Cognition 113 (2009) 1–13

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Number estimation relies on a set of segmented objects

S.L. Franconeri [a,*], D.K. Bemis [b], G.A. Alvarez [c]

[a] Department of Psychology, Northwestern University, United States
[b] Department of Psychology, New York University, United States
[c] Department of Psychology, Harvard University, United States

ARTICLE INFO

Article history:
Received 9 August 2007
Revised 23 June 2009
Accepted 6 July 2009

Keywords:
Number
Attention
Segmentation
Numerosity
Estimation

ABSTRACT

How do we estimate the number of objects in a set? Two types of visual representations might underlie this ability – an unsegmented visual image or a segmented collection of discrete objects. We manipulated whether individual objects were isolated from each other or grouped into pairs by irrelevant lines. If number estimation operates over an unsegmented image, then this manipulation should not affect estimates. But if number estimation relies on a segmented image, then grouping pairs of objects into single units should lead to lower estimates. In Experiment 1 participants underestimated the number of grouped objects, relative to disconnected objects in which the connecting lines were 'broken'. Experiment 2 presents evidence that this segmentation process occurred broadly across the entire set of objects. In Experiment 3, a staircase procedure provides a quantitative measure of the underestimation effect. Experiment 4 shows that the strength of the grouping effect was equally strong for a single thin line, and the effect can be eliminated by a small break in the line. These results provide direct evidence that number estimation relies on a segmented input.

The visual system is able to quickly generate summary information about a scene or collection of objects, such as the gist of a scene (Oliva, 2005), the average size of its objects (Ariely, 2001; Chong & Treisman, 2003), the average location (Alvarez & Oliva, 2008), or spatial ensemble statistics (Alvarez & Oliva, 2009). Another salient example is our ability to immediately perceive the approximate number of objects in a collection. Approximate number is quickly perceived by adults (e.g., Miller & Baker, 1968), and even infants (Xu & Spelke, 2000), with an error distribution that increases proportionally to the number being judged (Gallistel & Gelman, 1992).

How do we generate these number estimates? Number is an abstract property of the visual world, and it is difficult to imagine a system that measures number as directly as luminance or orientation. Instead, number estimation is likely to rely on a combination of 'surrogate' features that approximate number.[1] For example, when estimating the number of coins on a desk, one approximation of the answer might be the total area of the desk covered by a metallic substance, divided by the size of a typical coin. This typical coin size could be derived from a quick sample of a single coin's size from the table, or a calculation of the average size of the coins on the table (Ariely, 2001; Chong & Treisman, 2003), or from long-term memory. Another approximation of number might arise from the spatial frequency profile of the collection of coins. Assuming the coins were scattered randomly,

* Corresponding author. Address: Department of Psychology, Northwestern University, 2029 Sheridan Rd., Evanston, IL 60208, United States. Tel.: +1 847 467 1259; fax: +1 847 491 7859.
E-mail address: franconeri@northwestern.edu (S.L. Franconeri).

[1] The present experiments address only the estimation system for large numbers of objects. Other work suggests a potentially different mechanism allowing fast accurate counts of small collections of up to 3–5 objects. The potential mechanism for this qualitatively different process includes visual indexes (Trick & Pylyshyn, 1993) or pattern recognition processes (Logan & Zbrodoff, 2003), see Dehaene and Cohen (1994) for discussion. Other work also suggests as estimates of large collections are increasingly supplemented by other cues such as texture density (Durgin, 1995).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.07.002

higher spatial frequency power should increase with greater numbers of coins, while lower spatial frequency power could be associated with the total area taken up by the collection (these relationships rely on some assumptions, including how the coins are distributed over space). Yet another possible correlate of number might be the texture density of a collection, a property that is dissociable from spatial frequency (Durgin, 2008; Durgin & Huk, 1997; Durgin & Proffit, 1996).

Such approximations are intriguing in that they violate the spirit of number estimation. Number is inherently discrete, a measure of unitized *things*. But cues such as covered area, spatial frequency, or texture density do not take discrete units as input, and instead operate over a continuous visual image. Note that these surrogate features must eventually lead to abstracted representations of number *per se*, to explain findings of accurate number estimation across modalities (Barth, Kanwisher, & Spelke, 2003) and visual adaptation effects for numerosity (Burr & Ross, 2008a; but see Burr & Ross, 2008b; Durgin, 2008 for debate on whether these effects reflect adaptation of number *per se*, or correlates of number such as local texture density).

In contrast, many models of number estimation make an assumption that estimation operates over a representation of objects segmented into discrete units. One model of number estimation specifies that a visual image first undergoes an 'object normalization' process (Dehaene & Changeux, 1993). Starting with a simplified input of differently sized 'blobs' of activation on a simulated retina, the model uses a primitive object segmentation algorithm to isolate each object and produce a single locus of activation for each, independent of object size.

Another account, the 'occupancy' model, treats 'objects' in a display as somewhat distinct, but it is not clear whether these objects are broadly segmented. This model outlines each object with a circle of a set radius[2] and sums the area of those circles to produce a correlate of the number of objects (Allik & Tuulmets, 1991). When objects are close, overlap in their occupancy reduces their contribution to the total estimate, explaining illusions where clustered displays tend to be perceived as containing fewer objects (e.g., Ginsberg & Goldstein, 1987). One way to implement this occupancy model would be to segment a display into discrete objects and then apply an 'occupancy radius' to each object. However, this model could also describe processes operating over an unsegmented image, where the smaller contributions of densely spaced objects could result from inter-object masking or crowding (Intriligator & Cavanagh, 2001; Pelli, Palomares, & Majaj, 2004) or changes in spatial frequency profile or texture density. The occupancy radius could also be a product of preferred object sizes stemming from the size of critical spatial filters in early visual analysis (Allik & Tuulmets, 1991; Watt & Morgan, 1985).

Either of these potential mechanisms would need to account for the influence of several display variables on number estimates. For example, number estimates are affected by the amount of area in a display covered by objects (Vos, van Oeffelen, Tibosch, & Allik, 1988). Number estimates are higher when objects are spread across a larger area (Allik et al., 1991; Bevan, Maier, & Helson, 1963; Hollingsworth, Simmons, Coates, & Cross, 1991; Krueger, 1972; Sophian & Chu, 2008; Vos et al., 1988), or when the size of a frame defining the display is increased (Bevan & Turner, 1964). Estimates are also higher for smaller objects (Ginsberg & Nichols, 1988; Miller & Baker, 1968; Sophian, 2007), which might create the illusion of a larger area relative to the size of the objects. The object size manipulation alters number estimates because it changes display context as a whole – when objects are presented sequentially in isolation, manipulations of object size do not affect estimates (Beran, Taglialtela, Flemming, James, & Washburn, 2006). It is important to note that manipulations of spacing and object size are often difficult to interpret because they are confounded with changes in density, object contour length, or spatial frequency profile. Such confounding factors may be the reason for different effects of object size and spacing in other work (e.g., Birnbaum, Kobernick, & Veit, 1974; see Sophian & Chu, 2008 for discussion).

Both types of number estimation mechanism might be consistent with these effects. For estimation by processes that rely on non-segmented representations, the spacing manipulation would directly alter the surrogate features used for estimation, such as covered area, spatial frequency profile, or texture density. As a specific example, under the occupancy model (Allik & Tuulmets, 1991), the occupancy radius of each object could be defined relative to the overall size of the collection (Allik et al., 1991), and critical spatial filters for the image could be defined relative to the size of the relevant scene (Nakayama, 1990). Mechanisms that rely on a segmented representation would have more difficulty explaining these spacing and object size effects because the manipulations should not alter the number of units present in a scene. However, it is possible that spacing and size manipulations affect estimates at a processing stage after the recovery of the number estimate, such as in the comparison or response. Image properties could be computed in parallel with number and could substitute for (e.g., Stroop, 1935) or compete with the number response (Allik et al., 1991). For example, when judging which of two collections had the larger number of objects, observer judgments were faster and more accurate when the relative sizes of those two collections were congruent with their relative numerosity differences, relative to when they were incongruent, even when object size was irrelevant to the task (Hurewitz, Gelman, & Schnitzer, 2006; but see also Barth, 2008). Responses might also be influenced by other associations between image properties and number estimates, such as an association between smaller objects (e.g., berries or ants) and greater numerosity, as opposed to larger objects (e.g., houses or chairs) and relatively lower numerosity (Sophian, 2007).

As a second example, the amount of clustering in a collection's arrangement has a strong influence on number estimates. Relative to randomly scattered objects, estimates are lower when objects are locally clustered (Frith & Frith, 1972; Ginsberg & Goldstein, 1987) or spatially or-

[2] This radius had been suggested as a set retinal angle (Allik & Tuulmets, 1991), a distance relative to the total size of the collection (Allik, Tuulmets, & Vos, 1991), as well as relative to the textural density of the object collection (Durgin, 1995).

ganized into global shapes (Taves, 1941). Even randomly spaced collections contain clumps of objects, and they produce lower estimates than displays where objects are equally spaced in a regular pattern (Ginsberg, 1978). Estimation processes that rely on non-segmented representations could explain such effects through mechanisms like an occupancy radius (Allik & Tuulmets, 1991) or disruptions in the spatial frequency profile of the collection.

Segmented representations might also be affected if the segmentation process produced a more 'global' level for each unit, forcing clusters to be used as the counting unit instead of individual objects. Observers do have some control over the collection that they will count. Even within a single view, observers can isolate a collection of objects if they are featurally distinct from others (by color, orientation, or shape), and then quickly complete exact counts (Trick & Enns, 1997; Trick & Pylyshyn, 1993) or number estimates (Halberda, Sires, & Feigenson, 2006) almost as efficiently as when no other objects are present. Selecting a collection may be possible by increasing the activation of networks that respond to that collection's distinguishing feature (Cohen, Aston-Jones, & Gilzenrat, 2004; Desimone & Duncan, 1995).

However, if a segmented representation included both global clusters and lower-level units in a hierarchical arrangement, observers might not have the ability to isolate a preferred level of this hierarchy during number estimation. Instead, the requirement to broadly select an entire collection would likely produce a bias toward the global level of organization. Global biases are frequently observed in hierarchically organized objects, such as letters made of smaller letters (Navon, 1977), and the existence of this bias is dependent on various observer and stimulus factors (Lamb & Robertson, 1989, 1990). Global biases are also found in auditory stimuli, where changes to a more global pattern of notes can be easier to detect than changes to a more local pattern (Schiavetto, Cortese, & Alain, 1999). The global bias does not exist in all domains. For example, in the representation of complex events (Kurby & Zacks, 2008), there is evidence for hierarchical organization of event sequences, but observers seem to be able to represent multiple levels of the hierarchy (Zacks, Tversky, & Iyer, 2001).

In summary, our ability to quickly estimate visual number might rely on either a set of image properties that serve as a useful set of surrogates for number, or it might rely on a segmented array of discrete objects. Both types of representations might also play a role, either at processing stages where number is computed or at later stages that involve comparison and response. The purpose of the present studies is to show that, at minimum, a segmented representation contributes to number estimation.

To directly show that number estimation relies at least in part on a set of discretely segmented objects, we grouped objects in ways that created minimal change at the image level, with little effect on cues like covered area, spatial frequency, or texture density, yet had a strong effect on image segmentation. We found that grouping objects into pairs with thin lines strongly reduced the perceived number of objects in rapidly presented displays, confirming that the number estimation process is at least partially based on a representation of the display that is broadly segmented into groups.

In Experiment 1, participants were asked which of two consecutive displays contained more squares. When pairs of squares in one of the displays were connected by irrelevant lines, observers underestimated the number of squares in that display relative to displays where the pairs were not connected (the lines were present but 'broken' at the midpoint between the objects). Observers appear to be influenced by the number of connected groups instead of solely the number of squares. In Experiment 2, we present additional evidence suggesting that this segmentation occurs broadly across the display. Experiment 3 uses a staircase method to provide a more direct estimate of the number of connected objects that is perceptually equivalent to a set of unconnected objects. Finally, in Experiment 4, we test a smaller connectivity manipulation and find that breaks of only a few pixels are sufficient to eliminate the grouping effect.

1. Experiment 1

Experiment 1 tested whether connecting pairs of squares with a set of lines reduces estimates of the number of individual squares in a display. Grouping multiple regions of a display so that they become physically continuous is a particularly strong grouping manipulation that may even precede other types of grouping, such as color or shape similarity (Palmer & Rock, 1994). This connectivity grouping also appears to be mandatory for small groups of objects that are in the current focus of attention. For example, when asked to keep track of a small number of moving target squares among moving distractor squares, joining target and distractor squares within lines into connected 'pairs' impaired tracking accuracy. It appeared that participants were forced to treat each target–distractor pair as a single object, causing them to lose track of the target 'end' of the pair (Scholl, Pylyshyn, & Feldman, 2001). However, because only a small number of objects were processed at once, it is unclear whether this segmentation process operates broadly across the entire display, and whether enumeration operates over this segmented input. In the present experiment, we test whether these grouping cues are indeed processed broadly across the display, by testing whether they affect number estimation processes.

Participants viewed two brief displays of squares separated by a short interval. This task is often preferred over asking the observer to simply name the number of objects in a single display because the additional requirement of translating the number value into verbal form could add error to the number estimate. One of the displays always consisted of unconnected squares and the other display had 0%, 25%, 50%, 75%, or 100% of its squares connected. The task was to determine whether the first or second display had more squares (see Fig. 1a and b). If connecting pairs of squares forces the estimation process to operate over paired groups instead of individual squares, then participants should underestimate the number of squares in connected displays. This underestimation should *improve* performance on trials where the connected display con-

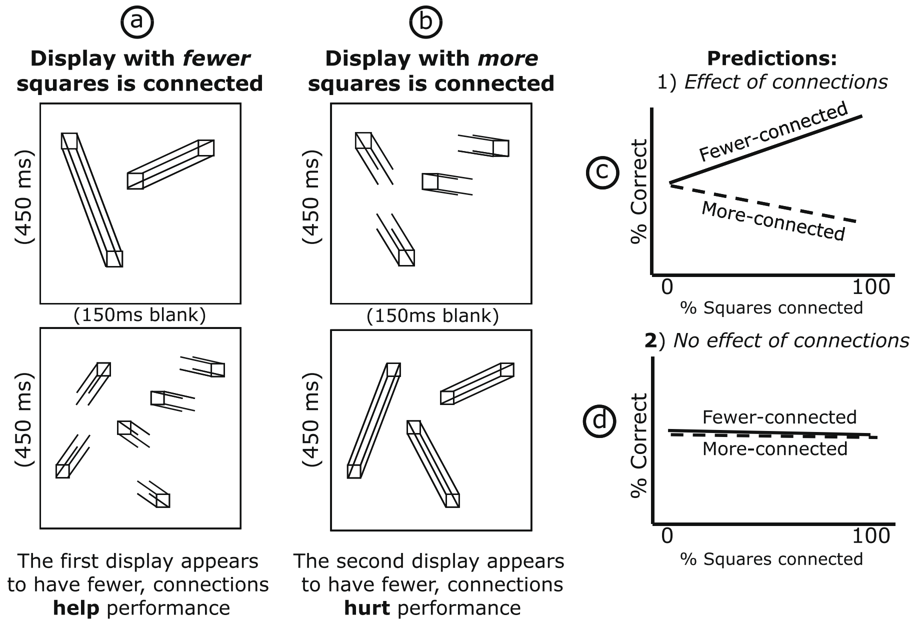

Fig. 1. In Experiment 1, observers saw two consecutive displays of squares separated by a brief blank period, and judged which display contained more squares. (a) An example of how connecting the squares in one display can help performance. (b) An example of how connecting the squares in one display can hurt performance. (c) If connecting the squares helps and hurts performance, then as a higher percentage of squares in the connected display are connected to each other, the effect should increase. (d) If connecting the squares has no effect on number estimates, then connecting the display with fewer or more squares will have no effect on performance. Displays are not drawn to scale.

tains *fewer* squares than the unconnected display (Fig. 1a) and *impair* performance on trials where the connected display contains *more* squares (Fig. 1b). Note also that the four squares in each of the first displays (and six squares in each of the second displays) of Fig. 1a and b are identical, yet without serially counting the squares, the number of squares in the displays appears different.

Fig. 1c illustrates a sample of how accuracy should change as more squares are connected to each other in a display if participants obligatorily count connected groups instead of individual squares. For trials where the connected display has *fewer* squares, underestimation of that display should cause the two displays to appear more dissimilar in number, leading to higher accuracy. For trials where the connected display actually contains *more* squares, the higher the percentage of squares connected to each other, the more a participant should underestimate the number of squares, causing the two displays to appear more similar in number, and leading to lower accuracy. In contrast, Fig. 1d shows that if the connecting lines have no effect on the estimation processes, then connecting squares to one another should have no effect on estimation accuracy.

1.1. Participants

Eighteen undergraduates participated in exchange for either course credit or $6.

1.2. Stimuli

The experiment was controlled by an iMac computer using custom software made with the VisionShell library (http://www.visionshell.com/). Although head position was not restrained, the display subtended $32.6^\circ \times 24.4^\circ$ at an approximate viewing distance of 50 cm, with a 1024×768 pixel resolution. All displays contained pairs of squares (0.7° on a side, with lines 2 pixels or 3.8 arcmin thick) that were connected by their vertices to each other by a set of four parallel lines (1 pixel or 1.9 arcmin thick). The first display contained 6, 10, 20, or 40 squares, and the second contained plus or minus 33%, (to the nearest even number, leading to actual differences of 33%, 40%, 30%, 30% for each set size, respectively).

Of these two displays, one was an unconnected 'baseline' display and one was a 'connected' display, in which 0%, 25%, 50%, 75% or 100% of square pairs were connected (the order of these displays was counterbalanced). To ensure that objects and lines would never occlude each other, each square was initially placed into one cell of an invisible grid containing 14 columns and between 1 and 4 rows, depending on the total number of objects. Squares adjacent in rows were then connected together and randomly moved vertically within the cell. A horizontal jitter of at most 0.63° was then added to these placements in order to avoid any collinear alignment between separate pairs. Unconnected pairs were then created by bisecting the connecting lines and pushing the cut ends apart vertically, such that they were separated by 2.2°.

1.3. Procedure

At the beginning of each trial, the first display was presented for 450 ms, followed by a blank screen (150 ms), the second display (450 ms), and a mask, consisting of 200 randomly placed squares and 100 randomly generated lines, which remained until the response. Each participant was instructed to determine whether the second screen contained more or fewer squares than the first screen. Critically, participants understood that they should count only the squares; they were told to ignore the lines entirely. Participants then pressed the 'M' key if they believed that there were more squares on the second display, or pressed 'L' ("Less") for fewer squares in the second display. To reduce any temptation to compensate for grouping effects with a high-level strategy (e.g., if a display is highly connected, divide estimate by some adjustment factor), responses were required within 1 s. Response times longer than 1 s resulted in an error message. Any key press would then begin the next trial.

Each session consisted of a practice block (used to make sure participants could perform the task above chance) and four test blocks. Each block consisted of 80 randomized trials, with four fully crossed factors: the number of items in the first display, whether the second display contained more or fewer squares than the first, the relative order of the 0% connected 'baseline' display and 'connected' display, and the percentage of squares connected in the 'connected' display (this factor is not relevant for the 0% connected trials, where both displays were completely unconnected). The experiment lasted about 25 min.

1.4. Results and discussion

Three participants did not perform the task (performance in practice blocks, or the 0% connected vs. 0% connected conditions, was not above chance) and were removed from the analysis. Participants made their relative number comparison quickly, with an average response time of 520 ms. Response times decreased in a linear fashion as set size increased. Trials where participants violated the 1 s response time limit (2.3% of trials) or responded in less than 100 ms (0.8% of trials) were omitted from the analysis.

Accuracy data (see Fig. 2) were submitted to a $4 \times 2 \times 4$ repeated-measures ANOVA with factors of set size (6, 10, 20, 40), whether the connected display had more or fewer squares, and percent of squares in the 'connected' display

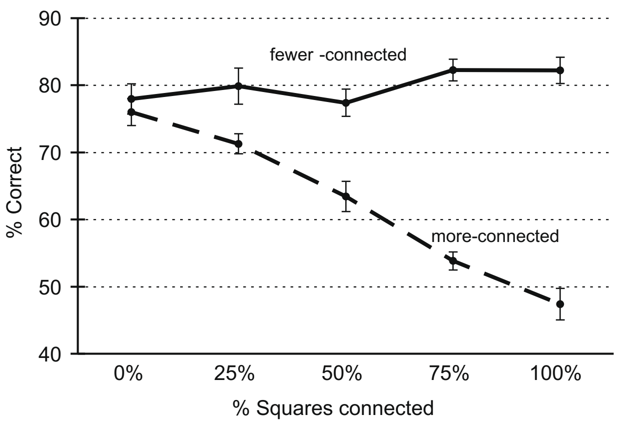

Fig. 2. In Experiment 1, connecting the squares in the display with fewer squares helped performance, while connecting squares in the display with more squares hurt performance. This effect increased dramatically in size as the percentage of connected squares in the connected display increased.

that were connected to their partner squares (25%, 50%, 75%, 100%). Trials where the 'connected' display had 0% connected pairs showed no significant difference in average accuracy ($p > .67$), as expected, and were omitted from this analysis.

There was a significant decrease in accuracy with greater set sizes ($F(3, 42) = 5.5$, $p = .003$, $h^2 = 0.2$), due to higher accuracy on set size 10 trials (75%) relative to other trials (70%), possibly caused by a greater number difference between displays at this set size (40%) relative to the difference between displays at other set sizes (30% or 33%). This difference resulted from the requirement that an even number of items be placed in each display (in order to construct connected pairs) and likely made the relative number discrimination easier for this set size.

The remaining results suggest that number estimates in the connected displays were influenced by the number of connected groups. Performance was better on fewer-connected trials ($M = 80\%$) than on more-connected trials ($M = 59\%$), $F(1, 14) = 176$, $p < 0.001$, $h^2 = 0.93$. This effect was strongest at the largest set size, as evidenced by a significant interaction between set size and whether the trial was fewer-connected or more-connected, $F(3, 42) = 5.1$, $p < 0.01$, $h^2 = 0.27$, (the differences were 19%, 13%, 14%, and 23% for the four set sizes). As the percentage of squares that were connected increased, so did the accuracy difference between fewer-connected and more-connected trials ($F(3, 42) = 22.5$, $p < 0.001$, $h^2 = 0.62$; subtracting more-connected accuracy from fewer-connected accuracy yields differences of 9%, 14%, 28%, and 35% for the 25%, 50%, 75%, and 100% connected conditions, resulting in a strong linear correlation, $r^2 = .97$. Accuracy also generally dropped with connectivity ($F(3, 42) = 13.9$, $p < .001$, $h^2 = 0.5$), due to a larger decrease in accuracy for the more-connected displays than an increase for fewer-connected displays. This effect is predicted by the changes in error expected when shifting two noisy number estimates closer or farther from each other. Moving two distributions closer to each other, as when the display with more squares is connected, will cause a larger increase in overlap between the distributions, (i.e. more error), than moving them apart will cause a decrease in overlap (i.e. less error), predicting that accuracy should drop overall as the percentage of connected squares increases.

Connecting pairs of squares caused participants to underestimate the number of squares in a display, relative to displays of unconnected squares. As predicted, accuracy increased when the connected display had fewer squares, but decreased when the connected display had more squares.

2. Experiment 2

In Experiment 1, participants estimated the number of squares in each display within 450 ms, suggesting that they did not count them serially (Gallistel & Gelman, 1992; Trick & Pylyshyn, 1993). While these results suggest that participants generated their estimates from a broad snapshot of each display, it is also possible that participants relied on a sampling process in which only a few

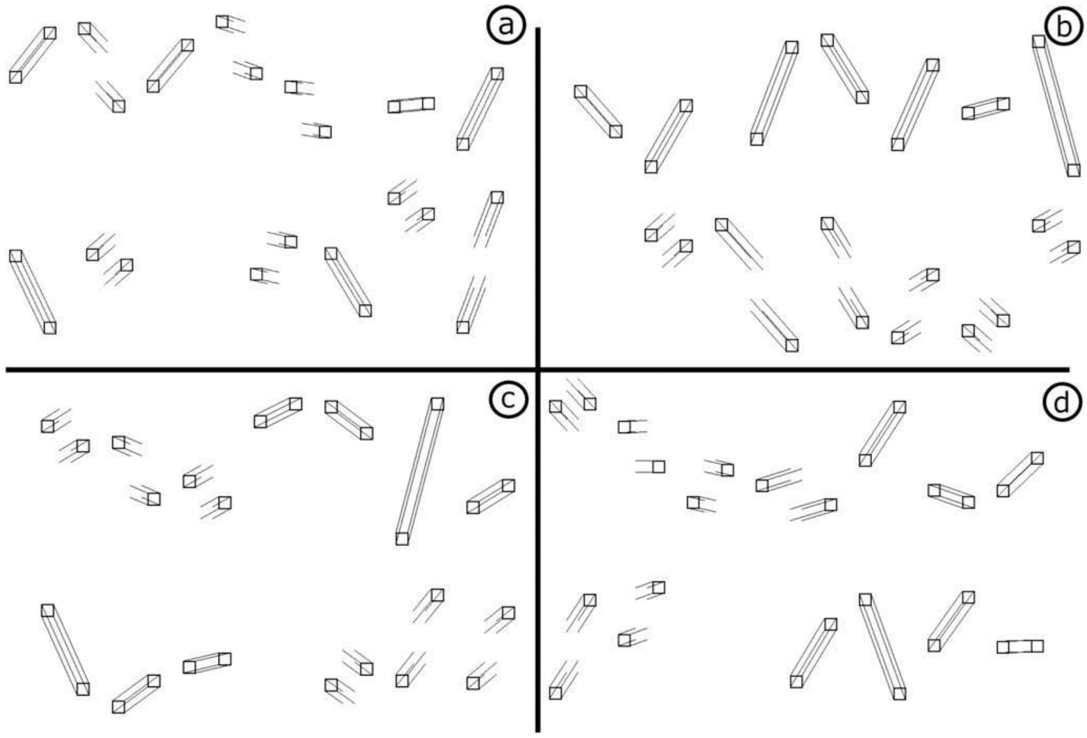

Fig. 3. In Experiment 2, all connected displays had 50% of square pairs connected. (a) An example of an unconstrained display, identical to a 50% connected display from Experiment 1. The remaining panels show restricted displays, with connected items restricted to (b) the two top quadrants, (c) the upper right and bottom left quadrants, and (d) the two right quadrants.

items from a small portion of each display were counted and compared for each trial. A similar ambiguity exists in the literature demonstrating our ability to generate statistical summaries of the properties of multiple objects, such as average size judgments. Initial reports of this ability suggested that the average was created by broadly selecting an entire collection of objects (Ariely, 2001; Chong & Treisman, 2003). However, other recent studies suggest that sampling a subset of objects could produce estimates of average size without taking input from the collection broadly (Myczek & Simons, 2008).

Experiment 2 rules out a simple version of a spatial sampling explanation by always connecting 50% of the squares in the connected display but varying whether the connected squares were clustered together or appeared in random locations (see Fig. 3). If participants estimate over a spatially defined sample of squares, then the variance in their estimates should be greater when the connected items are clustered close to one another, relative to when they are randomly distributed. That is, if the sample happens to come from a 100% connected region, number estimates for that display will be low, and if the sample happens to come from a 0% connected region, number estimates will be high. In contrast, when connected items are randomly distributed, number estimates should be more uniform (see Fig. 4 for examples of these predictions).[3]

In Experiment 1, participants indicated whether the second display had 'more' or 'less' squares than the first display. However, because variance does not exist for binomial distributions, we expanded the response options such that participants could respond "much less", "less", "more", or "much more". If participants generate their estimates from samples of small areas of each display, then there should be higher variance among the estimates (leading to a higher proportion of "much less" and "much more" responses) in the condition where the 0% and 100% connected squares are grouped into quadrants, relative to when they are randomly arranged.

2.1. Participants

Fifteen undergraduates participated in exchange for either course credit or $2.

2.2. Stimuli

Except where noted, the stimuli were identical to those in Experiment 1, with the modification that only one level of connectedness was used – in each trial an unconnected display was compared with a 50% connected display. There were two trial types (see Fig. 3). The unrestricted trials were identical to the original 50% connected conditions. On restricted trials, all of the connected objects were constrained to two randomly chosen quadrants of the screen.

2.3. Procedure

The procedure was the same as for Experiment 1 with two exceptions. First, participants had four response options: "much less", "less", "more", and "much more". Second, because only one level of connectedness (50%) was

[3] This design cannot rule out all sampling strategies. However, as sampling strategies become more sophisticated, by using samples of either greater variety or complexity, sampling becomes indistinguishable from an estimation strategy that operates broadly across the display.

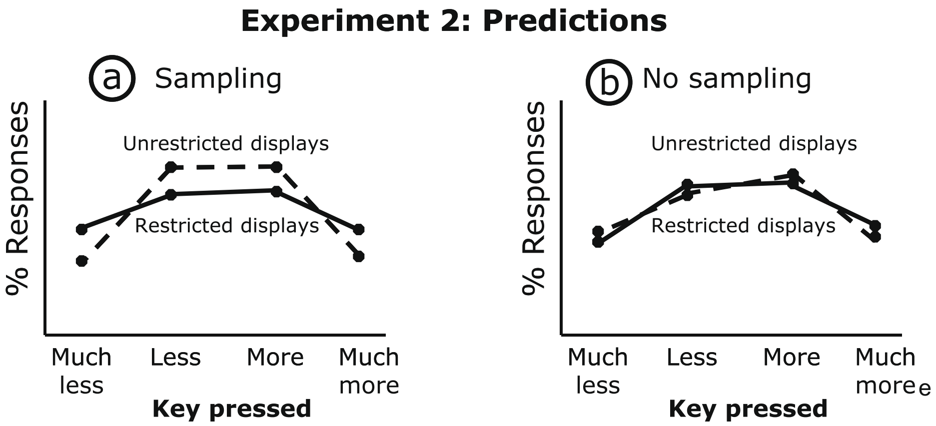

Fig. 4. Predictions for Experiment 2. If observers estimate visual number by a simple spatial sampling strategy, then restricting the connected squares to quadrants should increase the variance of that sample, (a) causing more variance in responses on the restricted displays than on the unrestricted displays. (b) If participants estimate number without sampling, by a process that operates broadly across the display, then restricting connected items to quadrants should not affect response distributions.

tested, the experiment consisted of only one test block (and one practice block which was not analyzed). The experiment lasted approximately 15 min.

2.4. Results and discussion

One participant did not perform the task above chance levels and was removed from the analysis. As shown in Fig. 5, the two distributions of responses arising from the restricted and unrestricted trials were virtually identical, suggesting that the estimation process does not rely on simple spatial sampling. There were no differences in the usage rates of each response key between the two types (all $t < 1$, $p > 0.64$, $d < 0.09$, except 'much less', $t = 1.5$, $p = 0.16$, $d = 0.24$). The lack of a difference in response variance between the restricted and unrestricted displays suggests that participants did not use a simple spatial sampling strategy to gauge which display contained more squares.

3. Experiment 3

The first two experiments show that participants rapidly segment pairs of squares that are connected together, in a spatially broad way, leading to an underestimation of the number of distinct squares appearing in a connected display. While the linear correlation between accuracy differences and percent connectedness suggests that the magnitude of the effect depends on the percentage of connected pairs, it should be possible to quantitatively determine the amount of underestimation. The present experiment used a staircase method to determine the subjective point of equality between a set of unconnected squares and a set of connected squares. That is, how many unconnected squares must be deleted for an observer to perceive the set as equal to N connected squares?

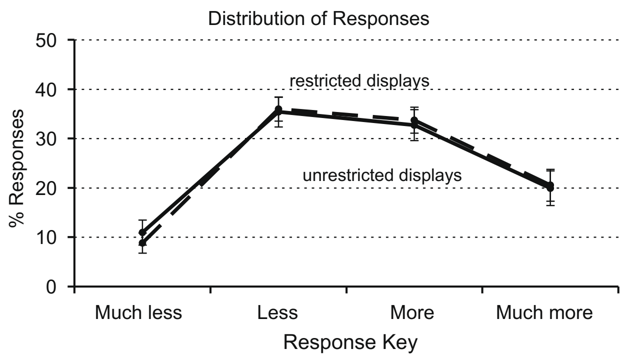

Fig. 5. In Experiment 2, restricting connected items to quadrants had no effect on response distributions.

3.1. Participants

Twenty six undergraduates participated in exchange for course credit, in addition to two authors.

3.2. Stimuli

The stimuli were identical to that of Experiment 1, except that each connected display contained either 12, 24, or 48 squares. The number of squares in the unconnected display varied according to the staircase procedure described below.

3.3. Procedure

Participants saw two displays of squares and made more/less responses, as in Experiment 1. For each of the 15 combinations of set size (12, 24, or 48 squares in the connected display) and percentage of objects connected (0%, 25%, 50%, 75%, 100%), we determined the perceptually equivalent number of unconnected squares using a staircase method. Initially, the number of objects in the unconnected display was equal to that of the connected display. Each time the participant indicated that the unconnected display had fewer squares, the number of squares in that display was increased by one for the following presentation. Otherwise, the number of squares was decreased by one. On each trial, one of the 15 trial types was randomly chosen until each trial type's staircase had reversed in direction 21 times, at which point that trial type was dropped from the remainder of the experiment. The entire experiment averaged 35 min in length.

3.4. Results and discussion

Four participants were removed from the analysis because they did not perform the task at above chance levels either in the practice blocks, or the 0% connected vs. 0% connected conditions. Participants made their estimations quickly, with an average response time of 475 ms. Response times decreased linearly as set size increased. Trials where participants violated the 1 s response time limit (2.1% of trials) or responded in less than 100 ms (1.6% of trials) were omitted from the analysis.

The perceived point of equality for 0% connected trials was, on average, 12.6, 23.8, and 47.3 for set sizes 12, 24, and 48, respectively (with standard deviations of 1.84, 2.67, and 3.07). The remaining data were submitted to a 3×4 ANOVA with set size (12, 24, 48 objects in the connected display) and percent connected (25%, 50%, 75%, 100%) as factors. There were no main effects of either set size ($F < 1$) or an interaction between the two factors ($F < 1$). However, as in Experiments 1 and 2 there was a significant main effect of percent connectedness, $F(3, 69) = 9.10$, $p < .001$, $h^2 = .283$, suggesting that as a higher percentage of objects were connected into pairs, fewer unconnected objects were needed to produce a perceptually equivalent display. Fig. 6 depicts this decrease in the perceptually equivalent number of unconnected squares for each set size. To normalize the *y*-axis for each set size, underestimation is expressed as a percentage relative to the comparison set size (12, 24, 48). As in Experiment 1, this decrease shows a strong linear correlation to the percentage of connected squares ($r^2 = .96$).

On the 100% connected displays, averaging across set sizes, participants perceived a display of fully connected squares to have 19% fewer squares than an unconnected display. If observers were counting exclusively over segmented groups of connected squares, we would expect the perceived number of squares in a fully connected display to be 50% the value of an unconnected display. Why would the value from the present experiment be lower? One reason is that the staircase procedure is designed to lead to near-chance responses on most trials. Participants may have been frustrated by their continual inability to distinguish differences between the displays and lost interest in this difficult task. Such a lack of focus at the most difficult (most similar) display combinations would lead the staircase to terminate prematurely, leading to a smaller observed effects of grouping strength.

To attain some measure of this difficulty effect, we revisited the results of Experiment 1. Although we did not systematically manipulate the relative ratio of unconnected and connected squares in that experiment, their ratios did vary across set sizes (due to the constraint of needing an even number of squares in the connected displays, see the Stimuli section of Experiment 1). Thus, the ratio of unconnected to connected squares ranged from 60% to 167%, with 10 values in between. We plotted the probability that the participant thought that the connected display contained more squares across each of these ratios and extrapolated the subjective point of equality. That is, at what ratio are there 50% 'more' responses and 50% 'less' responses? We found that participants perceived the fully connected displays to contain 32% fewer squares than the unconnected displays. This reduction is closer to the expected 50% and suggests that the overall difficulty of the staircase method likely resulted in an underestimation of the true effect size.

Other explanations for this "under-underestimation" apply to both Experiments 1 and 3. One possibility is that maintaining a response limit of 1 s does not completely influence strategic compensation for the grouping of objects into pairs. Even at the fast presentation times, there is a salient subjective difference between the connected and unconnected displays. The experience of this difference could cause participants to employ any number of strategies to compensate for the underestimation effect, such as inflating their perceptual estimate of the number of squares in the connected display.

Finally, it is possible that participants have access to segmented displays at both the grouped pair level and the individual square level. If so, their estimate of the connected display could be based on the average of these two estimates, leading to a hybrid estimate in between the two. Observers may try to select the individual squares, but when attempting to broadly select the whole collection, may be forced to include the more global grouped pair level (Navon, 1977). Future studies might explore this last possibility by manipulating the number of objects that are paired into a group, which should systematically change both the estimate of the segmented display and an average of the segmented and unsegmented displays. Note that even if both levels contribute to the representation used for number estimation, these results still imply that the global level *exists*, and that number estimation relies on such a segmented representation.

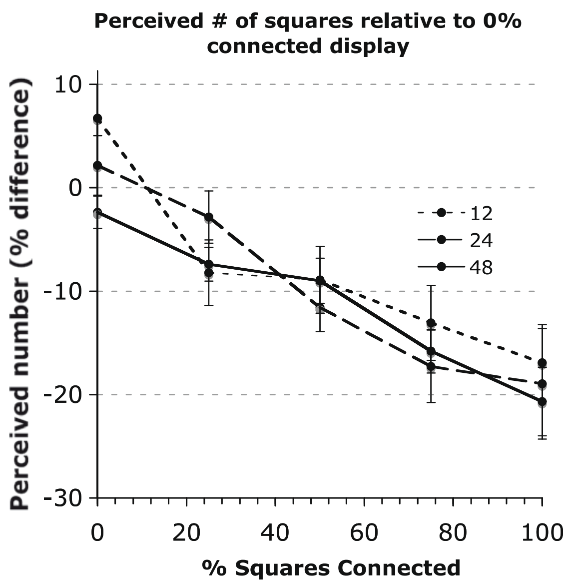

Fig. 6. In Experiment 3, as more pairs of squares were connected, the perceived number of unconnected squares judged to be equivalence in number dropped. This decrease was strongly linearly correlated to the increase in connected squares ($r^2 = .96$).

4. Experiment 4

The previous experiments demonstrate that connecting squares with a set of irrelevant lines causes observers to underestimate the number of squares in the display. However using four lines may have led observers to perceive the squares and lines as parts of as a three-dimensional extended cube. Also, using four parallel lines might have led to a stronger low-spatial frequency representation of a single large bar, which might alter unexpected aspects of the displays, such as object size or homogeneity of density.

To minimize the effect of connectedness on the display's subtended area and low-spatial frequency representation, Experiment 4, tests the effect of a smaller grouping cue – a single line 'broken' by a variable amount at its center. To avoid any three-dimensional interpretations of the objects, we connected objects at only their closest points and used circles instead of squares.

4.1. Participants

Twenty-two undergraduates participated in exchange for course credit, in addition to two authors.

4.2. Stimuli

The stimuli for this experiment were identical to that of Experiment 3, except that circles (0.7° in diameter with lines 1 pixel or 1.9 arcmin thick) were substituted for squares. Also, instead of four connecting lines, a single line (1 pixel or 1.9 arcmin thick) extended from the edge of each circle, originating at the point closest to its (potentially connected) pair. Connectedness was then manipulated by subdividing a connecting line at the midpoint between two circles and rotating the ends away from each other by 0°, 2°, 7°, 22°, or 50° clockwise relative to each object. In the unconnected displays, all connecting lines were subdivided by the maximum 50° rotation.

4.3. Procedure

The procedure was that same as in Experiment 3, except that only 17 reversals were required to terminate each staircase. The experiment lasted approximately 25 min.

4.4. Results and discussion

Two participants had below-chance performance in the 0% connected vs. 0% connected conditions and were removed from the analysis. Participants made their relative number comparison quickly, with an average response time of 491 ms. Response times decreased as set size increased. Trials where participants violated the 1 s response time limit (1.6% of trials) or responded in less than 100 ms (0.5% of trials) were omitted from the analysis.

When pitted against 50° displays, the perceived number of circles on 50° displays was 11.7, 24.8, and 47.1 circles for the 12, 24, and 48 circle displays. The remaining accuracy data were submitted to a 3 × 4 ANOVA with set size (12, 24, 48 objects in the connected display) and angle of separation (0°, 2°, 7°, 22°) as factors. There was a main effect of varying connectedness (i.e. rotating the angle of separation), $F(3, 63) = 13.63$, $p < .001$, $h^2 = .394$, but no main effect of set size ($F = 1.37$).

There was also an interaction between the two factors, $F(6, 126) = 2.26$, $p = .042$, $h^2 = .097$, indicating that decreasing the angle of separation in the connected displays led to greater underestimation for set sizes 12 and 24, relative to set size 48. This interaction was due solely to results at the 0° condition, for which there was more underestimation for set sizes 12 and 24 than for set size 48. One possible explanation for this effect is that, as only a single object was added or subtracted to the displays for each staircase step, the staircase may not have been able to shift far enough for this most extreme combination of set size and level of angle difference. (This effect might not have surfaced in Experiment 3 because the staircase was allowed to progress for a larger number of trials.) There was still a significant underestimation effect for displays at this set size, as the overall effect of angle of separation was strong for all three set sizes: set size 12, $F(3, 63) = 7.86$, $p < .001$, $h^2 = .272$, set size 24, $F(3, 63) = 4.97$, $p = .004$, $h^2 = .191$, and set size 48, $F(3, 63) = 3.35$, $p = .024$, $h^2 = .138$.

Even with a smaller segmentation cue, the amount of underestimation in Experiment 4 (18.7% fewer objects in the connected displays) was identical to the amount of underestimation found in Experiment 3. As depicted in Fig. 7, the largest change in the magnitude of the grouping effect is exhibited when the line rotation is decreased from 7° to 2°.

5. General discussion

In Experiment 1, we found that connecting a display of squares into pairs with lines led participants to greatly underestimate the number of squares present in the display, relative to when the connecting lines were broken. Participants underestimated connected displays despite explicit instructions that the lines were irrelevant and should be ignored throughout the task. Even practiced participants, including the authors, could not avoid greatly underestimating displays of connected elements. Number estimation appears to be influenced by a representation of a collection that is segmented into discrete objects. The results of Experiment 2 are inconsistent with the predictions of a spatial sampling strategy, suggesting that this segmentation is processed broadly across the display. Experiment 3 provides a direct method for assessing the strength of underestimation by measuring the perceived number of squares in a connected group.

Experiment 4 began to use number estimation as a tool to explore the rules governing the segmentation process. Instead of using squares connected by four lines, pairs of circles were connected by a single, pixel-thin line, which was 'broken' at a variably small angle. The underestimation effect appeared as strongly as before, in the maximally connected condition, but disappeared when the angle break reached 7°, suggesting that the mechanism underlying a sophisticated segmentation process sensitive to small manipulations.

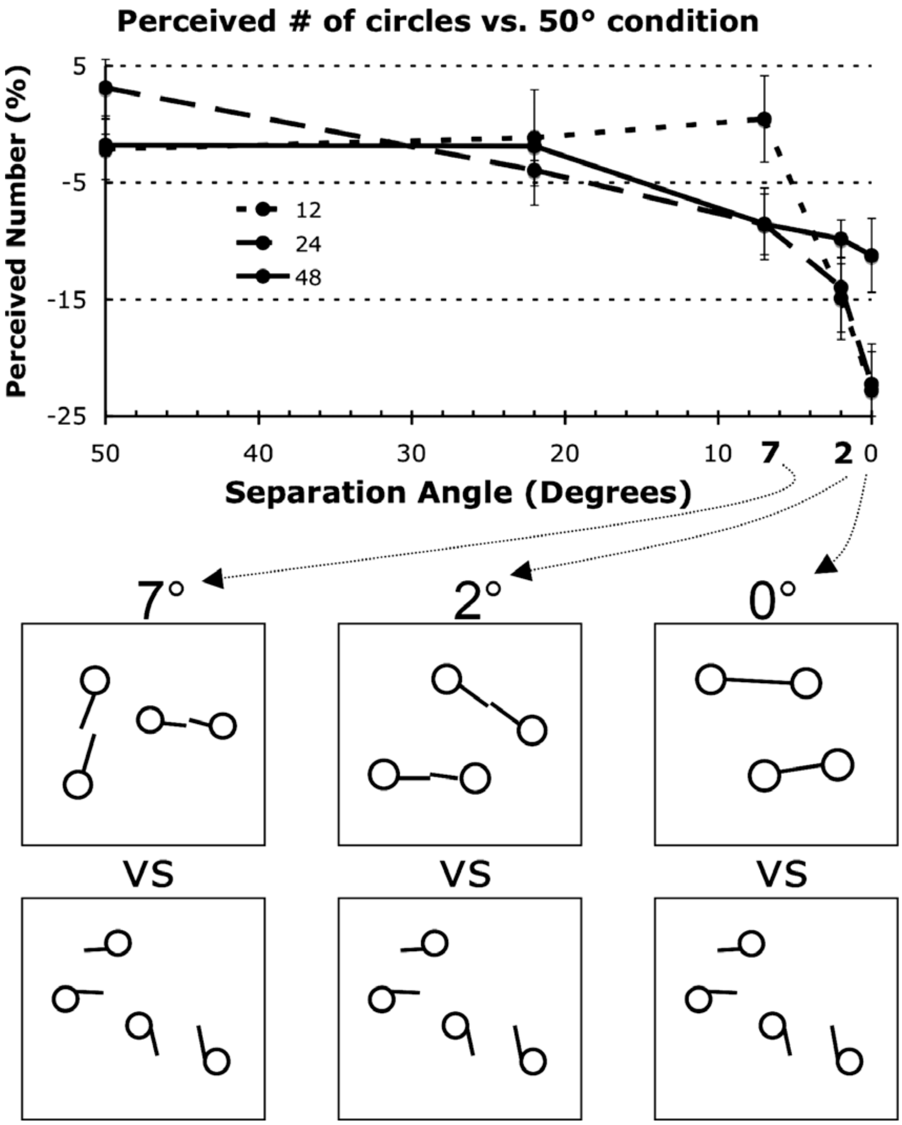

Fig. 7. In Experiment 4, the displays were altered from squares connected with four lines to circles connected by a single line, and the underestimation effect remained just as strong on average for the maximally connected condition. Also, as can be seen, the new displays allowed for a smaller manipulation of connectivity, demonstrating that a surprisingly small adjustment sufficed to produce the effect.

These grouping manipulations have only minimal effects on image properties that could be used as 'surrogates' for number perception. Factors such as covered area and spatial frequency profile were only slightly different across unconnected and connected displays in Experiments 1–3, and were virtually identical across display types in Experiment 4. One critique might be that the grouping manipulation changed another surrogate feature, the density of the display, by causing connected displays to appear less dense. This is only true if density is defined as objects per unit of area, which assumes a segmented representation. There are other definitions of density that do not require segmentation, such as those based on a spatial frequency profile. Under such a definition, there is only a tiny difference between unconnected and connected displays. Other density definitions such as texture density rely on second-order luminance comparisons at small spatial scales (Durgin, 2008; Durgin & Huk, 1997; Durgin & Proffit, 1996). While this cue is likely to be a useful correlate of number in high density arrays of small dots (Durgin, 1995), they would not predict the large-scale connectivity effect shown here.

Number estimates might also be based on the sum of fixed area contributions from each object, with the total area diminishing in spatially dense collections as area contributions increasingly overlap (Allik & Tuulmets, 1991). If this fixed contribution of each object stems ultimately from size of critical spatial filters in early vision (Allik & Tuulmets, 1991; Watt & Morgan, 1985), then the connectivity manipulation should not affect estimates. In contrast, if the fixed contribution stems from each object contributing a fixed amount of activation to a count (Dehaene & Changeux, 1993), then the occupancy model is by definition based on a segmented representation.

Visual number estimation appears to be influenced by a representation of the visual display that is broadly segmented into objects. Other image properties are still likely to play a strong role in the creation of estimates (Allik &

Tuulmets, 1991; Durgin, 1995), and there is no reason to ignore properties that correlate with number. But the present results show that there must also be a role for discrete objects representations.

Cognitive psychologists are often interested in isolating numerosity *per se* from other display factors like density, covered area, collection circumference, or spatial frequency. When determining whether young infants can represent numerical quantities, complex controls are needed to ensure that they are not relying instead on other display factors (e.g., Xu, Spelke, & Goddard, 2005). Other experiments test whether the visual system computes local representations of numerosity in the visual field, and additional control experiments are needed to differentiate these representations from those of other display factors (e.g., Burr & Ross, 2008a). Cognitive neuroscientists who seek to functionally localize mechanisms that process and compare number (e.g., Nieder, 2004; Pinel, Piazza, Le Bihan, & Dehaene, 2004; Shuman & Kanwisher, 2004) struggle to deconfound numerosity from other display factors. Manipulating perceived number through grouping cues may serve as an important future paradigm for researchers that study number perception and representation.

6. Broad segmentation

The finding that number estimation operates over a segmented collection is also important to our understanding of visual processing more generally because it suggests that the larger visual world is segmented broadly. A fundamental question about any visual process is whether it can occur broadly over the visual field, or whether the scope of the process is restricted to a narrow subset of incoming visual information (Neisser, 1967). In general, simpler processes are thought to occur broadly, while more complex processes are often relatively limited in extent. For example, uniquely colored objects can be found easily among homogenous distractors of a different color (Egeth, Jonides, & Wall, 1972), while searches requiring more sophisticated processing, such as finding an object with a certain spatial configuration (Wolfe & Bennett, 1997), are tediously slow, suggesting these processes operate over only subsets of visual information at a time.

One might expect that object segmentation cues as small as those shown in Experiment 4 might be implemented over only restricted subsets of visual information. At a local level, segmentation of an object from its background would help restrict processes like object recognition to only the area occupied by the object (Schneider, 1993). More generally, individual segmented objects (or small numbers of objects) can act as the inputs to other visual processes (see Scholl, 2001 for a review).

Yet, the present results suggest that a form of object segmentation is implemented for broad collections containing dozens of objects. Why would the visual system go to the computational trouble of implementing such sophisticated processing so broadly? The first reason, of course, might be to aid in number estimation. Real-world number estimation tasks do not always involve simple displays or circles, but instead more realistic 3D objects with complex part structures as well as figure-ground relationships and mutual occlusion. Cues such as covered area or spatial frequency might not suffice to generate reliable estimates in more naturalistic situations. Instead, estimation would be more reliable if each object were first represented as a single discrete unit (Dehaene & Changeux, 1993). There are proposals for computational or human visual mechanisms that could provide similarly abstracted representation of objects, such as medial axes (e.g., Blum, 1973), medial-point representations (e.g., Kovacs, Fehér, & Julesz, 1998), core representations (e.g., Burbeck & Pizer, 1995), and shock graphs (e.g., Siddiqi, Shokoufandeh, Dickinson, & Zucker, 1999). But to our knowledge there are no proposed mechanisms that take an unbroken visual image and extract a single fixed point for each object.

Another set of potential advantages of broad segmentation might be to provide an intermediate level of organization of visual information that can then be selected for subsequent processing. When performing a visual search, it would be helpful to know which features in a scene belong to actual objects and not the background (Wolfe, 1996). Other visual search experiments suggest that sophisticated grouping mechanism 'bundles' or 'clusters' local feature information into primitive objects broadly across the visual field, (Enns & Rensink, 1990; Rensink & Enns, 1995; Sun & Perona, 1996; Trick & Enns, 1997; Wolfe & Bennett, 1997). In one study, participants were asked to find a cross made up of a green vertical bar and a red horizontal bar among crosses with the opposite color combination, and search was slow and difficult. However, when all horizontal bars were linked together with lines, the task became easy, as if the horizontal bars were bundled into a single object that could be ignored, leaving the participant with the easy task of finding a uniquely colored green vertical bar (Wolfe & Bennett, 1997).

Another potential benefit of broad segmentation might be to *hide* information, by encapsulating the details of an object's structure and features, restricting initial analyses of global aspects of a scene to the 'big picture' (Rensink & Enns, 1995). In a visual search task, lines of unique length become difficult to find when embedded within Mueller–Lyer stimuli, but this detriment vanishes if a small gap is inserted at the end of each line segment, breaking the internal connectivity of the object (Rensink & Enns, 1995).

Because broad segmentation might serve an array of purposes in vision, future research should explore whether the segmented representation, and the cues used to create it, are identical across segmentation for separate processes such as number estimation, search guidance, or information encapsulation. Because the informational requirements of these processes differ, they might require subtly or strongly different representations of the visual field. The broad representation of objects in the visual field will likely also differ from the representations that we create for objects within the focus of attention, so it will be important to compare the effects of different cues for broad segmentation with those that effect phenomena such as object-based attention (Scholl, 2001), grouping (Palmer & Rock, 1994), and visual search (Wolfe & Bennett, 1997).

Acknowledgements

G.A.A. was supported by NEI/NIH fellowship F32 EY016982. We thank Patrick Cavanagh, Stanislas Dehaene, Laura Ortega, Julia Mossbridge, Lance Rips, Liz Spelke, and Jeremy Wolfe for helpful discussion.

References

Allik, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception and Psychophysics, 49*(4), 303–314.
Allik, J., Tuulmets, T., & Vos, P. G. (1991). Size invariance in visual number discrimination. *Psychological Research, 53*(4), 290–295.
Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. In *Proceedings of the National Academy of Sciences*, *106*(18), 7345–7350.
Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble features outside the focus of attention. *Psychological Science, 19*(4), 392–398.
Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12*(2), 157–162.
Barth, H. (2008). Judgments of discrete and continuous quantity: An illusory Stroop effect. *Cognition, 109*, 251–266.
Barth, H., Kanwisher, N., & Spelke, E. S. (2003). The construction of large number representations in adults. *Cognition, 86*, 201–221.
Beran, M. J., Taglialtela, L. A., Flemming, T. M., James, F. M., & Washburn, D. A. (2006). Nonverbal estimation during numerosity judgements by adult humans. *Quarterly Journal of Experimental Psychology, 59*, 2065–2082.
Bevan, W., Maier, R., & Helson, H. (1963). The influence of context upon the estimation of number. *American Journal of Psychology, 76*, 464–469.
Bevan, W., & Turner, E. (1964). Assimilation and contrast in the estimation of number. *Journal of Experimental Psychology: Human Perception and Performance, 67*, 458–462.
Birnbaum, M. H., Kobernick, M., & Veit, C. T. (1974). Subjective correlation and the size-numerosity illusion. *Journal of Experimental Psychology: Human Perception and Performance, 102*(3), 537–539.
Blum, H. J. (1973). Biological shape and visual science (part I). *Journal of Theoretical Biology, 38*, 205–287.
Burbeck, C. A., & Pizer, S. M. (1995). Object representation by cores: Identifying and representing primitive spatial regions. *Vision Research, 35*, 1917–1930.
Burr, D., & Ross, J. (2008a). A visual sense of number. *Current Biology, 18*(18), 425–428.
Burr, D., & Ross, J. (2008b). Response: Visual number. *Current Biology, 18*(18), R857–R858.
Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43*, 393–404.
Cohen, J. D., Aston-Jones, G., & Gilzenrat, M. S. (2004). A systems-level perspective on attention and cognitive control: Guided activation, adaptive gating, conflict monitoring, and exploitation vs. exploration. In M. I. Posner (Ed.), *Cognitive neuroscience of attention* (pp. 71–90). New York: Guilford Press.
Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience, 5*, 390–407.
Dehaene, S., & Cohen, L. (1994). Dissociable mechanisms of subitizing and counting: Neuropsychological evidence from simultanagnosic patients. *Journal of Experimental Psychology: Human Perception and Performance, 20*(5), 958–975.
Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.
Durgin, F. H. (1995). Texture density adaptation and the perceived numerosity and density of texture. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 149–169.
Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Current Biology, 18*(18), R855–R856.
Durgin, F. H., & Huk, A. C. (1997). Texture density aftereffects in the perception of artificial and natural textures. *Vision Research, 37*, 3273–3282.
Durgin, F. H., & Proffit, D. R. (1996). Visual learning in the perception of texture: Simple and contingent aftereffects of texture density. *Spatial Vision, 9*, 423–474.
Egeth, H., Jonides, J., & Wall, S. (1972). Parallel processing of multielement displays. *Cognitive Psychology, 3*(4), 674–698.
Enns, J. T., & Rensink, R. A. (1990). Influence of scene-based properties on visual search. *Science, 247*, 721–723.
Frith, C. D., & Frith, U. (1972). The solitaire illusion: An illusion of numerosity. *Perception and Psychophysics, 11*, 409–410.
Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44*, 43–74.
Ginsberg, N. (1978). Perceived numerosity, item arrangement, and expectancy. *American Journal of Psychology, 91*, 267–273.
Ginsberg, N., & Goldstein, S. R. (1987). Measurement of visual cluster. *The American Journal of Psychology, 100*, 193–203.
Ginsberg, N., & Nichols, A. (1988). Perceived numerosity as a function of item size. *Perceptual and Motor Skills, 67*, 656–658.
Halberda, J., Sires, S. F., & Feigenson, L. (2006). Multiple spatially-overlapping sets can be enumerated in parallel. *Psychological Science, 17*(7), 572–576.
Hollingsworth, H. H., Simmons, J. P., Coates, T. R., & Cross, H. A. (1991). Perceived numerosity as a function of array number, speed of array development, and density of array items. *Bulletin of the Psychonomic Society, 29*, 448–450.
Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences, 103*, 19599–19604.
Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43*, 171–216.
Kovacs, I., Fehér, A., & Julesz, B. (1998). Medial-point description of shape: A representation for action coding and its psychophysical correlates. *Vision Research, 38*, 2323–2333.
Krueger, L. E. (1972). Perceived numerosity. *Perception and Psychophysics, 11*, 5–9.
Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences, 12*, 72–79.
Lamb, M. R., & Robertson, L. C. (1989). Do response time advantage and interference reflect the order of processing of global- and local-level information? *Perception and Psychophysics, 46*(3), 254–258.
Lamb, M. R., & Robertson, L. C. (1990). The effect of visual angle on global and local reaction times depends on the set of visual angles presented. *Perception and Psychophysics, 47*(5), 489–496.
Logan, G. D., & Zbrodoff, N. J. (2003). Subitizing and similarity: Toward a pattern matching theory of enumeration. *Psychonomic Bulletin and Review, 10*, 676–682.
Miller, A. L., & Baker, R. A. (1968). The effects of shape, size, heterogeneity, and instructional set on the judgment of visual number. *American Journal of Psychology, 81*(1), 83–91.
Myczek, K., & Simons, D. J. (2008). Better than average: Alternatives to statistical summary representations for rapid judgments of average size. *Perception & Psychophysics, 70*(5), 772–788.
Nakayama, K. (1990). The iconic bottleneck and the tenuous link between early visual processing and perception. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 411–422). Cambridge University Press.
Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*, 353–383.
Neisser, U. (1967). *Cognitive psychology*. Englewood Cliffs, NJ: Prentice-Hall.
Nieder, A. (2004). The number domain – Can we count on parietal cortex? *Neuron, 44*(3), 407–409.
Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *The Encyclopedia of Neurobiology of Attention* (pp. 251–256). Elsevier.
Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin and Review, 1*(1), 29–55.
Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*(12), 1136–1169.
Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron, 41*(6), 983–993.
Rensink, R. A., & Enns, J. T. (1995). Preemption effects in visual search: Evidence for low-level grouping. *Psychological Review, 102*(1), 101–130.
Schiavetto, A., Cortese, F., & Alain, C. (1999). Global and local processing of musical sequences: An event-related brain potential study. *Neuroreport, 10*(12), 2467–2472.
Schneider, W. X. (1993). Space-based visual attention models and object selection: Constraints, problems, and possible solutions. *Psychological Research, 56*(1), 35–43.
Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition, 80*(1/2), 1–46.

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition, 80*(1/2), 159–177.

Shuman, M., & Kanwisher, N. (2004). Numerical magnitude in the human parietal lobe: Tests of representational generality and domain specificity. *Neuron, 44*(3), 557–569.

Siddiqi, K., Shokoufandeh, A., Dickinson, S., & Zucker, S. W. (1999). Shock graphs and shape matching. *International Journal of Computer Vision, 35*, 13–32.

Sophian, C. (2007). Measuring spatial factors in comparative judgments about large numerosities. In D. Schmorrow & L. Reeves (Eds.), *Foundations of augmented cognition: Third International Conference.* (pp. 157–165). Secaucus, NJ: Springer.

Sophian, C., & Chu, Y. (2008). How do people apprehend large numerosities? *Cognition, 107*, 460–478.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–661.

Sun, J., & Perona, P. (1996). Early computation of shape and reflectance in the visual system. *Nature, 379*, 165–168.

Taves, E. H. (1941). Two mechanisms in the perception of visual numerousness. *Archives of Psychology, 37*, 265.

Trick, L. M., & Enns, J. T. (1997). Clusters precede shapes in perceptual organization. *Psychological Science, 8*(2), 124–129.

Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 331–351.

Vos, P. G., van Oeffelen, M. P., Tibosch, H. J., & Allik, J. (1988). Interactions between area and numerosity. *Psychological Research, 50*(3), 148–154.

Watt, R. J., & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research, 25*(11), 1661–1674.

Wolfe, J. M. (1996). Vision: Resolving perceptual ambiguity. *Nature, 380*(18 April), 587–588.

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research, 37*(1), 25–43.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*(1), B1–B11.

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8*(1), 88–101.

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General, 130*, 29–58.