respectively for central and 15° of eccentricity; $p < 0.001$). This correlation shows that there is no speed–accuracy tradeoff, and that the improvements of performance in the density-perception regime go together with a speeding of responses.

Inspection of the reaction time curves of Figure 3 suggests that they were generally faster in peripheral than central viewing. This difference was significant with a bootstrap sign test ($p < 0.05$). We have no clear explanation for this difference. However, the two conditions were measured in different sessions, with partly different groups of participants. It is possible that different strategies were employed, resulting in a faster reaction time for the peripheral stimuli.

Discussion

This study provides further evidence for the existence of three different regimes for numerosity perception. We asked subjects to estimate numerosity as quickly and accurately as possible, and measured both reaction times and precision. The precision estimates confirmed our previous study showing that while Weber's law describes well the results for relatively sparse numerosities, it gives way to a square-root law for higher numerosities (Anobile et al., 2014). Importantly, however, reaction times, which have traditionally defined subitizing, also followed the same trend. They increased from the subitizing range as numerosities increased, but only up to a point, then decreased, in a similar manner to the CV.

The point where reaction times started to decrease was similar to where Weber's law gives way to a square-root law (a possible signature of the texture–density system). Indeed, the two measures—reaction times and CVs—correlate positively with each other, with no "speed–accuracy trade-off." These data extend and reinforce our previous findings (Anobile et al., 2014; Anobile et al., 2015) indicating that the lowering of precision at high numerosities is genuine and does not come at the expenses of higher reaction times.

As the point where CV gives way to a square-root law depends on eccentricity (Anobile et al., 2015), we measured reaction times at two different eccentricities. As predicted, reaction times began to decrease at a lower numerosity for the more eccentric stimuli. Again, the correlation between reaction times and CVs for all ranges was strong at this eccentricity.

In this study, we used stimuli of the same area (6°). As density and numerosity are confounded with constant area stimuli, we cannot be certain that the transition from one region to the other is determined by increased numerosity, density, or both. However, we did test this by varying area in our previous paper (Anobile et al., 2014) and found that the transition was governed primarily, but not totally, by density.

Other evidence for separate mechanisms for small, medium, and high numerosities comes from studies showing that numerosity discrimination thresholds for intermediate (N24) numerosities correlate well with children's math scores, but those for very low (subitizing) or high (N250) numerosities do not (Anobile, Arrighi, & Burr, 2019; Anobile et al., 2018; Anobile, Castaldi, Turi, Tinelli, & Burr, 2016). Given the importance of mathematical skills in our society, understanding the three regimes of numerosity perception has practical as well as a theoretical relevance. It would seem that only numerosity estimation, not subitizing or density discrimination, may serve as "start-up tools" for acquiring mathematical skills (Piazza, 2010).

Recent experiments have suggested that numerosity, rather than density or area, is discriminated spontaneously (Cicchini et al., 2016, 2019). In one task subjects were asked to detect the odd stimulus in a field of three (odd-one-out: Cicchini et al., 2016); in the other to reproduce the essential characteristics of a dot array by physical matching (Cicchini et al., 2019). In both cases, errors were minimal when numerosity changed, and maximal for changes in area and density that led to no changes in numerosity. However, only for moderate dot densities was performance strongly dominated by dot numerosity rather than density or patch area: At high numerosities, where the elements where tightly packed, area and density, rather than numerosity, dominated the match, which is further evidence for the action of separate mechanisms.

It should be noted that we are not proposing the existence of three completely independent and non-overlapping mechanisms for subitizing, estimation, and texture. On the contrary, we believe, with good evidence, that there is considerable overlap of the mechanisms, but the most sensitive will dominate a particular task. For very small numbers, the attentional-based subitizing system is the most sensitive, and will dominate. However, under conditions of attentional deprivation, this system cannot operate, and numerosity estimates will be based on the estimation system. Evidence for this comes from studies showing that during dual task (which inhibits subitizing), adaptation (a signature of the estimation system) also occurs for numerosities in the subitizing range (Anobile et al., 2017). Similarly, the estimation system (which obeys Weber's law) is typically more sensitive at lower numerosities, giving way to a density-based system (following a square-root law) at higher densities. But under conditions of area mismatch, where one system cannot act as a proxy for the other, it can be shown that both systems cover a very wide range (Anobile et al., 2014). The current data expand on this suggestion by showing that advantages in precision are accompanied by faster response times: Again, it is the faster response that will dominate the reaction time measures, even though both systems may be activated.

Having demonstrated the existence of three regimes for numerosity discrimination, it is reasonable to ask what the mechanisms are behind each. Subitizing clearly relies on attentional mechanisms (Burr et al., 2010; Vetter, Butterworth, & Bahrami, 2008), which can operate on up to four to five elements. The high-density range is probably subserved by a mechanism sensitive to texture, which could be based on several properties, such as the Fourier energy or the statistical properties of interdot distance, both leading to a square-root dependence on numerosity (Anobile et al., 2014). The mechanism for numerosity discrimination is harder to define. There exist several models, such as the classic model of Dehaene and Changeux (1993), which is constrained to have Weber law behavior (quite different from models of texture discrimination). Similarly, it has been shown that numerosity discrimination, with Weber law properties, can emerge spontaneously from neural networks designed for other, nonnumerical functions (Hannagan, Nieder, Viswanathan, & Dehaene, 2017; Stoianov & Zorzi, 2012). A common principle of all of these models is a segregation stage, in which inhibitory surround regions play an important role (Sengupta, Surampudi, & Melcher, 2014; Stoianov & Zorzi, 2012). When segregation is impossible (for example at high densities), these purported mechanisms will clearly break down. However, while these models all are quite successful up to a point, it is clear that much more work is needed in uncovering the precise mechanisms responsible for estimation of numerosities, higher than those that can be subitized, but too low to define texture.

To recap, this study reinforces previous work suggesting separate (but largely overlapping) mechanisms for encoding subitizing, numerosity, and density. Not only do thresholds change from being very low, then proportional to numerosity to increasing with the square root of numerosity, but reaction times also follow this pattern.

Keywords: subitizing, density perception, numerosity perception, approximate number system, numerical cognition

Acknowledgments

This research was funded by the Italian Ministry of Health and by Tuscany Region under the project "Ricerca Finalizzata", grant number GR-2013-02358262 to GA, by the European Research Council FP7-IDEAS-ERC (grant number 338866 "Early Sensory Cortex Plasticity and Adaptability in Human Adults - ECSPLAIN"), by European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement number 801715-PUPILTRAITS and grant number 832813 "Spatio-temporal mechanisms of generative perception" GenPercept), by Italian Ministry of Education, University, and research under the PRIN2017 program (grant number 2017XBJN4F "EnvironMag" and grant number 2017SBCPZY "Temporal context in perception: serial dependence and rhythmic oscillations").

Commercial relationships: none.

Corresponding author: David C. Burr.
Email: davidcharles.burr@unifi.it.
Address: Department of Neuroscience, Psychology, Pharmacology and Child Health, University of Florence, Florence, Italy.

References

Anobile, G., Arrighi, R., & Burr, D. C. (2019). Simultaneous and sequential subitizing are separate systems, and neither predicts math abilities. *Journal of Experimental Child Psychology*, *178*, 86–103.

Anobile, G., Arrighi, R., Castaldi, E., Grassi, E., Pedonese, L., Moscoso, P. A. M., . . . Burr, D. C., (2018). Spatial but not temporal numerosity thresholds correlate with formal math skills in children. *Developmental Psychology*, *54*(3), 458–473.

Anobile, G., Castaldi, E., Turi, M., Tinelli, F., & Burr, D. C. (2016). Numerosity but not texture-density discrimination correlates with math ability in children. *Developmental Psychology*, *52*(8), 1206–1216.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2014). Separate mechanisms for perception of numerosity and density. *Psychological Science*, *25*(1), 265–270.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2016). Number as a primary perceptual attribute: A review. *Perception*, *45*(1–2), 5–31.

Anobile, G., Cicchini, G. M., Pomè, A., & Burr, D. C. (2017). Connecting visual objects reduces perceived numerosity and density for sparse but not dense patterns. *Journal of Numerical Cognition*, *3*, https://doi.org/10.5964/jnc.v3i2.38.

Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. C. (2012). The effects of cross-sensory attentional demand on subitizing and on mapping number onto space. *Vision Research*, *74*, 102–109.

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*(5):4, 1–12, https://doi.org/10.1167/15.5.4. [PubMed] [Article]

Atkinson, J., Campbell, F. W., & Francis, M. R. (1976). The magic number 4 ± 0: A new look at visual numerosity judgements. *Perception*, *5*(3), 327–334.

Balakrishnan, J. D., & Ashby, F. G. (1992). Subitizing—Magical numbers or mere superstition. *Psychological Research*, *54*(2), 80–90.

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

Burr, D. C., Anobile, G., & Arrighi, R. (2017). Psychophysical evidence for the number sense. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1740): 20170045, https://doi.org/10.1098/rstb.2017.0045.

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision*, *10*(6):20, 1–10, https://doi.org/10.1167/10.6.20. [PubMed] [Article]

Choo, H., & Franconeri, S. L. (2014). Enumeration of small collections violates Weber's law. *Psychonomic Bulletin & Review*, *21*(1), 93–99.

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

Cicchini, G. M., Anobile, G., & Burr, D. C. (2019). Spontaneous representation of numerosity in typical and dyscalculic development. *Cortex*, *114*, 151–163, https://doi.org/10.1016/j.cortex.2018.11.019.

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences, USA*, *108*(49), 19552–19557.

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities—A neuronal model. *Journal of Cognitive Neuroscience*, *5*(4), 390–407.

Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Current Biology*, *18*(18), R855–R856.

Fernberger, S. W. (1921). A preliminary study of the range of visual apprehension. *The American Journal of Psychology*, *32*, 121–133, https://doi.org/10.2307/1413479.

Fornaciai, M., & Park, J. (2017). Distinct neural signatures for very small and very large numerosities. *Frontiers in Human Neuroscience*, *11*, https://doi.org/10.3389/fnhum.2017.00021.

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.

Hannagan, T., Nieder, A., Viswanathan, P., & Dehaene, S. (2017). A random-matrix theory of the number sense. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1740): 20170253, https://doi.org/10.1098/rstb.2017.0253.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, *13*(8), U1020–U1152.

Jevons, W. S. (1871, February 9). The power of numerical discrimination. *Nature*, *3*, 281–282.

Kaufman, E. L., & Lord, M. W. (1949). The discrimination of visual number. *The American Journal of Psychology*, *62*(4), 498–525.

Mandler, G., & Shebo, B. J. (1982). Subitizing—An analysis of its component processes. *Journal of Experimental Psychology-General*, *111*(1), 1–22.

Morgan, M. J., Raphael, S., Tibber, M. S., & Dakin, S. C. (2014). A texture-processing model of the "visual sense of number." *Proceedings of the Royal Society B-Biological Sciences*, *281*(1790), https://doi.org/10.1098/rspb.2014.1137.

Oberly, H. S. (1924). The range for visual attention, cognition and apprehension. *The American Journal of Psychology*, *35*(3), 332–352, https://doi.org/10.2307/1414013.

Park, J., DeWind, N. K., Woldorff, M. G., & Brannon, E. M. (2016). Rapid and direct encoding of numerosity in the visual stream. *Cerebral Cortex*, *26*(2), 748–763.

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, *14*(12), 542–551.

Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition*, *121*(1), 147–153.

Poulton, E. C. (1973). Unwanted range effects from using within-subject experimental designs. *Psychological Bulletin*, *80*(2), 113–121.

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, *19*(6), 607–614.

Sengupta, R., Bapiraju, S., & Melcher, D. (2017). Big and small numbers: Empirical support for a single, flexible mechanism for numerosity perception. *Attention, Perception, & Psychophysics*, *79*(1), 253–266.

Sengupta, R., Surampudi, B. R., & Melcher, D. (2014). A visual sense of number emerges from the dynamics of a recurrent on-center off-surround neural network. *Brain Research*, *1582*, 114–124.

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

Teghtsoonian, R., & Teghtsoonian, M. (1978). Range and regression effects in magnitude scaling. *Perception & Psychophysics*, *24*(4), 305–314.

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*(6):8, 1–19, https://doi.org/10.1167/12.6.8. [PubMed] [Article]

Vetter, P., Butterworth, B., & Bahrami, B. (2008). Modulating attentional load affects numerosity estimation: Evidence against a pre-attentive subitizing mechanism. *PLoS One*, *3*(9), e3269.

Zimmermann, E., & Fink, G. R. (2016). Numerosity perception after size adaptation. *Scientific Reports*, *6*, 32810.