Applied Cognitive Psychology, *Appl. Cognit. Psychol.* (2010)

Enumeration: Experts Take Their Time

ROY ALLEN* and PETER MCGEORGE

The School of Psychology, College of Life Sciences and Medicine, University of Aberdeen, Kings College, Scotland, UK

Summary: Previous research has demonstrated that grouping information enhances enumeration, and that this advantage is significantly greater for observers with expertise in dynamic visuospatial tasks (e.g. air traffic controllers). We sought to elucidate whether this advantage is the result of over-learned, automated responding in an enumeration experiment where many of the stimuli were incongruent as to canonical arrangement and numerosity. If so, we predicted that experts' performance would be more severely affected than novices'; if not, and experts continued to perform better than novices, we predicted that their response times should increase, reflective of the additional cognitive load needed. We demonstrated the latter is so. Experts continued to out-perform novices but their response times were significantly slower suggesting that expertise is not rigid and automatic but, rather, is flexible and responsive to the specific situation, allowing experts to switch between strategies. Copyright © 2010 John Wiley & Sons, Ltd.

Enumeration tasks typically produce a distinctive pattern of response times and accuracy (e.g. Trick & Pylyshyn, 1993, 1994, but see, e.g. Balakrishnan & Ashby, 1992). When stimuli contain just a few items (up to 3 or 4), accuracy tends to be close to ceiling and response times are fast and relatively constant. Kaufman, Lord, Reese, and Volkmann (1949) coined the term *subitizing* to describe this type of rapid and accurate enumeration. With greater numbers, *counting* or *estimating* is required, and accuracy generally decreases and response times increase with each additional item in the display.

Mandler and Shebo (1982) have suggested that subitizing is achieved as a result of over-learned geometric cues in the arrangement of items in the display, leading to fast pattern recognition that accesses associated numerosity information. In an enumeration task, they presented displays arranged either randomly or in familiar patterns such as occur on the face of a die. Participants demonstrated a pattern recognition advantage in that they responded faster and more accurately to the familiar patterns than to random arrangements.

Why subitizing only appears to occur with displays containing a few items has been attributed to an increasing difficulty to discriminate between regular geometric shapes having more sides (i.e. pentagons, hexagons, etc., increasingly lose numerosity information as they tend towards being merely 'circular'). However, there is evidence to suggest that the advantage associated with pattern recognition can be extended through practice. For example, Mandler and Shebo (1982) reported that, after around 50, fixed-pattern trials, response times to displays containing more than 4 items fell significantly. This effect has also been demonstrated by Wolters, Van Kempen, and Wijlhuizen (1987), who tested participants' ability to enumerate displays consisting of between 4 and 18 items on each of 5 consecutive days. For one group, items were presented in different, random configurations on each day, for the other, in consistent patterns. Practice with the consistent-pattern stimuli led to both large decreases in response times and improvements in accuracy, while only small improvements were found with the random configurations.

Enumeration is improved not only following repeated exposure to specific patterns but also by experience of tasks in which there is a strong dynamic spatial component. For example, Green and Bavelier (2006) have shown that participants who played action-video games for just '. . . one hour per day for ten out of fifteen days . . ..' (p. 229), showed a significant improvement in their enumerating ability when compared to similar participants who played Tetris^{TM}, a non-action-video game, for the same period. Green and Bavelier linked the enhanced enumeration skills to changes in the availability and distribution of attentional resources resulting from the dynamic spatial nature of action-video games (but see Boot, Kramer, Simons, Fabiani, & Gratton, 2008). However, it is not clear whether these changes extend to a greater use of pattern information.

Green and Bavelier (2006) also demonstrated that action-video game players out-performed non-action-video game players at a multiple-object tracking (MOT) task, something that '. . . requires attention to be allocated to several objects over time' (p. 217) (see also Allen, McGeorge, Pearson, & Milne, 2004). This link between enumeration and MOT had been previously identified by Trick and Pylyshyn (1994) who suggested that enumeration and MOT both involve the deployment of a limited number of spatial visual indexes that 'tag' items in the visual field for subsequent processing.

To examine directly whether the enhanced enumeration skill seen after experience with dynamic spatial tasks was linked with a greater use of pattern information, Allen and McGeorge (2008) compared the performance of air traffic controllers (ATCs) and matched novices on an enumeration task. There are two aspects of the task performed by ATCs that lead to the expectation that they will show enhanced use of pattern information. Firstly, Yantis (1992) has previously shown that one way in which individuals track multiple objects is to group the objects into a higher-order virtual object and it is this virtual object that is then tracked. Yantis has also shown that individuals faced with tracking multiple objects discover this strategy with experience. Since the role of the

---

*Correspondence to: Roy Allen, The School of Psychology, College of Life Sciences and Medicine, University of Aberdeen, William Guild Building, Kings College, Scotland AB24 2UB, UK. E-mail: roy.allen@abdn.ac.uk

ATC is to track multiple objects on computer displays and, given the similarity between this task and the experimental paradigm used in MOT, it is not unreasonable to suggest that ATCs develop an enhanced ability to form and make use of grouping information. Secondly, unlike the movement of items in an experimental MOT task, the aircraft tracked by ATCs do not move in a random manner but instead obey lawful rules (i.e. fixed flight corridors, etc.). It might be expected that this experience would further enhance reliance on pattern information in tracking.

Allen and McGeorge (2008) presented ATCs and novices with stimuli (+'s) arranged in either regular geometric shapes, with each one being apical, or in straight lines. They hypothesized that, if the advantage shown by ATCs on the MOT task was related to pattern recognition processes, then any advantage in an enumeration task might be greater where the stimuli consist of regular/canonical patterns, particularly since we are biased towards regular forms (Feldman, 2000). They found that ATCs were significantly better than novices, but only where arrangement/grouping information was available to them (i.e. they were no better than novices on linear arrays). Allen and McGeorge speculated that whilst linear arrays do hold unique numerosity information, like canonical arrays, they only do so when comprised of just the minimum number of items necessary to define their 'shape' (i.e. for a line, no more than 2 items).

If shape does inform enumeration then this raises the possibility that, when arrangement and numerosity conflict, such as when a shape contains more than the number of items necessary to indicate its apices (e.g. 5 in a square), enumeration would be affected detrimentally. In addition, since experts such as ATCs seem to make greater use of shape to inform enumeration, this might suggest that shape/numerosity incongruence would affect them more profoundly. If the expertise effect is the result of an over-learned, automatic process, then accuracy should be reduced, whilst if expertise persists, suggesting it is associated with a greater attentional flexibility (Bellenkes, Wickens, & Kramer, 1997), then longer reaction times; such as have been observed in other experts, particularly when they have been presented with incongruent stimuli (Drost, Rieger, Brass, Gunter, & Prinz, 2005, Keller & Koch, 2008, Rieger, 2007), should be manifested. In order to settle this question, we examined the effect of incongruence between arrangement and numerosity, by presenting stimuli deployed in regular geometric shapes, which frequently contained more than the minimum number of items necessary to define them (e.g. 4 items arranged to form a triangle, etc.).

METHODS

Participants

Twenty-eight undergraduate students (14 females) at the University of Aberdeen aged between 19 and 27 ($M = 21.9$, $SD = 2.97$), and 7 military radar operators (4 females) aged between 18 and 29 ($M = 24.3$, $SD = 3.70$), all but one of whom had graduated from university, took part. The military personnel were tested at their base. All have normal or corrected-to-normal vision. Note that 4 novices (3 females) were excluded from the analyses because their performance was beyond 2 SDs of mean performance.

Materials

All stimuli were prepared in advance on Microsoft's Powerpoint and saved as bitmaps for presentation. Each stimulus, being black on a white background, contained a number of identical items ('+'s in bold, 18 point Times New Roman), subtending a visual angle of approximately 1.25° at a viewing distance of approximately 57 cm. Item numbers ranged from 2 to 8, with their extent being always the circumference of a notional 50 mm diameter circle (5° of visual angle) centred in the middle of the screen. All stimuli were arranged in regular geometric shapes (i.e. line, triangle, square, pentagon and hexagon) with equal sides. However, the number of items ('+'s) per shape was manipulated (see Table 1).

Linear patterns were lines of '+'s, always arranged equidistantly along the full length of the notional 50 mm circle's diameter. Where the sides of a geometric shape contained an extra item, this was placed at the centre of the side. Examples of the various stimulus configurations are shown in Figure 1. (Note, that whilst the co-linearity of parts of the pluses (+'s) might facilitate the percept of a square this was neither expected to be significant; nor does the situation arise from any other arrangement.)

Each trial began as a static frame consisting of a centralized black fixation letter 'o' subtending a visual angle of 0.42°, on a white background subtending a visual angle of 21.5°. After a delay of 500 milliseconds, a stimulus was presented for 20 milliseconds before being replaced by a screen display of black text on a white background with the instruction that participants should respond as to how many items the stimulus image had contained.

Trials were displayed using Superlab Pro (Cedrus Corporation, San Pedro, CA) running on a PC with a 17-in. monitor at a viewing distance of approximately 57 cm.

Table 1. Stimulus matrix

No. of items	Arrangement				
	Nominal shape	Nominal + 1 shape	Nominal + 2 shape	Nominal + 3 shape	Nominal + 4 shape
2	Line				
3	Triangle	Line			
4	Square	Triangle	Line		
5	Pentagon	Square	Triangle	Line	
6	Hexagon	Pentagon	Square	Triangle	Line
7		Hexagon	Pentagon	Square	
8			Hexagon	Pentagon	Square

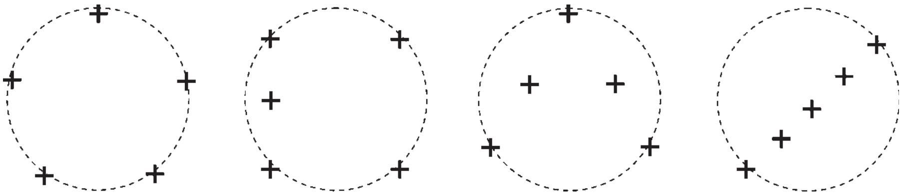

Figure 1. Notional 50 mm diameter circle at centre of stimulus image showing, from left to right, for 5 items, the nominal shape (pentagon), nominal + 1 shape (square), nominal + 2 shape (triangle), nominal + 3 shape (line)

Design

All participants undertook an enumeration task in which the number of items (2–8) and arrangement (i.e. from nominal to nominal + 4, see Table 1) were systematically manipulated. To minimize pattern learning during the study, several versions of every arrangement were created, the orientation of each differing by its degree of rotation. This was achieved by maintaining the orientation of the nominal shape and inserting the additional items into different locations. Order of presentation was completely randomized by the presenting software. In total, there were 630 trials, being 30 each for the 21 different number × arrangement configurations (see Table 1), that took approximately 30–45 minutes to complete. Accuracy and reaction time data were collected. No practice trials were administered.

Procedure

In order to minimize extraneous visual distractors, participants were tested individually in an unlit, windowless room. Each participant was instructed that they were to be presented with a series of images that each contained a number of identical items. For each image, once presented, they were simply to indicate, using the numeric keys, how many items it had contained. They were also cautioned that they might sometimes feel they have not been shown an image because the presentation time had been so short. They were assured that every trial contained an image, and that every image contained at least 1 and not more than 9 items, and that they should always give a response, even when they had not consciously registered an image, based on what they *felt* was the correct number. Participants were misled as to the maximum number of items per image so as to reduce the ‘end-effect’ where participants presented with a stimulus containing numerous items will simply respond with what they have been told is the upper limit of items in the display (Simon, Peterson, Patel, & Sathian, 1998).

RESULTS

As can be seen from Table 1, stimuli could not complete the entire matrix. Thus two regions were selected for analysis, one to examine a wide range of object number (3–6) across just nominal and nominal + 1 arrangements; the other to examine nominal, nominal + 1 and nominal + 2 arrangements across just two numbers (5, 6) (thus no analyses of nominal + 3 and nominal + 4 arrangements are reported here). In both cases, separate repeated-measures ANOVAs of accuracy and RTs to correct responses were carried out, with number of items and arrangement as within-subjects factors and group (novice/expert) as the between-subjects factor. An alpha level of .05 was used at all times and, wherever appropriate, the Greenhouse–Geisser correction was applied.

For the first region’s accuracy data, the expected main effect of number of items was found ($F(1.56, 45.32) = 10.24$, $p = .001$, $\eta_p^2 = 0.26$) in that accuracy declined as the number of objects increased. More importantly, there was also a significant effect of arrangement ($F(1, 29) = 4.23$, $p = .049$, $\eta_p^2 = 0.13$) such that participants were significantly more often correct when enumerating nominal, rather than nominal + 1, arrangements. Finally, there was a main effect of group ($F(1, 29) = 5.12$, $p = .031$, $\eta_p^2 = 0.15$) in that experts ($M = 97.02$, $SD = 1.39$) performed significantly better than novices ($M = 93.45$, $SD = 0.75$).

For the RT data, there were main effects of both number of items and arrangement, but these were moderated by a significant interaction ($F(1.85, 53.58) = 7.20$, $p = .002$, $\eta_p^2 = 0.20$). However, *post hoc t*-tests (p set to .01 for multiple comparisons) found only significant differences at the .05 level. There was also, again, a main effect of group ($F(1, 29) = 6.69$, $p = .015$, $\eta_p^2 = 0.19$) in that experts ($M =$ 1192 milliseconds, $SD = 63.91$) took significantly longer to respond than novices ($M = 1005$ milliseconds, $SD = 34.52$).

For the second region’s accuracy (% correct) data there was a main effect of arrangement ($F(2, 58) = 3.76$, $p = .029$, $\eta_p^2 = 0.11$) in that performance on nominal arrays ($M = 95.04$, $SD = 1.09$) was significantly greater than that for nominal + 2 arrays ($M = 89.13$, $SD = 1.92$) (pairwise comparison, $p = .013$). Further, there was a significant interaction between number of items and group ($F(1, 29) = 6.19$, $p = .019$, $\eta_p^2 = 0.18$). *Post hoc* paired-sample *t*-tests, based on the Bonferroni correction ($p < .025$ for multiple comparisons), showed that whilst experts’ accuracy remained constant across 5- and 6-item arrays ($t(6) = 1.88$, ns), novice performance declined significantly ($t(23) = 5.90$, $p < .001$).

Finally, for the same region’s RT data there were main effects of number of items, arrangement and group, but these were moderated by significant interactions. Firstly, there was an interaction of number of items × arrangement ($F(1.59, 46.26) = 3.66$, $p = .043$, $\eta_p^2 = 0.11$) in that RTs to 6-item arrays increased faster than those to 5-item arrays (see Figure 2).

Secondly, there was an interaction of arrangement × group ($F(2, 58) = 4.08$, $p = .022$, $\eta_p^2 = 0.12$). Subsequent independent-sample *t*-tests showed that whilst novices’ and experts’ reaction times did not vary significantly for congruent (nominal) trials ($t(29) = 1.92$, ns), experts were significantly slower than novices for all incongruent trials (nominal + 1 ($t(29) = 2.30$, $p = .029$); nominal + 2 ($t(29) = 3.14$, $p = .004$) (see Figure 3).

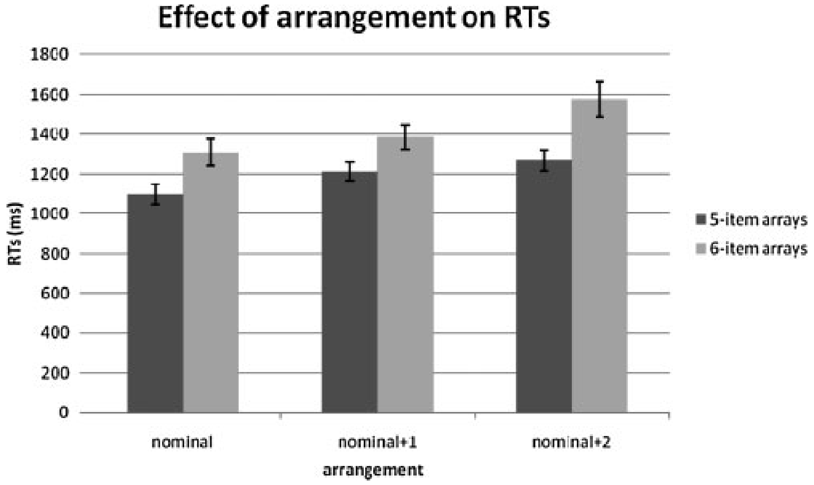

Figure 2. Participants' RTs as a function of arrangement for 5- and 6-item arrays

Note that the purpose of manipulating the orientation of the stimuli's elements was only to reduce the tendency of participants to spontaneously learn a pattern recognition strategy whilst they took part in the experiment. Number of trials in each orientation were, therefore, insufficient for any meaningful analysis. Note, also, that the total number of times participants responded '9' to a trial, due to being misled as to the maximum number of items in any one trial, varied from 0 to 25 ($M = 8.6$, $SD = 6.48$).

DISCUSSION

Allen and McGeorge (2008) carried out an enumeration experiment in which stimuli contained identical objects arranged either as the apical points of a canonical shape or as a linear array. They found that individuals (ATCs) who have accumulated experience at a visuospatial task outperformed novices, but only when the objects were arranged canonically; they were no better on linear arrays. Allen and McGeorge concluded that experts made more use of canonical shape information to inform numerosity.

Here, we hypothesized that reliance upon the link between shape and numerosity might affect experts more detrimentally when shape no longer informed numerosity. This detrimental effect was expected irrespective of increased crowding (as a result of increasing numerosity while maintaining display size) as this applied equally to both novices and experts. If such an ability is inflexible and automated, then we would predict that the expertise effect would either be significantly reduced or lost altogether; if not, then the cognitive cost of maintaining the expertise effect would necessitate the need for slower reaction times. To test this, we presented novice and expert participants with stimuli where either the objects' arrangement directly informed numerosity (congruent) or where it did not (incongruent). Experts were no more affected by incongruent trials than novices and continued to outperform them across all arrangements, achieving this by maintaining a high level of accuracy across increasing number of objects, as novice performance declined. In the current study the stimulus presentation was not followed by a pattern mask, only by a new screen containing text and therefore it is possible that performance might be enhanced by access to iconic memory representations. However, there are no reasons to expect that any effect of access to iconic memory will differentially influence novices and experts. Indeed, previous research using the related task of MOT (Barker, Allen, & McGeorge, 2010) has indicated that experts do not maintain

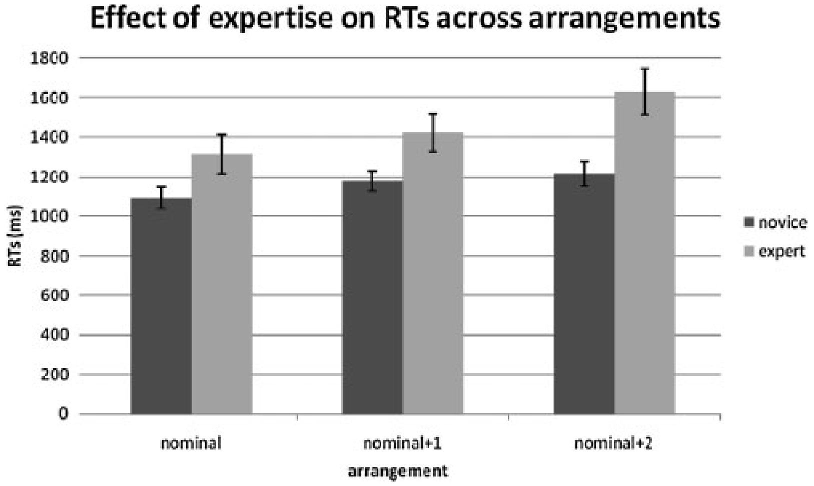

Figure 3. Participants' RTs as a function of arrangement for novices and experts

superior performance following a pre-response delay as might be expected if they were able to make superior use of iconic memory.

However, experts were consistently slower to respond than novices, sometimes by as much as 400 milliseconds. This is in sharp contrast to RT data collected during the study by Allen and McGeorge (2008), but not reported at the time, which showed no significant difference in reaction times between groups of experts ($N = 18$) and novices ($N = 18$) ($F(1, 34) = 2.94$, ns) on an enumeration task that *only* contained congruent stimuli (experts: $M = 1273$ milliseconds, $SD = 46.23$ milliseconds; novices: $M = 1161$ milliseconds, $SD = 46.23$ milliseconds). This, then, suggests that the expertise effect in enumeration is not automated but, rather, is flexible and permits attentional access.

A number of researchers have recently suggested a specific attentional component for subitizing (Egeth, Leonard, & Palomares, 2008; Olivers & Watson, 2008; Poiese, Spalek, & Di Lollo, 2008) and reaction times are often seen as an attentional metric (i.e. the longer a task takes the more attentionally demanding it is considered to be). Given that experts in this experiment maintain a greater accuracy than novices, we cannot interpret their slower responses as being purposeless guesses, as might be alleged for novices' quicker responses. It would also be unusual to suggest that professional ATCs might be less motivated than student participants. Since it takes time to mobilize and focus attention, clearly experts are doing something more attentionally demanding in order to achieve greater accuracy. And, since we have shown previously that experts are more accurate by dint of making greater use of shape for numerosity information, it seems reasonable to infer that, when faced with shape/numerosity incongruence, experts are devoting additional time for a strategically beneficial purpose.

Allen and McGeorge (2008) have demonstrated that ATCs make greater use of canonical shape information when enumerating items in briefly presented displays. It is the enhanced sensitivity to grouping information as a result of the nature of their experience which may serve to highlight the mismatch between a canonical shape, having uniform distances between the adjacent apices, and arrangements in which this rule is violated (i.e. a square with an extra item on one of its sides) alerting experts to switch to a less pattern orientated strategy such as counting and thus resulting in increased response times.

REFERENCES

Allen, R., & McGeorge, P. (2008). Enumeration: Shape information and expertise. *Acta Psychologica*, *129*, 26–31.

Allen, R., McGeorge, P., Pearson, D., & Milne, A. B. (2004). Attention and expertise in multiple target tracking. *Applied Cognitive Psychology*, *18*, 337–347.

Balakrishnan, J. D., & Ashby, F. G. (1992). Subitizing-magical numbers or mere superstition. *Psychological Research-Psychologische Forschung*, *54*, 80–90.

Barker, K., Allen, R., & McGeorge, P. (2010). Multiple object tracking: Enhanced visuospatial representations as a result of experience. *Experimental Psychology*, *57*, 208–214.

Bellenkes, A. H., Wickens, C. D., & Kramer, A. F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation Space and Environmental Medicine*, *68*, 569–579.

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica*, *129*, 387–398.

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005). When hearing turns into playing: Movement induction by auditory stimuli in pianists. *Quarterly Journal of Experimental Psychology*, *58A*, 1376–1389.

Egeth, H. E., Leonard, C. J., & Palomares, M. (2008). The role of attention in subitizing: Is the magical number 1? *Visual Cognition*, *16*, 463–473.

Feldman, J. (2000). Bias toward regular form in mental shape spaces. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 152–165.

Green, C. S., & Bavelier, D. (2006). Enumeration versus multiple object tracking: The case of action video game players. *Cognition*, *101*, 217–245.

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

Keller, P. E., & Koch, I. (2008). Action planning in sequential skills: Relations to music performance. *Quarterly Journal of Experimental Psychology*, *61*, 275–291.

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

Olivers, C. N. L., & Watson, D. G. (2008). Subitizing requires attention. *Visual Cognition*, *16*, 439–462.

Poiese, P., Spalek, T. M., & Di Lollo, V. (2008). Attentional involvement in subitizing: Questioning the preattentive hypothesis. *Visual Cognition*, *16*, 474–485.

Simon, T.J., Peterson, S., Patel, G., & Sathian, K. (1998). Do the magnocellular and parvocellular visual pathways contribute differentially to subitizing and counting? *Perception & Psychophysics*, *60*, 451–464.

Rieger, M. (2007). Letters as visual action-effects in skilled typing. *Acta Psychologica*, *126*, 138–153.

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*, 331–351.

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*, 1–23.

Wolters, G., Van Kempen, H., & Wijlhuizen, G. J. (1987). Quantification of small numbers of dots-subitizing or pattern-recognition? *American Journal of Psychology*, *100*, 225–237.

Yantis, S. (1992). Multi-element visual tracking: Attention and perceptual organization. *Cognitive Psychology*, *24*, 295–340.