

Effects of color heterogeneity on subitization

DERRICK G. WATSON and ELIZABETH A. MAYLOR
University of Warwick, Coventry, England

Small numbers of items can be enumerated rapidly and accurately via a process termed *subitizing*. In two experiments, we examined the effect of target color heterogeneity on subitizing efficiency. In contrast to the findings of Puts and de Weert (1997), we found that observers were no less efficient at subitizing displays containing red and green items than they were at subitizing displays of a single color. We propose that these findings are consistent with subitization operating on items represented within a location master map that codes where objects are but not what they are. The data are discussed in relation to recent findings concerning the detection of single-feature targets and the functional architecture of early visual processing.

There is a substantial difference in performance in the visual enumeration of relatively small numbers of items in comparison with that of larger ones. Up to about three or four items, enumeration is rapid and accurate and there is relatively little increase in response times (RTs) as a function of numerosity (typically, <40 msec/item). Beyond this range, RTs increase substantially (~350 msec/item) and errors are more likely to occur (see, e.g., Trick & Pylyshyn, 1993, 1994). Such efficient enumeration for small numerosities has been termed *subitization* (Kaufman, Lord, Reese, & Volkman, 1949).

One relatively recent account of subitizing (Trick & Pylyshyn, 1993, 1994) proposes that it arises as a consequence of the visual system's ability to tag a limited number (approximately four) of preattentively individuated elements with indexes termed FINSTs (fingers of instantiation; Pylyshyn, 1989, 2001; Pylyshyn & Storm, 1988; for an overview, see Pylyshyn, 2000). Focused serial attention can then be rapidly moved around the tagged items, allowing further, more detailed processing. The FINST system provides a mechanism whereby multiple identical objects in a scene can be individuated and allows a number of other visual functions (e.g., determining the spatial relationships between objects) to be performed. According to the FINST theory of subitizing, a small number of items can be simultaneously tagged, and then numerosity can be derived by linking each bound FINST to a number name. However, when target numerosity exceeds the number of available tags, the FINSTs must be reassigned, serial attention must be moved, a running total must be maintained, and previously enumerated items must be prevented from being reprocessed. These additional op-

erations lead to substantial increases in RT for each additional item enumerated.

Of most relevance to the present work, Trick and Pylyshyn (1993) found that observers were able to successfully subitize targets among distractors provided that the targets differed from the distractors at the level of a single feature difference (e.g., targets and distractors differed in color). In contrast, when targets differed from distractors only by a unique combination of two features (i.e., conjunction targets; Treisman & Gelade, 1980), subitization did not occur (see also Watson, Maylor, & Bruce, 2005a, and Watson, Maylor, & Manson, 2002, for related findings). Thus, FINSTs can be selectively applied to items represented within a particular primitive feature map in order to exclude distractors represented in other feature maps (Trick & Pylyshyn, 1993).

In contrast to Trick and Pylyshyn (1993), Puts and de Weert (1997) examined whether or not there was a cost to subitizing items represented in two different feature maps (enumerating red and green dots) in comparison with items represented in a single feature map (enumerating red dots). In other words, would subitizing be less efficient when the stimuli were of two colors than when they were of a single color? In FINST terms, this amounts to whether or not it is more difficult to apply FINSTs to stimuli represented across two different feature maps than to stimuli represented within a single feature map. The results showed that, overall, observers were slower to determine that a display contained four dots (as opposed to three or five dots) when the dots were red and green in comparison with when they were all red. However, this occurred only when the dots were arranged so that they did not form reliable shape cues for different numerosities. When the dots were presented in patterns that provided a cue to each numerosity, observers were no slower to subitize displays of mixed-color dots (red + green) than displays of dots of a single color (red). Puts and de Weert concluded that subitization in mixed-color displays was less efficient than subitization in single-color displays unless the dots formed patterns that could be used as reliable

This research benefited from Grant R000239180 from the Economic and Social Research Council of Great Britain. We thank Khalilah Khan for assistance with data collection and Lana Trick, Jeremy Wolfe, and an anonymous reviewer for useful comments on an earlier version of the article. Correspondence concerning this article should be addressed to D. G. Watson, Department of Psychology, University of Warwick, Coventry CV4 7AL, England (e-mail: d.g.watson@warwick.ac.uk).

cues to numerosity. In this case, each dot configuration effectively became a primitive feature, and enumeration proceeded via pattern recognition (see Mandler & Shebo, 1982).

On the basis of previous findings, reduced subitization efficiency for mixed-color displays is somewhat unexpected. For example, in visual search each of two or more single-feature defined targets can be detected in a spatially parallel way and as efficiently as one single-feature defined target, provided that they are defined within the same stimulus dimension (e.g., orientation; Müller, Heller, & Ziegler, 1995; Treisman & Gelade, 1980; Wolfe, Butcher, Lee, & Hyle, 2003). This suggests that monitoring for activity across multiple feature maps of the same stimulus dimension does not produce a performance cost. Similarly, there appears to be no cost to subitizing squares of unequal size in comparison with squares of equal size (Trick & Pylyshyn, 1993). Indeed, in some studies stimulus heterogeneity has been shown to improve enumeration performance, especially in the postsubitizing range (Frick, 1987). Finally, Randall (1998) showed that enumerating red and green dots was no less efficient (in terms of either overall RTs or subitization slope) than selectively enumerating dots of one of the colors.

However, there are several features of the Puts and de Weert (1997) study that may have increased the possibility of finding an artifactual difference between single-color and mixed-color subitization. First, the mixed-color condition contained red and green dots and performance in this condition was compared to performance in a single-color condition that consisted of only red dots. If observers were slower to process the green dots (perhaps because they were less salient), this could lead to longer RTs in the mixed-color than in the single-color condition (see Randall, 1998, for a relevant related finding). However, this would not be attributable to the need to process items from different feature maps, but rather to processing one of the colors more slowly than the other and to the fact that the mixed-color condition was compared only to the single-color condition in which the RTs were shorter. Also, the stimuli were briefly presented and were not masked, which could have exaggerated any differences in color salience.

Second, in order to keep dot configuration from acting as a numerosity cue, Puts and de Weert (1997, Experiment 2) presented all numerosities in the form of triangles. However, global properties of the displays (e.g., item density, area, amount of redness) could still act as cues to numerosity, and such properties are likely to be less salient in mixed-color than in single-color displays due to disruption of the overall configuration by local color-based grouping and segmentation processes.

Third, a two-alternative forced choice (2AFC) procedure was used, in which observers indicated whether the display contained four items as opposed to either three items or five items. This procedure would allow observers to perform the task by setting upper and lower limits on some global property of the display, such as whether it contained a given amount of overall luminance or redness

(equivalent to four items), rather than requiring subitizing per se. This would again favor conditions in which display properties such as stimulus color and luminance were homogeneous; the brief display exposure may have encouraged the use of such strategies.

Fourth, several previous studies have shown that the subitizing limit may range from three to four items (see, e.g., Trick & Pylyshyn, 1993). Thus, it is likely that the range of numerosities of three to five items presented by Puts and de Weert (1997) fell outside the subitizing range for many observers. Finally, it is of note that Puts and de Weert measured overall RTs, which allowed them to determine whether or not subitizing single- versus mixed-color displays was slower *overall*. However, their study could not determine whether the *rate* of subitizing was affected. It remains possible that even if mixed displays produce an overall increase in RTs, the rate and, therefore, the efficiency of subitization remains equivalent to those in single-color displays.

In the present study, we reexamined the subitization of single- versus mixed-color stimuli. In contrast to Puts and de Weert (1997), we (1) determined single-color subitization for both sets of colors used in the mixed-color condition in order to rule out the possibility that differences in the processing rates of each individual color could account for poorer performance in the mixed-color condition, (2) used an unlimited but postresponse masked display procedure, (3) used a larger range of response alternatives (4AFC as opposed to 2AFC) and examined subitization for the typical range of subitizable numerosities (1–4), and (4) used displays in which the configuration of the dots was generated randomly on each trial in order to rule out the possibility of using reliable spatial properties of the displays as a cue to numerosity. Finally, in addition to measuring the effect of single- versus mixed-color subitization on overall RTs, we examined the effect of manipulating display heterogeneity on subitizing rates.

EXPERIMENT 1

There were two main conditions in Experiment 1. In the single-color condition, observers enumerated one to four dots that were unpredictably either all red or all green on any given trial. In the mixed-color condition, the displays consisted of a mix of both red and green dots.

Method

Participants. Eighteen students (14 female) 18–43 years of age ($M = 21.8$ years) from the University of Warwick took part in the study for payment of £3. All had self-reported normal color vision and normal or corrected-to-normal visual acuity.

Stimuli and Apparatus. Stimuli were generated and presented by custom programs running on a 1-GHz Pentium-based computer attached to a 19-in. monitor (visible diagonal, 445 mm [40.7°]). The stimuli were generated by placing one to four solid-color discs (10-mm [0.95°] diam) randomly into the cells of an invisible 6×6 matrix. The stimulus colors were red (RGB = 200,0,0; CIE $xy = 0.619, 0.345$; luminance = 13.2 cd/m²) and green (RGB = 0,170,0; CIE $x_1y = 0.285, 0.609$; luminance = 25.7 cd/m²).¹ Center-to-center interelement spacing was 54 pixels (24 mm [2.29°]) horizontally and

58 pixels (26 mm [2.48°]) vertically, and elements were jittered by up to ±15 pixels along the *x*- and *y*-axes. On each trial, a mask stimulus consisting of 36 discs was generated, with the color of each disc chosen randomly between red and green. All stimuli were presented against the black screen of the computer monitor (0.5 cd/m²). Display resolution was 800 × 600 pixels, and display changes were synchronized to the screen retrace (60 Hz). The experiment was conducted in a sound-attenuated, dimly illuminated laboratory, and displays were viewed from a distance of approximately 60 cm, although no mechanical means was used to restrict head movements.

Design and Procedure. The experiment included two conditions: single color and mixed color. In the single-color condition, the stimuli were either all red or all green and all numerosities and color combinations were presented equally often. In the mixed-color condition, the stimuli consisted of a mix of red and green discs, and each possible ratio of red:green discs was presented equally often at each numerosity (with the exception of numerosity 1, which was equally often a single red disc or a single green disc). In both conditions, a trial consisted of a blank screen (500 msec) followed by a central white fixation point (3 × 3 mm [0.29° × 0.29°], 1,000 msec) followed by an enumeration display containing one to four solid discs. The participants enumerated the discs and then pressed the space bar, which served to replace the stimulus display with the mask. They then entered their response by pressing the “1” to “4” keys on the computer keypad, which initiated the next trial. RTs were measured from the onset of the enumeration stimulus until the pressing of the space bar. This method eliminates many of the problems associated with using a voice key methodology (Rastle & Davis, 2002) or directly assigning numerosities to multiple response keys (for previous uses of this method, see Atkinson, Campbell, & Francis, 1976; Svenson & Sjöberg, 1983; Trick & Enns, 1997; Watson & Humphreys, 1999; Watson et al., 2002).

The participants completed one block of 144 trials for each of the single- and mixed-color conditions. Every 36 trials, the participants were prompted to take a short break and then to press any key to continue. Condition order was counterbalanced across participants, and a 24-trial practice block directly preceded each full block.

Results

RTs and slopes. Fourteen anticipatory RTs of <100 msec (<0.3%) were discarded. Median correct RTs² were then calculated for each cell of the design and for each participant, and the overall means of these medians are shown in Figure 1. As in previous studies (Trick & Pylyshyn, 1993, 1994; Watson et al., 2002), the largest numerosity (in our case, 4) was not included in any subsequent analyses because of possible end effects. In addition, we wanted to maximize the possibility that we were measuring subitizing performance in all observers. As is indicated in Figure 1, RTs tended to increase as a function of numerosity, but there were no overall RT or slope differences across conditions. This was confirmed by a 3 (stimulus type: red vs. green vs. mixed) × 3 (numerosity: 1–3) within-subjects ANOVA,³ which revealed a marginally significant main effect of numerosity [$F(1,16, 19.64) = 3.27, MS_e = 3,112.31, p = .081$]; RTs increased overall by 8.5 msec/item.⁴ However, neither the main effect of condition (overall means: single color red, 433.3 msec; single color green, 434.4 msec; mixed color, 433.1 msec; $F < 1$) nor the numerosity × condition interaction [$F(4,68) = 1.54, MS_e = 469.33, p = .200$] approached significance. Subitizing slopes for each condition were calculated for each individual participant (on the

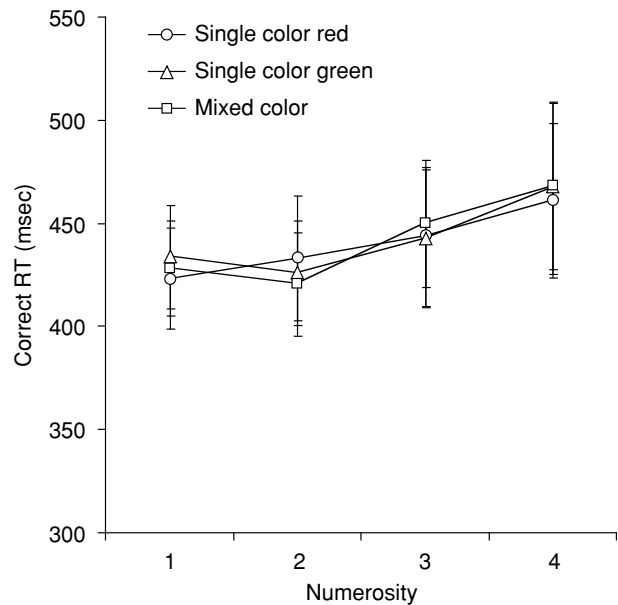


Figure 1. Mean correct median response times (RTs) as a function of condition and numerosity for Experiment 1. Error bars indicate ±1 standard error of the mean.

basis of numerosities 1–3). The overall means were 10.1, 4.5, and 10.9 msec/item for the single-color red, single-color green, and mixed-color displays, respectively, and these did not differ significantly [$F(2,34) = 1.50, MS_e = 145.76, p = .238$].⁵

Errors. Errors were rare overall (0.93%) and were not analyzed further.

Intertrial priming in single-color conditions. Median correct RTs were determined for trials in which the dot color was the same as or different from the dot color of the preceding trial, and mean correct median RTs are shown in Table 1. There was no evidence that RTs were shorter when stimulus colors were repeated across trials than when they changed across trials. This was confirmed by a 2 (repetition: same vs. different color trial) × 3 (numerosity: 1–3) within-subjects ANOVA, which showed that neither the main effect of repetition ($F < 1$) nor the repetition × numerosity interaction [$F(2,34) = 1.57, MS_e = 547.73, p = .223$] approached significance.

Discussion

The main finding from Experiment 1 was that, in contrast to Puts and de Weert (1997), enumerating items in mixed-color displays was no less efficient than enumerating items in single-color displays in terms of either overall RTs or subitizing rates. Overall RTs were within 1 msec of each other for enumerating items in single-color red, single-color green, and mixed-color displays. Likewise, the subitizing slope in the mixed-color condition was less than 1 msec different from the steeper single-color subitizing slope. A second finding was that there was no evidence of priming, since the subitization of stimuli that

Table 1
Mean Correct Median Response Times (RTs), and Standard Deviations,
as a Function of Whether the Stimulus Color Was Changed or
Repeated Across Two Consecutive Trials in Experiment 1

Stimulus Color	Numerosity								<i>M</i>
	1		2		3		4		
	RT	<i>SD</i>	RT	<i>SD</i>	RT	<i>SD</i>	RT	<i>SD</i>	
Nonrepeated	428.5	92.7	428.9	116.2	451.3	144.9	454.2	156.4	440.7
Repeated	432.6	115.5	424.0	114.8	435.9	139.6	465.0	178.2	439.3

were the same color as those on the preceding trial was no more efficient in terms of either slopes or overall RTs than the subitization of stimuli of a different color than those in the preceding trial. Thus, overall, our data suggest that subitizing stimuli represented across different color feature maps is no less efficient than subitizing stimuli within a single feature map.

However, one possible reason that we failed to obtain a single-color subitizing advantage is that, although we presented both red and green single-color displays, these were presented within a single block of trials. Thus, in our single-color condition, observers were not able to predict in advance whether the next trial would contain red or green stimuli. It might be that a single-color advantage arises only when observers can set themselves in advance to a particular color, perhaps by distributing limited-capacity attentional weights to bias processing of a specific color map (see, e.g., Wolfe et al., 2003). In addition, repeated and predictable single-color trials may lead to automatic and involuntary priming of target features (Maljkovic & Nakayama, 1994, 1996). In contrast, when the stimulus color of a single-color display is unpredictable, then (1) limited attentional weights may have to be shared between two color maps (as might be necessary with mixed-color displays), (2) weights may have to be dynamically redistributed from one feature map to the other when the current color weighting does not match the stimulus color, or (3) automatic intertrial feature priming would be reduced (see, e.g., Wolfe et al., 2003).

According to the accounts presented above, we might have predicted that there should be evidence of intertrial priming in the single-color condition when the stimulus colors repeated over trials (Maljkovic & Nakayama, 1994, 1996), which we did not find. However, note that following an observer's response, the display was masked by a field of both red and green dots. The rapid onset of this dual-color mask display might have acted to exogenously reset attentional weights across the available color maps and/or reduce/balance any color-specific feature-based priming.

In order to assess these possibilities, in Experiment 2 we presented the single-color conditions in separate blocks of trials and changed the masks so that they matched the color of the enumeration stimulus. Thus, in the red single-color block the masks were red, and in the green single-color block they were green. In this way, observers knew the color of the displays in the single-color conditions and so had maximum opportunity to set themselves to process the

relevant color, and the masks would no longer be able to exogenously reset observers to a competing color.

EXPERIMENT 2

Experiment 2 was essentially the same as Experiment 1 except that the single-color stimuli were presented in separate blocks of trials and the masks were the same color as the stimuli.

Method

Participants. Eighteen new participants (12 female) 19–32 years of age ($M = 22.8$ years) from the University of Warwick took part in the study for payment of £3. All had self-reported normal color vision and normal or corrected-to-normal visual acuity.

Stimuli and Apparatus. The general stimuli and apparatus were identical to those of Experiment 1.

Design and Procedure. The design was similar to that of Experiment 1 except that the single-color stimuli were presented in individual blocks and the masks were of a single color, which matched that of the stimuli. The blocks of single-color stimuli consisted of 72 trials, with each numerosity presented equally often (18 times) in a randomized order. Thus, as in Experiment 1, the total number of single-color and mixed-color trials was equal over the entire experiment. The mixed-color condition was identical to that presented in Experiment 1. Condition order was completely counterbalanced across participants, and a 24-trial practice block immediately preceded each full block.

Results

RTs and slopes. Six anticipatory RTs of <100 msec (<0.2%) were discarded. Median correct RTs (see note 2) were then calculated for each cell and each participant; the overall means of these median RTs are shown in Figure 2. RTs increased with numerosity at an overall rate of 8.2 msec/item [$F(1.17, 19.91) = 7.53, MS_e = 1,281.09, p = .010$]. However, neither the main effect of condition (overall means: red, 418.1 msec; green, 419.2 msec; mixed, 418.7 msec; $F < 1$) nor the condition \times numerosity interaction [$F(2.23, 37.93) = 1.92, MS_e = 601.50, p = .156$] approached significance.⁶

Subitizing slopes were calculated for each observer (on the basis of numerosities 1–3); overall means were 12.1, 1.6, and 11.0 msec/item for the single-color red, single-color green, and mixed-color displays, respectively. There was a marginally significant difference between these slopes [$F(2,34) = 2.87, MS_e = 207.01, p = .071$]. Planned comparisons revealed that the slope for the single-color green condition differed from that for the mixed-color condition [$t(17) = 2.27, p = .037$] and differed marginally

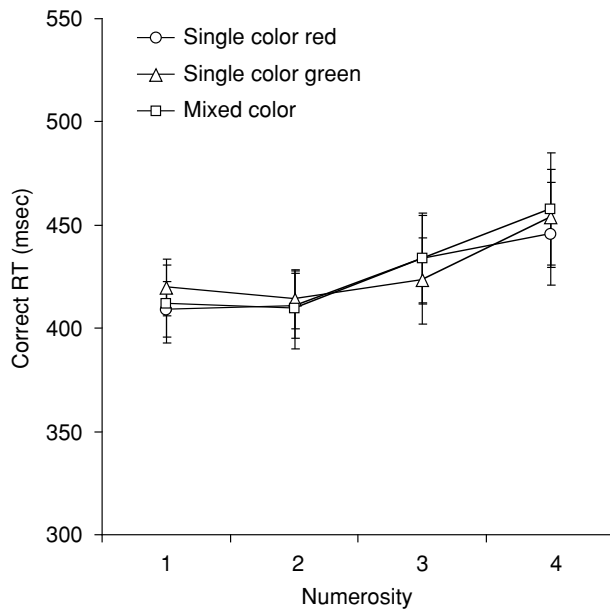


Figure 2. Mean correct median response times (RTs) as a function of condition and numerosity for Experiment 2. Error bars indicate ± 1 standard error of the mean.

ally from the single-color red condition [$t(17) = 1.83, p = .084$]. However, the difference between the single-color red and mixed-color condition slopes did not approach significance [$t(17) = 0.25, p = .802$].

Errors. Errors were rare (0.75% overall) and were not analyzed further.

Discussion

As in Experiment 1, overall RTs across conditions were within 1 msec of each other and did not differ. The subitizing rate in the single-color green condition was marginally higher than in the single-color red condition, and it was reliably higher than in the mixed-color condition. Crucially, however, the subitizing rate in the mixed-color condition was no less efficient than that in the single-color red condition. This underscores the importance of including both single-color conditions. If we had assessed only the single-color green condition, we would have concluded that mixed displays were less efficiently subitized than single-color displays. In fact, the mixed-color condition was subitized no less efficiently than the less efficient single-color condition (red). The difference between the single-color red and green conditions likely reflects subtle differences in the visual qualities of the red stimuli versus those of the green stimuli.

Experiment 2 therefore replicated our previous finding that mixed-color subitizing was no less efficient than single-color subitizing. Experiment 2 also showed that the results of Experiment 1 were not due to the inability of the observers to predict the stimulus color in the single-color conditions or to reduced intertrial priming (Maljkovic & Nakayama, 1994, 1996)⁷ when single-color conditions were presented in a single block.

GENERAL DISCUSSION

Puts and de Weert (1997) found that there was an overall RT cost when mixed-color displays are subitized in comparison with when single-color displays are, and they attributed this cost to the need to integrate information from two different color feature maps. In contrast, we did not find any cost to subitizing mixed-color in comparison with single-color displays in terms of either overall RTs or subitizing rates. One obvious issue is the extent to which our findings rely on a null effect. We do not believe that the equivalence of subitizing homogenous and heterogeneous color displays observed here is due simply to a lack of power or to noise in the data. First, our experiments included more participants than did those of Puts and de Weert, and they involved comparable numbers of trials. Second, there was sufficient power to detect relatively small effects, such as the overall 8-msec/item subitizing slope in Experiment 2. Third, there was no hint of a difference between the mixed-color and single-color conditions, with overall RTs within 1 msec of each other. Finally, the equivalence of subitizing single- and mixed-color displays was observed across two independent experiments.

Instead, we consider the contrasting results of our own experiments and those of Puts and de Weert (1997) to have been caused by the methodological differences detailed earlier. In brief, it is possible that Puts and de Weert found a difference between their single- and mixed-color conditions for one (or more) of at least three reasons. First, the mixed-color (red and green) condition was compared to only one of two possible single-color conditions (red but not green). As is shown in the present Experiment 2, any differences in the processing of either color alone could produce an apparent though unreal disadvantage for the mixed-color condition. Second, we examined a range of numerosities (1–3) that typically produces subitizing in most observers. In contrast, Puts and de Weert presented a range of numerosities (3–5) that likely encompassed both the subitizing and the counting ranges for most of their observers. Indeed, it can be noted that RTs in the present study were generally shorter than those observed by Puts and de Weert, which is consistent with their participants' using a more time-consuming enumeration method. Third, Puts and de Weert employed a brief presentation method and a 2AFC procedure, which, as was detailed earlier, might have encouraged participants to use a number of strategies to determine numerosity, such as estimating a global property of the display. It is likely that such strategies could have been employed more effectively in the single-color conditions, in which global display properties would be more reliable predictors of numerosity. In contrast, we used a larger range of response alternatives (one to four) and presented masked displays of unlimited duration.

There were also some differences between the stimuli across the two studies. Puts and de Weert's (1997) stimuli were approximately 0.6° , with a total configuration size of 2° , and they were presented on a white background. It is possible that these stimulus differences also contrib-

uted to the different findings. However, if anything, the lower contrast of their stimuli and the reduced display area might also have reduced the possibility that their displays were being subitized (Hunter & Sigler, 1940; Intriligator & Cavanagh, 1997; Trick & Pylyshyn, 1993).

Accounting for the Data

Our finding of equally efficient subitizing of mixed- and single-color displays is compatible with recent work assessing visual search for multiple targets that are defined within, in comparison with across, stimulus dimensions. Müller et al. (1995) showed that observers were able to detect the presence of any one of three possible single feature targets among a homogeneous set of distractors in a spatially parallel manner, regardless of whether the targets were defined within a dimension (the target could have any one of three different orientations) or across dimensions (the target was defined by size, orientation, or color). However, RTs were longer overall in the across-dimensions condition than in the within-dimension condition (see also Treisman, 1988). Müller et al. proposed that the longer RTs in the across-dimensions condition reflected time needed to redistribute attentional weights between different dimension feature saliency maps (Wolfe, 1994; Wolfe, Cave, & Franzel, 1989) in order to detect the target. In contrast, for within-dimension search, attentional weight changes were not needed in order for the target to be detected.

According to guided search theory (Wolfe, 1994; Wolfe et al., 1989), each feature dimension (color, orientation, size, etc.) possesses a saliency map, which is based on representations arising from the relevant simple primitive feature maps. Activity at each location within a saliency map depends on how closely the features of the item at that location match a target (top-down component) and how different that item is from neighboring items within the same dimension (bottom-up component). Each feature dimension thus produces an individual saliency map. The activity from each specific saliency map then feeds to a master location map, and search proceeds via examination of master map locations in order of activation value. According to the dimensional weighting account (Müller et al., 1995), detecting a target requires a sufficiently large activation at a location within the master location map. For within-dimension target searches, the weight between the relevant dimension saliency map and the master map can be set to a maximum value. This results in rapid transfer of activity from that saliency map to the master map and thus rapid target detection. In contrast, for between-dimension target searches, weights have to be dynamically reassigned (or shared) across several saliency maps, thereby slowing the transfer of activity to the master map and producing an RT cost.

From this account, search responses are made on the basis of activity that represents differences in saliency of items within a display rather than activity within a particular feature map. However, even though activity within a saliency map may indicate that a target is present, it does not directly indicate the dimension or feature values

of the target. Indeed, other evidence (Müller et al., 1995) suggests that the visual system does not have direct access to specific feature maps (e.g., the map coding a particular orientation). For example, when targets were defined by the presence of either of two particular feature values but a third was to be treated as a nontarget (i.e., it was to elicit an absent response), there was a substantial RT cost in the within-dimension condition in comparison with when all three feature values were to be treated as targets. In this situation, master map activity alone cannot be used to generate a response because it is derived from a dimension-specific saliency signal that does not provide information about the identity of the feature difference. Instead, determining the identity of the feature difference appears to require additional processes, such as a serial check through each dimension-specific feature map (e.g., each color map; Müller et al., 1995). Thus, it appears that weights cannot be easily shifted in advance to particular feature values within a dimension.⁸

Watson, Maylor, and Bruce (2005b) provided converging evidence that the visual system does not have direct access to specific feature maps. In a series of experiments, they required observers to enumerate colors (rather than items) present in a display. When the number of colors did not correlate with the number of items, enumeration was particularly slow and there was little evidence of subitization. This is consistent with the visual system's inability to monitor activity within specific feature maps within a particular stimulus dimension. In other words, it has no direct parallel access to individual feature maps within a dimension.

The present data mesh with the previous findings. In particular, if enumeration proceeds on the basis of activation within a master location map (Found & Müller, 1996), which receives information via dimension saliency maps, we would expect little difference between enumerating items in mixed-color displays in comparison with single-color displays. (This prediction does not hold when items differ across dimensions; see Found & Müller, 1996, for relevant subitizing data.) If anything, mixed-color displays might be processed more quickly because a mixed display might produce a larger bottom-up component within the saliency map, since, on average, neighboring items would be more different from each other.

However, recall that Trick and Pylyshyn (1993) found that items of one color could be selectively processed and subitized among distractors of a different color. If there is no direct access to specific feature maps (in other words, if FINSTs cannot be applied to active locations within a specific color map), how is color-selective subitizing achieved? In addressing a similar question, Watson et al. (2005b) proposed that color-selective subitization could be achieved by the top-down inhibition of distractor feature maps (Treisman & Sato, 1990; Watson & Humphreys, 1998) or activation of locations coding target features (as in guided search; Wolfe, 1994; Wolfe et al., 1989). This would effectively reduce the activation of the distractors or enhance the activation of the targets in the

relevant dimension saliency map and consequently produce a selective activation of target items in the master location map to which FINSTs were applied. These additional processes would require time and likely reduce the rate of accumulation of target-specific activity within the master map (similar to the extra time required to determine the identity of a simple feature difference; Müller et al., 1995). Consistent with this possibility, Watson et al. (2002) found that observers could subitize single-feature, shape-defined targets (that were detected efficiently in simple visual search) but that there was an overall cost in RT (see Randall, 1998; Trick & Pylyshyn, 1993, for similar findings with color differences).

Conclusion

Our findings suggest that subitization is based on representations that, by default, show where objects are but not what they are (i.e., a location-based master map). Such a map is initially generated on the basis of signals from dimension-specific saliency maps that do not provide information about within-dimension feature content. As such, subitization (and therefore early visual tagging; Pylyshyn, 2000) can be rapid, efficient, and unaffected by item heterogeneity (at least for within-dimension heterogeneity). When selective enumeration of targets possessing a particular feature is required, activity from a specific target feature map within a dimension must be enhanced (or activity from distractors must be inhibited), and this requires additional time-consuming processes (as in equivalent visual search tasks; Müller et al., 1995), leading to an overall RT performance cost.

REFERENCES

- ATKINSON, J., CAMPBELL, F. W., & FRANCIS, M. R. (1976). The magic number 4 ± 0 : A new look at visual numerosity judgements. *Perception*, **5**, 327-334.
- FOUND, A., & MÜLLER, H. J. (1996). Searching for unknown feature targets on more than one dimension: Investigating a "dimension-weighting" account. *Perception & Psychophysics*, **58**, 88-101.
- FRICK, R. W. (1987). The homogeneity effect in counting. *Perception & Psychophysics*, **41**, 8-16.
- HUNTER, W. S., & SIGLER, M. (1940). The span of visual discrimination as a function of time and intensity of stimulation. *Journal of Experimental Psychology*, **26**, 160-179.
- INTRILIGATOR, J., & CAVANAGH, P. (1997). Individuation and the resolution of attention [Abstract]. *Investigative Ophthalmology & Visual Science*, **38**, 3209.
- KAUFMAN, E. L., LORD, M. W., REESE, T. W., & VOLKMAN, J. (1949). The discrimination of visual number. *American Journal of Psychology*, **62**, 498-525.
- MALJKOVIC, V., & NAKAYAMA, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition*, **22**, 657-672.
- MALJKOVIC, V., & NAKAYAMA, K. (1996). Priming of pop-out: II. The role of position. *Perception & Psychophysics*, **58**, 977-991.
- MANDLER, G., & SHEBO, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, **111**, 1-22.
- MÜLLER, H. J., HELLER, D., & ZIEGLER, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics*, **57**, 1-17.
- PUTS, M. J. H., & DE WEERT, C. M. M. (1997). Does color influence subitization? *Acta Psychologica*, **97**, 71-78.
- PYLYSHYN, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, **32**, 65-97.
- PYLYSHYN, Z. W. (2000). Situating vision in the world. *Trends in Cognitive Sciences*, **4**, 197-207.
- PYLYSHYN, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, **80**, 127-158.
- PYLYSHYN, Z. W., & STORM, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, **3**, 179-197.
- RANDALL, G. (1998). *The temporal binding of visual features and their role in enumeration*. Unpublished doctoral thesis, University of Birmingham, U.K.
- RASTLE, K., & DAVIS, M. H. (2002). On the complexities of measuring naming. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 307-314.
- SVENSON, O., & SJÖBERG, K. (1983). Speeds of subitizing and counting processes in different age groups. *Journal of Genetic Psychology*, **142**, 203-211.
- TREISMAN, A. [M.] (1988). Features and objects: The Fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, **40A**, 201-237.
- TREISMAN, A. M., & GELADE, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97-136.
- TREISMAN, A. [M.], & SATO, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 459-478.
- TRICK, L. M., & ENNS, J. T. (1997). Clusters precede shapes in perceptual organization. *Psychological Science*, **8**, 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 & 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**, 80-102.
- WATSON, D. G., & HUMPHREYS, G. W. (1998). Visual marking of moving objects: A role for top-down feature-based inhibition in selection. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 946-962.
- WATSON, D. G., & HUMPHREYS, G. W. (1999). The magic number four and temporo-parietal damage: Neurological impairments in counting targets amongst distractors. *Cognitive Neuropsychology*, **16**, 609-629.
- WATSON, D. G., MAYLOR, E. A., & BRUCE, L. A. M. (2005a). Effects of age on searching for and enumerating targets that cannot be detected efficiently. *Quarterly Journal of Experimental Psychology*, **58A**, 1119-1142.
- WATSON, D. G., MAYLOR, E. A., & BRUCE, L. A. M. (2005b). The efficiency of feature-based subitization and counting. *Journal of Experimental Psychology: Human Perception & Performance*, **31**, 1449-1462.
- WATSON, D. G., MAYLOR, E. A., & MANSON, N. J. (2002). Aging and enumeration: A selective deficit for the subitization of targets among distractors. *Psychology & Aging*, **17**, 496-504.
- WOLFE, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, **1**, 202-238.
- WOLFE, J. M., BUTCHER, S. J., LEE, C., & HYLE, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 483-502.
- WOLFE, J. M., CAVE, K. R., & FRANZEL, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 419-433.

NOTES

1. Color measurements were obtained with a Minolta CS-1000 spectroradiometer directed toward a solid circular patch of color that completely filled the receptive field of the device from a distance of 60 cm.
2. Analyzing mean RTs (as Puts & de Weert, 1997, did) rather than medians produced essentially the same set of findings.

3. The Greenhouse–Geisser correction was applied when the sphericity assumption was violated.

4. Although subitizing is sometimes described as a spatially parallel process, there is often a reliable RT–numerosity slope. According to Trick and Pylyshyn (1994), the slope arises as a result of assigning number names to each bound FINST even though FINST assignment itself may be a spatially parallel process.

5. We thank Lana Trick for pointing out that at numerosity 1 the displays were always homogeneous and, thus, calculating slopes on the basis of one to three items confounds display heterogeneity with numerosity. However, note that this would tend to exaggerate any slope differences between the homogeneous and heterogeneous conditions and hence provides a more conservative test of the effects of color heterogeneity. Nonetheless, we reanalyzed the data excluding numerosity 1 and, although this strengthened the main effect of numerosity [$F(1,17) = 6.98$, $MS_e = 1,389.68$, $p < .05$], neither the main effect of condition ($F < 1$) nor the condition \times numerosity interaction [$F(2,34) = 1.38$, $MS_e = 606.35$, $p = .266$] approached significance.

6. Excluding numerosity 1 produced the same results: There was a reliable effect of numerosity [$F(1,17) = 13.09$, $MS_e = 725.31$, $p < .005$], but neither the main effect of condition ($F < 1$) nor the numerosity \times condition interaction [$F(2,34) = 1.45$, $MS_e = 432.22$, $p = .249$] approached significance.

7. We note incidentally that perhaps intertrial priming of target features is effective only when targets must be detected or selected among competing distractors. In contrast, in our displays all items had to be processed, irrespective of feature values.

8. Alternatively, Wolfe et al. (2003) proposed that multiple weights for feature values within a dimension might be maintained with little cost, provided that neither stimulus value ever plays the role of a distractor. However, by either account, one would predict little cost to subitizing mixed-feature within-dimension defined targets.

(Manuscript received July 16, 2004;
revision accepted for publication May 26, 2005.)