

Enumeration of small collections violates Weber's law

H. Choo · S. L. Franconeri

Published online: 4 July 2013
© Psychonomic Society, Inc. 2013

Abstract In a phenomenon called *subitizing*, we can immediately generate exact counts of small collections (one to three objects), in contrast to larger collections, for which we must either create rough estimates or serially count. A parsimonious explanation for this advantage for small collections is that noisy representations of small collections are more tolerable, due to the larger relative differences between consecutive numbers (e.g., 2 vs. 3 is a 50 % increase, but 10 vs. 11 is only a 10 % increase). In contrast, the advantage could stem from the fact that small-collection enumeration is more precise, relying on a unique mechanism. Here, we present two experiments that conclusively showed that the enumeration of small collections is indeed “superprecise.” Participants compared numerosity within either small or large visual collections in conditions in which the relative differences were controlled (e.g., performance for 2 vs. 3 was compared with performance for 20 vs. 30). Small-number comparison was still faster and more accurate, across both “more–fewer” judgments (Exp. 1), and “same–different” judgments (Exp. 2). We then reviewed the remaining potential mechanisms that might underlie this superprecision for small collections, including the greater diagnostic value of visual features that correlate with number and a limited capacity for visually individuating objects.

Keywords Subitizing · Number Perception · Numerosity

The perceived intensity of a stimulus depends on its context. Lighting a single candle generates a salient change in a dark room, but is barely noticed in a well-lit room. Weber's law more precisely specifies this relationship—our ability to detect

differences in a signal depends on the ratio between the difference and the signal's baseline level. If a viewer is 75 % accurate at detecting a length difference between lines that are 2 and 2.2 cm in length, a proportionally larger difference would be needed to obtain the same level of performance for lines that are ten times longer: 20 and 22 cm. This ratio signature of Weber's law has been observed across the perception of many continuous dimensions, such as length, weight, and the pitch of pure tones (Henmon, 1906).

There is controversy over whether the perception of numerosity is an exception to this law. Although Weber's law states that discrimination should be noisy at all values, numerosity perception is near-perfect within its smallest baseline values. In a phenomenon called *subitizing*, previous studies have demonstrated that people can rapidly make nearly perfect counts of up to three to four objects (Jevons, 1871; Trick & Pylyshyn, 1994). In one of the first demonstrations of subitizing, Jevons threw handfuls of beans across a table, each time making an immediate count of the subset that landed in a small tray. His performance was perfect for one to four beans, with only a few errors for five beans, and then consistently high error for six to 16 beans. Since this first demonstration, subitizing has been widely observed across a variety of studies (e.g., Kaufman, Lord, Reese & Volkman, 1949; Trick & Pylyshyn, 1994).

This evidence of high precision for small collections does not necessarily mean that numerosity violates Weber's law (Dehaene, 2003). People can estimate number within collections of any size, but as predicted by Weber's Law, with a level of error proportional to the number of objects (Gallistel & Gelman, 1991; Whalen, Gallistel & Gelman, 1999). This proportional level of precision may extend down through the smallest collections, but a special property of numerosity may make this proportional noise hard to detect. Unlike the continuous dimensions of length, weight, and pitch, number is discrete, setting a minimum value of 1 on the size of between-collection differences that can be presented. This minimum difference of 1 unit may be sufficiently large to mask any uncertainty about number judgments within small (e.g., one to

H. Choo (✉)
Department of Psychology, The Ohio State University,
1872 Neil Avenue, Columbus, OH 43210, USA
e-mail: choo.38@osu.edu

S. L. Franconeri (✉)
Department of Psychology, Northwestern University,
2029 Sheridan Road, Evanston, IL 60208, USA
e-mail: franconeri@northwestern.edu

three) collections, yet also small enough to prevent confident discrimination among larger (e.g., four or more) collections.

An account that explains subitizing through the interaction of proportionally scaled noise masked by a minimum difference value between numbers (one unit) has the advantage of parsimony over those claiming special status for small collections. It also has empirical support, as some measures have suggested proportional “mental spacing” even for small collections. When discriminating between small collections, response times slowed for smaller ratios, consistent with Weber’s law (e.g., 1:2, 2:3, 3:4; Lemmon, 1927). Response times were invariant, however, for discriminations between collections with different “baselines” but a constant ratio (1:2, 2:4, 3:6; Crossman, 1955). In another study, participants arranged cards containing visual dot patterns into equally psychologically spaced categories, and their responses reflected ratio spacing even across small collections of objects (Buckley & Gillman, 1974).

Yet also, some evidence supports the alternative interpretation that the perception of numerosity in small collections is “superprecise” relative to the predictions from Weber’s law. Such superprecision could stem from special processing mechanisms that are available to the visual system for small collections. Evidence for these special mechanisms has stemmed from suspiciously consistent capacity limitations of three to four objects or locations (for a discussion, see Franconeri, Alvarez & Enns, 2007) across phenomena such as object tracking (Pylyshyn & Storm, 1988) and visual memory (Luck & Vogel, 1997). Studies on infants have provided particularly powerful evidence of the special processing of small collections. For example, infants can discriminate two objects from three, but not one object from four, suggesting that small and large collections (for infants, four appears to be a ‘large’ collection) may rely on incompatible representations (Feigenson, 2007). Therefore, determining whether small-collection enumeration is truly superprecise holds strong implications for how we understand the capacity limits on the visual system, as well as for the human ability to differentiate discrete objects from continuous substances (Hespos, Ferry & Rips, 2009) and the developmental mechanism that bootstraps our understanding of number (Carey, 2010).

How could the accounts above be dissociated? One possibility would be to model the precision of the estimation process and to evaluate fits for the competing accounts. But one such analysis failed to find satisfactory fits to performance for versions of both the Weber’s law and “special-mechanism” accounts (Balakrishnan & Ashby, 1991), suggesting that modeling may not be the most fruitful route. One possibility would be to compare identical ratios at different baselines (e.g., 2:3 vs. 20:30) at which Weber’s law predicts identical performance. The first study to use this approach revealed perfect correspondence with Weber’s law (Crossman, 1955). However, this study used a high ratio of 1:2, leaving strong

potential for performance ceiling limitations. In the present study, we will use a wider range of ratios.

Another study featured a comprehensive range from one to eight and ten to 80 objects, comparing naming times and accuracy (Revkin, Piazza, Izard, Cohen & Dehaene, 2008). Small collections of one to four showed near-perfect counts, in contrast to the difficulty with large collections of ten to 40. However, this relative advantage for small collections could have stemmed from postperceptual stages of the task. In particular, the advantage could stem from strong existing mappings from perceptual information related to number, due to verbal labels for those numbers. The participants had a lifetime of practice with the verbal labeling of small collections (i.e., linguistic frequencies dramatically decrease with numerosity intensity; Dehaene & Mehler, 1992), as compared to limited training with the naming of larger collections. Linguistic frequency relationships might underlie other effects that ostensibly reflect the structure of number representations (e.g., the SNARC effect; Hutchinson, Johnson & Louwse, 2011) and can serve as a proxy for distance relationships more generally—indeed, a map of Middle Earth can be generated solely from the co-occurrence of geographic locations in the text of *The Lord of the Rings* (Louwse & Benesh, 2012).

Here, we isolated the perceptual limits of enumeration by using a visual comparison task that did not require verbal labeling. In this way, we were able to show conclusively that the enumeration of small collections is *perceptually* superprecise relative to the predictions of Weber’s law.

Experiment 1

Observers judged the relative numerosities of collections with either small (3 vs. 1, 2, 4, or 5) or large (30 vs. 10, 20, 40, or 50) baselines. Weber’s law predicts that performance should be identical across these baselines, as long as the ratios are equal. If Weber’s law were violated, response times (RTs) should be faster for comparisons of 1:3 and 2:3, relative to 10:30 and 20:30, respectively. Such advantages may only be present for 2:3 relative to 20:30 because the smaller ratio creates a difficult comparison. This advantage would manifest as an interaction between baseline size and ratio size within the lower values in each baseline size. If such differences were absent in the higher values of each baseline (3:4, 3:5; 30:40, 30:50), it would point even more definitively to superprecision in the range from one to three (marked by a triple interaction between baseline size, ratio size, and comparison direction).

Methods

Participants A group of 25 Northwestern undergraduates (15 females, 10 males) participated in this experiment. All of the participants were naive and reported corrected-to-normal

vision. To ensure that all participants could “subitize,” they also performed a separate number-naming task¹ after the experiment, and one participant (male) was excluded from further analysis due to accuracy rates lower than 90 % for collections of one to three objects.

Stimuli and apparatus All of the stimuli were created and displayed using MATLAB with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) on an Intel Macintosh running OS X 10.6. All stimuli were displayed on a 17-in. ViewSonicE70fB CRT monitor (1,024 × 786, 75 Hz). The viewing distance was approximately 57 cm. In a fixation display, a green (91 cd/m²) bar with a size of 0.29° (width) × 21.43° (height) appeared for 200 ms on a black background (42 cd/m²). The subsequent dot collection display (see Fig. 1) added two collections (white 101-cd/m² dots, each subtending 0.30°) scattered within imaginary circles (11° diameter) with center points 8.93° to the left and right of fixation. One collection always had 3 or 30 dots, and the other collection either 1, 2, 4, or 5 or 10, 20, 40, or 50 dots. The dots were randomized without overlap by choosing 198 random locations, with an additional random jitter ranging from 0° to 0.14°.

Procedure Participants reported whether the variably sized collection contained fewer or more dots than the reference (3 or 30). The baseline size order was blocked across participants with an ABAB or BABA counterbalancing, and the location of the reference collection was blocked as AABB or BBAA (Fig. 1). Each trial began with a 100-ms beep simultaneous with a 200-ms fixation display, and then the dot collection display was presented until response. Each variably sized collection (e.g., 1, 2, 4, or 5) was equally likely within a block. Participants pressed keys labeled “more” or “fewer” in order to make relative judgments, with the label mapping being counterbalanced to the “k” and “o” keys across participants. Errors resulted in feedback (the word INCORRECT presented for 2,500 ms). The intertrial interval (ITI) was 800 ms. The total number of trials was 288: 2 baseline sizes (3 or 30) × 2 ratio sizes (larger [1:3, 3:5] or smaller [2:3, 4:5]) × 2 comparison directions (fewer or more) × 2 reference collection locations (left or right) × 18 repetitions.² The experiment lasted approximately 30 min, including 16 practice trials and the subsequent number-naming task.

¹ Participants were asked to name the number of one to seven dots. This test was identical to the main experiment, except that a single collection was presented at the center of the screen for 200 ms and then masked by a random pattern, preventing further visual processing. Participants performed 210 trials (7 collections × 30 repetitions).

² One of the participants performed 384 trials, with 24 repetitions of each condition.

Results and discussion

For each participant, RTs higher than three *SDs* over the mean were discarded, as well as an equal number from the opposite side of the distribution ($M = 4.3\%$, $SD = 1.3\%$). The average accuracy was 94.6 % ($SD = 2.6\%$). The patterns of error rates qualitatively matched the RT data.

The RTs for correct trials were entered into a repeated measures ANOVA with three factors: Baseline Size (small vs. large), Comparison Direction (fewer vs. more), and Ratio Size (smaller vs. larger) (see Fig. 2a). The RT for small-baseline-size collections ($M = 561$ ms, $SE = 17$ ms) was not significantly faster than the one for large-baseline collections ($M = 573$ ms, $SE = 20$ ms), $F(1, 23) = 2.9$, $p = .10$, $\eta^2 = .11$. The RT was faster for the “fewer” ($M = 551$ ms, $SE = 18$ ms) than for the “more” ($M = 583$ ms, $SE = 19$ ms) comparison direction, $F(1, 23) = 17.1$, $p < .001$, $\eta^2 = .43$, and the RT was also faster for larger ($M = 540$ ms, $SE = 17$ ms) than for smaller ($M = 594$ ms, $SE = 20$ ms) ratio sizes, $F(1, 23) = 134.9$, $p < .001$, $\eta^2 = .85$. These results are both consistent with the idea that smaller target-to-reference ratios can increase the difficulty of number comparisons.

Critically, baseline size interacted significantly with each of the other two factors: Baseline Size × Comparison Direction, $F(1, 23) = 5.9$, $p = .02$, $\eta^2 = .20$; Baseline Size × Ratio Size, $F(1, 23) = 15.6$, $p < .001$, $\eta^2 = .40$; and most importantly, Baseline Size × Ratio Size × Comparison Direction, $F(1, 23) = 8.7$, $p = .007$, $\eta^2 = .27$. This predicted three-way interaction shows that 20:30 comparisons were slower than 10:30 comparisons by 91 ms ($SE = 7$ ms), but that such distance effects were smaller between the 1:3 and 2:3 comparisons, at 45 ms ($SE = 6$ ms). This three-way interaction also shows that the distance effect difference was not present outside of the subitizing range: The 30:40 comparisons were slower than the 30:50 comparisons by 44 ms ($SE = 5$ ms), and the difference in the distance effect was comparable between 3:4 and 3:5, at 36 ms ($SE = 6$ ms). Figure 3 depicts this three-way interaction in a more straightforward format, showing the large increase in precision for “fewer,” but not for “more,” judgments of small collections. These results suggest that small collections of one to three are superprecise, in contrast to Weber’s law.

Experiment 2

In Experiment 2, we replicated these effects using a second type of comparative judgment: Instead of reporting “more” or “fewer,” we asked participants to report “same” or “different.” The patterns of RTs and accuracy for the *different* trials reflected the same critical three-way interaction found in Experiment 1.

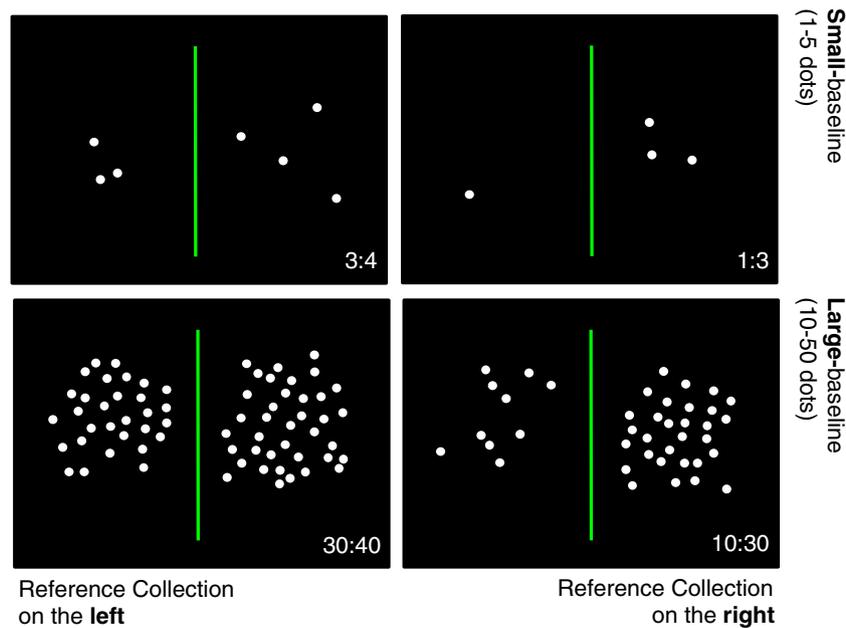


Fig. 1 Four schematic dot displays used in Experiments 1 and 2. The ratios printed at the bottom right corner of each example display were not displayed during the experiments. A fixation was presented for 200 ms, with a 100-ms beep. Then, a dot display was presented until

response. The size of the baselines and the location of the reference collection (always three or 30 dots) were blocked, and the block order was counterbalanced across participants

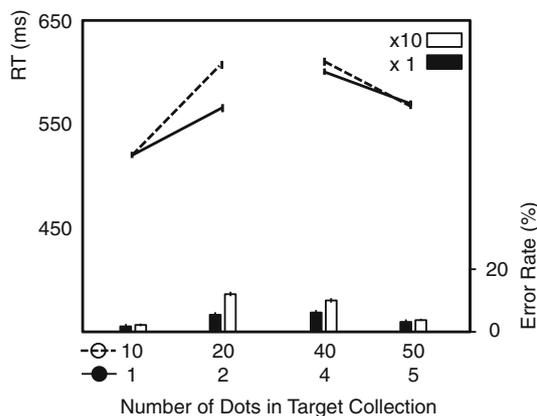
Method

Participants A group of 28 Northwestern undergraduates (14 females, 14 males) participated in Experiment 2. All participants reported normal or corrected-to-normal vision and were naive to the purpose of this experiment. The same number-naming control task was used, leading to the exclusion of four participants (two males, two females) who could not perform at 90 % at rapid number naming for small collections (one to three objects).

Stimuli and apparatus These were identical to those aspects of Experiment 1.

Procedure The procedure was identical to that of Experiment 1, except for the following changes. The task was to judge whether the two collections had either the same or different numbers of dots. In addition to the displays with *different* numbers of dots, an equal frequency of *same* trials was added in which both collections had either three or 30 dots. The experiment consisted of 384 *same* trials [2 baseline sizes (3 vs.

A) Exp 1: RTs for More/Fewer Judgments than 3 or 30



B) Exp2: RTs for Same/Different Judgments to 3 or 30

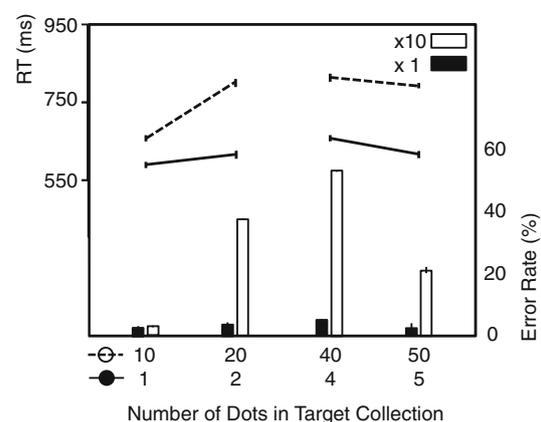


Fig. 2 Average response times (RTs) and error rates for Experiments 1 and 2 for both the small-baseline ($\times 1$) and large-baseline ($\times 10$) collections. Filled circles connected by solid lines represent the small-baseline condition (one to five), and open circles connected by dashed lines

represent the large-baseline condition (10–50). Average error rates are also shown with the RTs; black bars represent the small-baseline condition, and white bars represents the large-baseline condition. Error bars depict standard errors

30) \times 192 repetitions] and 384 *different* trials [2 baseline sizes (3 vs. 30) \times 2 comparison directions (more vs. fewer) \times 2 ratio sizes (larger vs. smaller) \times 2 reference collection locations (left vs. right) \times 24 repetitions]. Experiment 2 lasted about 50 min, including 16 practice trials and the subsequent number-naming task.

Results and discussion

Using the same screening procedure as in Experiment 1, we discarded 3.1 % of the trials ($SD = 1.0$ %). The average accuracy of the remaining trials was 84.0 % ($SD = 19.96$ %). The patterns of error rates qualitatively matched the RT data.

For the *different* trials, the same ANOVA as in Experiment 1 again revealed that RTs were significantly faster in the “fewer” comparison direction ($M = 696$ ms, $SE = 15$ ms) than in the “more” comparison direction ($M = 766$ ms, $SE = 18$ ms), $F(1, 23) = 17.1$, $p < .001$, $\eta^2 = .43$. RTs were also faster for larger ($M = 695$ ms, $SE = 15$ ms) than for smaller ($M = 768$ ms, $SE = 18$ ms) ratios, $F(1, 23) = 134.9$, $p < .001$, $\eta^2 = .85$ (see Fig. 2b). Unlike in Experiment 1, we found a strong effect of baseline size, $F(1, 23) = 72.2$, $p < .001$, $\eta^2 = .76$ (small, $M = 652$ ms, $SE = 15$ ms; large, $M = 811$ ms, $SE = 22$ ms). A similar effect was present in the *same* trials, in which the average RT was slower for the large baseline ($M = 823$ ms, $SE = 14$ ms) than for the small baseline ($M = 642$ ms, $SE = 11$ ms), $t(23) = 7.6$, $p < .001$. The main effect of baseline size may have occurred because the same–different discrimination task in Experiment 2 entailed more fine-grained decision thresholds around the neighboring responses: “Same–different” decisions require discrimination between smaller ratios—for example, among 2:3/20:30, 3:3/30:30, and 3:4/30:40—than do “more–fewer” decisions—between 2:3/20:30 and 3:4/30:40. Most importantly, baseline size interacted with all of the other main factors: Baseline Size \times Comparison Direction, $F(1, 23) = 19.6$, $p < .001$, $\eta^2 = .46$; and Baseline Size \times Ratio Size, $F(1, 23) = 29.8$, $p < .001$, $\eta^2 = .56$. The three-way interaction of Baseline Size \times Ratio Size \times Comparison Direction was again significant, $F(1, 23) = 24.0$, $p < .001$, $\eta^2 = .51$, showing that the size of distance effects was reduced only for relative judgments on one to three objects, but not for relative judgments on three to five objects, as compared with their matched large collections (see Fig. 3).

General discussion

Our results suggest that number discrimination in small collections of one to three objects is *visually* more precise than is predicted by Weber’s law. If the precision of each collection’s number representation were indeed proportional to its value, number discriminations between small collections (e.g., 2:3)

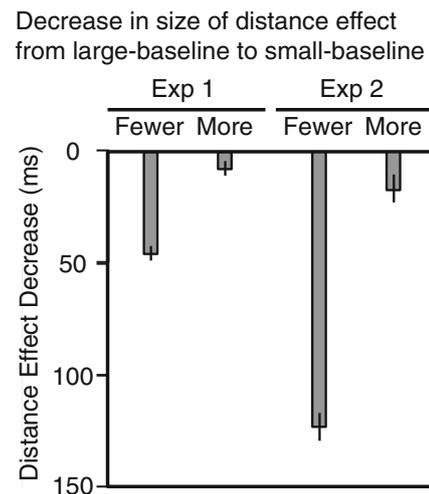


Fig. 3 Differences in the slopes between the small-baseline condition (solid lines) and the large-baseline condition (dashed lines) from Fig. 2 (a proxy for each of the two-way interactions visible in that figure), represented as bars showing single values. If enumeration is based on Weber’s law, these values should not deviate from zero, because distance effects should be equal for the large- and small-baseline judgments. In contrast, distance effects were far smaller in small-baseline judgments for “fewer” judgments, in both Experiments 1 and 2. These patterns plotted above together suggest a superprecise enumeration for collections of one to three objects, well beyond the precision predicted by Weber’s law

should have shown performance equal to that in larger discriminations with the same ratio (e.g., 20:30). Instead, RTs for decisions about the small collections were far faster than those for large collections, both for relative numerosity discrimination (Exp. 1) and for same–different numerosity discrimination (Exp. 2). To our knowledge, these results are the first demonstration that small-collection “superprecision” is indeed rooted in the visual stages of numerosity judgment, isolating this precision from later stages that associate those representations with symbolic or verbal codes.

Where does this superprecision come from? We know of two classes of explanation that remain if the Weber’s law explanation is ruled out. The first is that the *diagnostic value* of visual information across small collections is inherently higher than for large collections. Number perception may rely on correlations between number and other visual features, such as a collection’s covered area (e.g., the group’s circumference), textural density, spatial frequency profile, or other operations over primitive image segmentation (see Franconeri, Bemis & Alvarez, 2009, for a discussion). For example, a change from one dot to two dots causes powerful changes to the spatial frequency profile of a collection. Changing from two to three dots causes an equally powerful difference to the second spatial dimension of this spatial frequency profile (since the original two dots could only be arranged in a single spatial dimension).

Consistent with this idea, adding additional objects in a linear arrangement makes subitizing abilities vanish (Allen &

McGeorge, 2008). Critically, for large collections, changes of the same ratio sizes, such as 10:20 or 20:30, create nowhere near the same magnitude of signal difference among these dimensions. A collection of 20 essentially fills an entire display or container, and a collection of 30 adds little more surface area or difference in the spatial frequency profile. Such cues may be the reason why participants rate individual numerosities as being similar to each other (e.g., three-object patterns all look similar), but different from flanking values (e.g., three-object patterns look dissimilar from four-object patterns). These patterns reverse for larger collections—the similarity among seven-object patterns disappears, so that seven starts to look similar to eight (Logan & Zbrodoff, 2003).

A second class of explanation is that something special about the architecture of the visual system outputs discrete value for each small collection of one to three objects. For example, the perception of small collections may tap a shape recognition system that treats dots or objects as the vertices of a polygon, yielding long-term memory associations between those shapes and the number or vertices (or sides) that they contain (Mandler & Shebo, 1982). A dot entails one object, a line two, a triangle three, and a square or diamond four. Larger numbers do not signal prototypical shapes, and therefore should not produce efficient performance. The disruption of subitizing by linear arrangement has been used to support this account, because shape is also disrupted (Allen & McGeorge, 2008). And subitizing limits can be raised when patterns signal familiar shapes (Peterson & Simon, 2000), such as dice patterns (Mandler & Shebo, 1982).

Another possibility is that the visual system may contain an object individuation mechanism that is limited to three to four objects (Trick & Pylyshyn, 1994), which could rise from limitations in the cortical spacing of objects (Franconeri, Alvarez & Cavanagh, 2013; Franconeri, Jonathan & Scimeca, 2010). The most compelling evidence for this possibility comes from a surprisingly clear trade-off between performance in a multiple-object tracking task and a concurrent subitizing task: While attentively tracking multiple moving objects, subitizing limits decreased by one for each additional object tracked (Chesney & Haladjian, 2011). A similar trade-off was also found between a visual memory task and a concurrent subitizing task, but not between the visual memory task and concurrent numerosity judgments on large collections (Piazza, Fumarolar, Chinello & Melcher, 2011). Similarly, subitizing was significantly compromised under attention-demanding concurrent tasks (Olivers & Watson, 2008; Railo, Koivisto, Revonsuo & Hannulae, 2008), in which some common processing resource may have been recruited for individuation (Intriligator & Cavanagh, 2001). Subitizing limits might also decrease if collections are cramped in a rather smaller cortical space—for example, within a visual quadrant—thus hindering efficient individuation (Delvenne, Castronovo, Demeyere & Humphreys, 2011).

Author Note We thank Audrey Lustig and Priti Shah for helpful feedback. This work was supported by NSF CAREER Grant No. BCS-1056730 (to S.L.F.).

References

- Allen, R., & McGeorge, P. (2008). Enumeration: Shape information and expertise. *Acta Psychologica*, *129*, 26–31. doi:10.1016/j.actpsy.2008.04.003
- Balakrishnan, J. D., & Ashby, F. G. (1991). Is subitizing a unique numerical ability? *Perception & Psychophysics*, *50*, 555–564.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436. doi:10.1163/156856897X00357
- Buckley, P. B., & Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology*, *103*, 1131–1136.
- Carey, S. (2010). *The origin of concepts*. New York: Oxford University Press.
- Chesney, D. L., & Haladjian, H. H. (2011). Evidence for a shared mechanism used in multiple-object tracking and subitizing. *Attention, Perception, & Psychophysics*, *73*, 2457–2480. doi:10.3758/s13414-011-0204-9
- Crossman, E. R. F. W. (1955). The measurement of discriminability. *Quarterly Journal of Experimental Psychology*, *7*, 176–195. doi:10.1080/17470215508416692
- Dehaene, S. (2003). The neural basis of the Weber–Fechner law: A logarithmic mental number line. *Trends in Cognitive Sciences*, *7*, 145–147. doi:10.1016/S1364-6613(03)00055-X
- Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*, 1–29.
- Delvenne, J.-F., Castronovo, J., Demeyere, N., & Humphreys, G. W. (2011). Bilateral field advantage in visual enumeration. *PLoS ONE*, *6*, e17743. doi:10.1371/journal.pone.0017743
- Feigenson, L. (2007). The representations underlying infants' choice of more: Object file versus analog magnitude. *Psychological Science*, *13*, 150–156.
- Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. C. (2013). Flexible cognitive resources: Competitive content maps for attention and memory. *Trends in Cognitive Sciences*, *17*, 134–141. doi:10.1016/j.tics.2013.01.010
- Franconeri, S. L., Alvarez, G. A., & Enns, J. T. (2007). How many locations can be selected at once? *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1003–1012. doi:10.1037/0096-1523.33.5.1003
- Franconeri, S. L., Bemis, D. K., & Alvarez, G. A. (2009). Number estimation relies on a set of segmented objects. *Cognition*, *113*, 1–13. doi:10.1016/j.cognition.2009.07.002
- Franconeri, S. L., Jonathan, S., & Scimeca, J. M. (2010). Tracking multiple objects is limited only by spatial interference, not speed, time, or capacity. *Psychological Science*, *21*, 920–925.
- Gallistel, C. R., & Gelman, R. (1991). Subitizing: The preverbal counting process. In F. I. M. Craik, W. Kessen, & A. Ortony (Eds.), *Essays in honor of George Mandler* (pp. 65–81). Hillsdale: Erlbaum.
- Henmon, V. A. C. (1906). The time of perception as a measure of differences in sensation. *Archives of Philosophy, Psychology, and Scientific Methods*, *8*
- Hespos, S. J., Ferry, A. L., & Rips, L. J. (2009). Five-month-old infants have different expectations for solids and substances. *Psychological Science*, *20*, 603–611. doi:10.1111/j.1467-9280.2009.02331.x
- Hutchinson, S., Johnson, S., & Louwerse, M. M. (2011). A linguistic remark on SNARC: Language and perceptual processes in spatial-numerical association. In L. Carlson, C. Hölscher, & T. F. Shipley (Eds.), *Expanding the space of cognitive science: Proceedings of*

- the 33rd annual conference of the cognitive science society* (pp. 1313–1318). Austin: Cognitive Science Society.
- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, *43*, 171–216. doi:10.1006/cogp.2001.0755
- Jevons, W. S. (1871). The power of numerical discrimination. *Nature*, *3*, 281–282.
- Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkman, J. (1949). The discrimination of visual number. *American Journal of Psychology*, *62*, 498–525.
- Lemmon, V. W. (1927). The relation of reaction time to measures of intelligence, memory, and learning. *Archives of Psychology*, *15*.
- Logan, G. D., & Zbrodoff, N. J. (2003). Subitizing and similarity: Toward a pattern-matching theory of enumeration. *Psychonomic Bulletin & Review*, *10*, 676–682. doi:10.3758/BF03196531
- Louwerse, M. M., & Benesh, N. (2012). Representing spatial structure through maps and language: *Lord of the rings* encodes the spatial structure of middle earth. *Cognitive Science*, *36*, 1–14. doi:10.1111/cogs.12000
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281. doi:10.1038/36846
- Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*, 1–22. doi:10.1037/0096-3445.111.1.1
- Olivers, C. N. L., & Watson, D. G. (2008). Subitizing requires attention. *Visual Cognition*, *16*, 439–462.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442. doi:10.1163/156856897X00366
- Peterson, S. A., & Simon, T. J. (2000). Computational evidence for the subitizing phenomenon as an emergent property of the human cognitive architecture. *Cognitive Science*, *24*, 93–122.
- Piazza, M., Fumarolar, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition*, *121*, 147–153.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*, 179–197. doi:10.1163/156856888X00122
- Railo, H., Koivisto, M., Revonsuo, A., & Hannulae, M. (2008). The role of attention in subitizing. *Cognition*, *107*, 82–104.
- Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, *19*, 607–614. doi:10.1111/j.1467-9280.2008.02130.x
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, *101*, 80–102. doi:10.1037/0033-295X.101.1.80
- Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science*, *10*, 130–137. doi:10.1111/1467-9280.00120