Numerosity underestimation with item similarity in dynamic visual display

Ricky K. C. Au
Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan
Japan Society for the Promotion of Science, Tokyo, Japan

Katsumi Watanabe
Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan

The estimation of numerosity of a large number of objects in a static visual display is possible even at short durations. Such coarse approximations of numerosity are distinct from *subitizing*, in which the number of objects can be reported with high precision when a small number of objects are presented simultaneously. The present study examined numerosity estimation of visual objects in dynamic displays and the effect of object similarity on numerosity estimation. In the basic paradigm (Experiment 1), two streams of dots were presented and observers were asked to indicate which of the two streams contained more dots. Streams consisting of dots that were identical in color were judged as containing fewer dots than streams where the dots were different colors. This underestimation effect for identical visual items disappeared when the presentation rate was slower (Experiment 1) or the visual display was static (Experiment 2). In Experiments 3 and 4, in addition to the numerosity judgment task, observers performed an attention-demanding task at fixation. Task difficulty influenced observers' precision in the numerosity judgment task, but the underestimation effect remained evident irrespective of task difficulty. These results suggest that identical or similar visual objects presented in succession might induce substitution among themselves, leading to an illusion that there are few items overall and that exploiting attentional resources does not eliminate the underestimation effect.

Introduction

When viewing a complex scene that contains a large number of objects, humans have no difficulty approximating the number of the objects (i.e., numerosity). The visual system is so proficient at numerosity estimation that it is as if it can sense number directly (Ross & Burr, 2010). Similar to other sensory systems, the sense of numerosity is subject to the influence of adaptation (Burr & Ross, 2008). We can judge numerosity in various ways, such as one-by-one serial counting of elements or rough estimation based on instantaneous impressions. We can also discriminate numerosity between two scenes. Our ability to discriminate numerosities starts to develop in the early stages of life, and emerges even earlier than language or symbolic counting (Lipton & Spelke, 2003). There is also evidence that the ability to judge numerosity is not limited to vision but also exists in audition (Lipton & Spelke, 2003, 2004) and touch (Plaisier, Tiest, & Kappers, 2010), and across modalities perhaps relying on a unitary amodal system (Gallace, Tan, & Spence, 2007).

It has been suggested that the visual system possesses two separate systems for processing numerosities of different ranges (small vs. large number of objects) without counting. People are quite accurate at rapidly judging the exact number of objects up to about four. This rapid and accurate numerosity judgment for a range of small numbers without counting is referred to as *subitizing* (Dehaene, 1992; Kaufman, Lord, Reese, & Volkmann, 1949). The other system deals with larger numbers of items and involves approximate estimation—the rapid coarse approximation of numerosity, which is less accurate and precise than subitizing. Accumulating evidence suggests that there are two separate processing systems (Ansari, Lyons, van Eimeren, & Xu, 2007; Hyde & Spelke, 2011; Palomares & Egeth, 2010; Revkin, Piazza, Izard, Cohen, & Dehaene, 2008). However, a few studies have also shown that adaption and manipulation of attentional load affect numerosity judgments in both high (estimation) and low (subitizing) ranges of numerosities,

Citation: Au, R. K. C., & Watanabe, K. (2013). Numerosity underestimation with item similarity in dynamic visual display. *Journal of Vision*, *13*(8):5, 1–15, http://www.journalofvision.org/content/13/8/5, doi:10.1167/13.8.5.

doi: 10.1167/13.8.5 Received March 26, 2013; published July 2, 2013 ISSN 1534-7362

suggesting that numerosity may be processed by a single system (Burr, Anobile, & Turi, 2011; Vetter, Butterworth, & Bahrami, 2008). Whether numerosity judgments depend on a single process or multiple processes is still under debate.

Numerosity estimation has been found to depend on the configuration of the scene and the items' visual features. The spatial arrangement of objects can interact with numerosity estimation (Allik, Helsper, & Vos, 1991; Ginsburg, 1991; Sophian & Chu, 2008). Some studies have suggested that a higher density (i.e., a smaller array area) of elements leads to underestimation of numerosity (Hollingsworth, Simmons, Coates, & Cross, 1991; Krueger, 1972; Shuman & Spelke, 2006; Vos, van Oeffelen, Tibosch, & Allik, 1988), while others have suggested that density does not matter (Burgess & Barlow, 1983). Density of texture does not appear to mediate numerosity judgments, while decreasing luminance could lead to increased perceived numerosity (Ross & Burr, 2010). In response to Ross and Burr (2010), Raphael, Dillenburger, and Morgan (2013) argued that numerosity in textures of circular dots can be judged based on density or size (i.e., area) of the overall field of dots when they are correlated with the total number of elements. In this case, the visual system may infer the total number of elements based on estimates of the size and density of the field. Since observers judge numerosity more accurately when it is correlated with changes in size than changes in density, Raphael et al. (2013) have suggested that multiple mechanisms are involved in judging numerosity (at least one based on size and one based on density). They further argued that if there is a direct mechanism that determines numerosity, it must be noisier than mechanisms that respond to changes in size or density.

In terms of visual features, people tend to overestimate numerosity in dot arrays when the dots have a smaller diameter, aggregate surface or density but a larger convex hull (Gebuis & Reynvoet, 2012). Previous findings also suggested that arrays with larger items tend to be perceived as less numerous (Ginsburg & Nicholls, 1988; Shuman & Spelke, 2006), although others suggested that perceived numerosity is size invariant (Allik, Tuulmets, & Vos, 1991). When asked to judge the number of objects that have a particular feature, people tend to overestimate when items with the target feature are clustered together (Goldstone, 1993). Goldstone (1993) explained that this might be due to the fact that visual features become more salient if they are highly concentrated in a given region. People also underestimate numerosity when objects in the field are connected by task-irrelevant lines (He, Zhang, Zhou, & Chen, 2009). Tokita and Ishiguchi (2010) demonstrated that practice might be a critical factor that mediates the effects of perceptual variables, including array size and element size, on numerosity judgments.

In the current literature, many studies have examined numerosity estimation in vision with one-shot static displays that require the observer to give an estimate or to compare numerosities between briefly presented stimuli. However, despite the fact that numerosity judgments in the real life often occur in dynamic visual contexts, very few investigations of numerosity judgments with dynamic displays have been conducted. Afraz, Kiani, Vaziri-Pashkam, and Esteky (2004) reported motion-induced overestimation of numerosity when items moved along a circular path on the screen. Observers exhibited a bias to overestimate the number of items when rotation speed increased. When the items were marked with different colors, this motion-induced overestimation effect was still observed, even though observers did not perceive a greater number of colors. Afraz et al. (2004) concluded that this effect was dependent on the product of rotation speed and the number of items shown. By asking observers to indicate which of two streams of dynamic dot displays appeared more numerous, Allik and Tuulmets (1993) found that perceived numerosity decreased when spatial and temporal proximity between the presented items increased. This finding demonstrates how temporal properties of dynamic visual events can interact with spatial properties to influence the perception of overall numerosity.

To add to the currently limited literature on numerosity judgments in dynamic displays, we examined whether the homogeneity of visual elements in dynamic displays affects numerosity judgments. We hypothesized that, in addition to spatial and temporal proximity (Allik & Tuulmets, 1993), proximity in feature space (i.e., similarity) would affect numerosity judgments. Our results supported this hypothesis and demonstrated a new phenomenon—numerosity under estimation[1] in dynamic displays with a large number of objects (far beyond the subitizing range) when a visual feature (color) is identical among items, compared to when items differ in color. Further, we demonstrate that the underestimation effect due to color homogeneity only occurs with fast presentation rate but neither with slow presentation rate (Experiment 1) nor in static displays (Experiment 2).

Previous studies also showed that subitizing can be affected by the availability of attentional resources (Burr, Turi, & Anobile, 2010; Egeth, Leonard, & Palomares, 2008; Olivers & Watson, 2008; Railo, Koivisto, Revonsuo, & Hannula, 2008). In Experiments 3 and 4, therefore, we manipulated attentional load by requiring observers to perform another task simultaneously with the numerosity judgment task and examined whether the availability of attentional resources affected the underestimation effect. Finally, we discuss possible explanations for the numerosity underestimation effect, and suggest that object substitution (Enns & Di Lollo, 1997) might occur among items that are close on both spatiotemporal and featural dimensions, causing the apparent perception of fewer objects.

Object substitution occurs when new incoming information is fed into the visual input rapidly before the system has fully processed previous information. As described in the three-layer computational model of object substitution by Di Lollo, Enns, and Rensink (2000; see also Enns & Di Lollo, 2000), a new visual event activates the input, working, and pattern layers in the system; reentrant connections (Bullier, McCourt, & Henry, 1988; Felleman & Van Essen, 1991) allow feedback communications from higher extrastriate areas (of the pattern layer) with the primary visual areas involved in the input and working space layers. Information stored in the pattern layer is constantly copied to the working space, and the system iteratively compares the information in the working space with the information in the input layer. When there is an information match, the signal in the pattern layer is strengthened by the weighted output from the input and working space layers, giving rise to a stable conscious perception. However, if the information in the input layer is changed rapidly, such that the pattern layer is still processing previous information, there will be a mismatch between the input and working space. Such a mismatch between reentrant activity and the ongoing input causes the signal associated with the previous information (for which processing is unfinished) to fade out and be replaced with the new, incoming information. Whether the old information can ultimately reach conscious perception depends on the number of iterations required to complete processing before the information has totally decayed. Therefore, if new incoming information appears rapidly enough, the old information will fade out without being consciously perceived (and possibly integrated into the representation of the object that is currently being processed). Such substitution effect of object commonly occurs when the processing system is required to perform rapid updating of object information on fast-changing visual input (e.g., Moore & Enns, 2004).

Experiment 1

The aim of Experiment 1 was to examine whether the homogeneity of visual elements in a dynamic display would influence numerosity judgment. Two streams of dynamic displays were presented in succession; one stream contained visual elements in a single color and the other contained visual elements in two different colors.

Methods

Observers

Twelve naïve participants were recruited to participate in the experiment. All observers had normal or corrected-to-normal visual acuity and gave informed consent prior to the experiment.

Stimuli

Stimuli were programmed in MATLAB R2012b (MathWorks, Natick, MA) with the Psychophysics Toolbox extension (version 3.0.10; Brainard, 1997; Pelli, 1997) and were displayed on a CRT monitor with a refresh rate of 100 Hz and resolution of 800 $\times$ 600 pixels. Stimulus presentation and response collection were controlled by a personal computer running the Windows 7 operating system. Observers viewed the display at a distance of 60 cm from the monitor and performed the experiment in a dimly illuminated and quiet room.

All stimuli were presented on a black background. The fixation stimulus was a cross composed of vertical and horizontal white lines (length = 0.317°, width = 0.045°), presented at the center of the screen. Dot stimuli were either red or green (controlled to be the same luminance, 0.47 cd/m^2) with a diameter of 0.273°. They appeared at one of 72 possible positions evenly distributed on an imaginary circle with a radius of 4.533°.

Procedure

Each trial began with a blank screen. The observer initiated each trial by pressing the space bar on the keyboard. The fixation then appeared and remained on the screen throughout the trial until a response was made. After 500 ms, two streams of dynamic visual stimuli appeared on the screen, separated by a blank screen of 500 ms (Figure 1). The two streams were always presented for 960 ms each. One of the streams consisted of dots that were all the same color (either all red or all green; same-color stream) while the other stream consisted of dots of both colors (different-color stream).

The experiment was composed of two blocks: one with a fast presentation rate of 40 ms per frame and the other with a slow rate of 240 ms per frame. Because the duration of each stream was fixed at 960 ms, a total of 24 frames were presented in the fast-rate (40 ms/frame) block, and four frames were presented in the slow-rate (240 ms/frame) block. In the different-color stream,

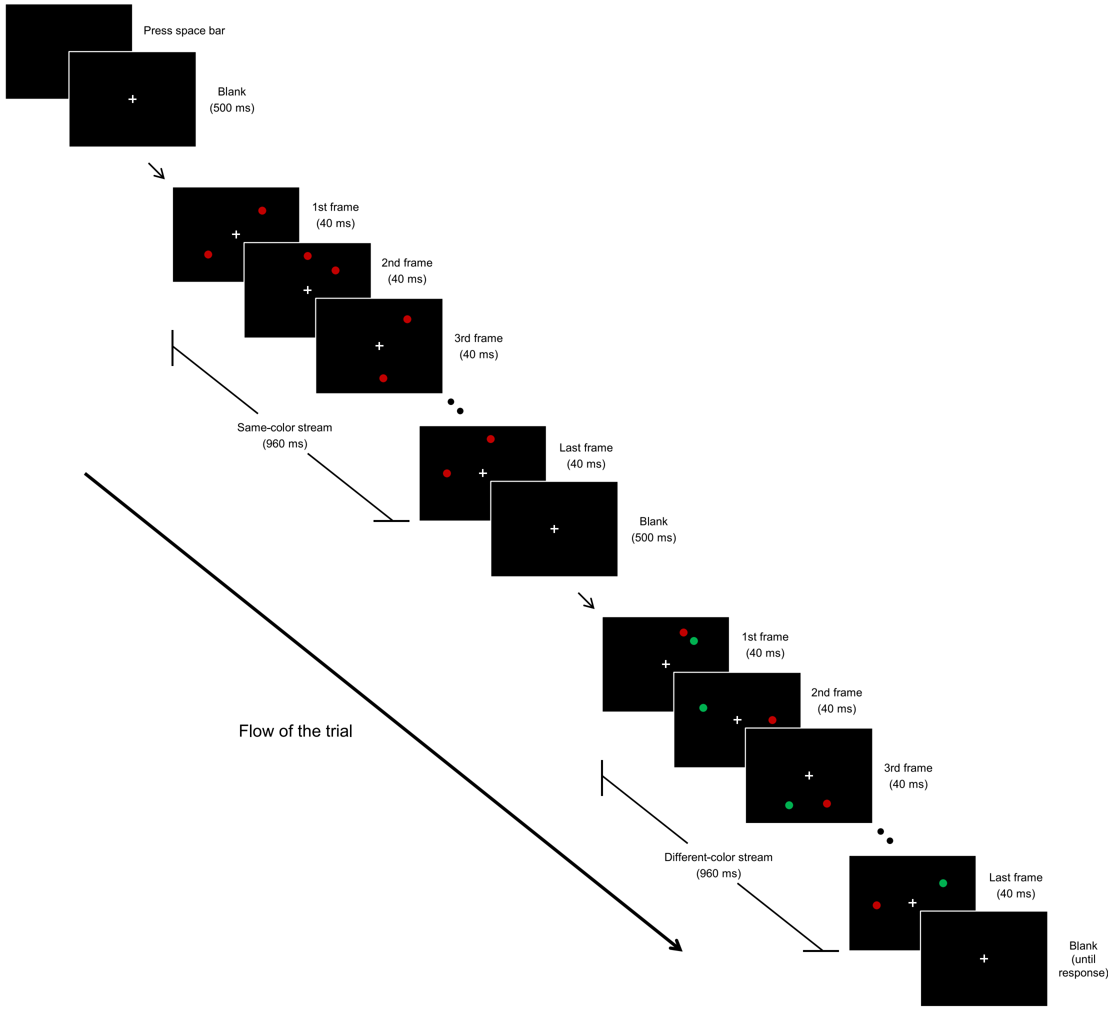

Figure 1. The flow of a trial in Experiment 1 (dynamic display session) at the frame duration of 40 ms.

two dots (one red, one green) were always presented in each frame at two randomly determined positions among the 72 possible positions, with the constraint that the two randomly selected positions in each frame had to be different from the two selected in the previous frame. Thus, for different-color streams, a total of 48 dots (2 dots $\times$ 24 frames) were presented in the fast-rate block, and 8 dots (2 dots $\times$ 4 frames) were presented in the slow-rate block.

In the same-color stream, the number of dots in each frame varied: one, two, or three dots were presented in each frame. Similar to the different-color streams, the positions of the dots were constrained such that the randomly selected positions differed from those in the previous frame. In each same-color stream, the percentage of frames that displayed one or three dots varied across a range of nine conditions: four conditions with 25%, 50%, 75%, or 100% of the frames showing one dot; four conditions with 25%, 50%, 75%, or 100% of the frames showing three dots; and the 0% condition, in which two dots were presented in all frames. For example, in the 25%-one dot condition in the fast-rate block, 25% of the frames (which equals 24 frames $\times$ 25% $=$ 6 frames) in the stream contained one dot, and 75% of the frames (18 frames) contained two dots. The two frame types were presented in a random sequence within the stream. Consequently, there were nine different number of dots conditions showing a total of 24, 30, 36, 42, 48, 54, 60, 66, or 72 same-color dots in streams in the fast-rate block, and 4, 5, 6, 7, 8, 9, 10, 11, or 12 same-color dots in streams in the slow-rate block (both are represented in percentage of numerosity relative to the standard, which are 48 and 8, respectively, in Figure 2 in the Results section).

Observers were instructed to indicate which of the two streams contained more dots by pressing one of two designated keys on the keyboard, after the second stream disappeared. For each block, observers completed a total of 180 trials (2 presentation orders $\times$ 9 percentages of frames or total number of dots in the same-color stream $\times$ 2 colors of the same-color stream $\times$ 5 repetitions). Observers completed the two blocks in a counterbalanced order separated by a 5-min break. The whole experiment lasted about 30 min.

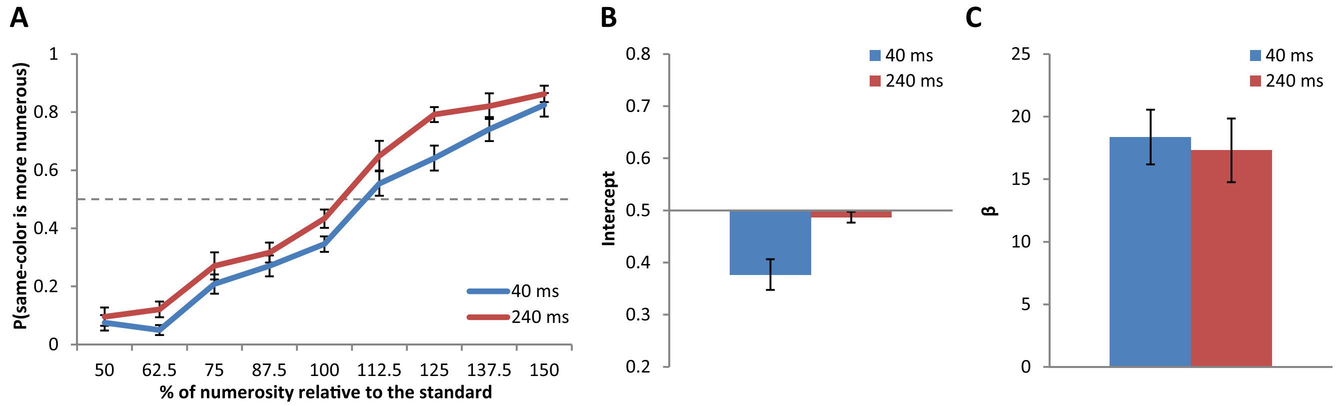

Figure 2. The proportion of trials the observers responded that the same-color stream contained a greater number of dots than the different-color stream in Experiment 1 (Panel A); average estimated proportion value (i.e., intercept) at $x = 100\%$ (Panel B); average β of the fitted logistic function (Panel C); error bars showing the standard error of the mean.

Results and discussion

The proportion of trials in which the observer indicated that the same-color stream contained more dots than the different-color stream was computed for each "percentage of numerosity relative to the standard" condition. The proportions averaged across the twelve observers for the fast-rate (40 ms/frame) and slow-rate (240 ms/frame) blocks are shown in Figure 2A.

An omnibus $2 \times 2 \times 9$ (Order $\times$ Presentation Rate $\times$ Numerosity) analysis of variance (ANOVA) was performed to examine whether there was any difference in proportion of response across the conditions. The main effects of Presentation Rate, $F(1, 11) = 13.856$, $p = 0.003$, and Numerosity, $F(8, 88) = 108.293$, $p < 0.001$, the Order $\times$ Presentation Rate interaction, $F(1, 11) = 10.770$, $p = 0.007$, and the Order $\times$ Numerosity interaction, $F(8, 88) = 5.247$, $p < 0.001$, were found to be statistically significant; while the Order main effect, $F(1, 11) = 0.442$, $p = 0.520$, the Presentation Rate $\times$ Numerosity interaction, $F(8, 88) = 0.953$, $p = 0.478$, and the Order $\times$ Presentation Rate $\times$ Numerosity interaction, $F(8, 88) = 0.801$, $p = 0.604$, were not. Simple main effect analyses showed that the Order main effect was only marginally significant in the slow-rate condition ($p = 0.057$) and not significant in the fast-rate condition ($p = 0.428$). Therefore, the order of presentation might be of some relevance in explaining the apparent bias in response (a proportion lower than the value of 0.5 in the 100% condition in Figure 2A) observed in the slow-rate condition, but the underestimation effect observed in the fast-rate condition is not likely to be explained by the bias in response.

For each observer, separate logistic functions were fitted to the data in the fast-rate and slow-rate blocks respectively, using the Bootstrap Inference function provided in the Psignifit toolbox for MATLAB version 3.0 (see http://psignifit.sourceforge.net/; Fründ, Haenel, & Wichmann, 2011; Wichmann & Hill, 2001). The α (the threshold obtained after adjustment of the lower and the upper bound), β (a parameter related to the slope of the fitted function, representing the variance), γ (the miss rate, governing the lower bound of the fitted function), λ (the lapse rate, governing the upper bound of the fitted function), and the intercept (the estimated proportion at the condition of 100%) of the fitted functions were determined (Figure 2B, 2C). The parameters γ and λ were allowed to vary within the range of 0 to 0.25 in the fitting procedure. The slope of a psychometric function represents the precision in making response to a particular stimulus level: When perfect judgments are made, the slope would be infinity (β would tend to zero) in the midway of the stimulus range; when judgments become imprecise, variations in response would emerge and the function would become S shape with increasingly shallow slope at all stimulus levels (i.e., a function with increasingly large β). An intercept with a value smaller than 0.5 would represent the case that the observer chooses the same-color stimulus to be more numerous than the different-color stimulus for a less proportion of trials, and thus would indicate an underestimation of numerosity of the same-color stimulus relative to the different-color stimulus in the particular condition. One-sample t tests showed that the intercept in the fast-rate condition was significantly smaller than 0.5, $t(11) = -4.203$, $p = 0.001$, while the intercept in the slow-rate condition was not, $t(11) = -1.234$, $p = 0.243$. A paired t test showed that the intercept in the fast-rate condition was significantly lower than the intercept in the slow-rate condition, $t(11) = -3.727$, $p = 0.003$. Thus, observers perceived fewer dots in the same-color than in the different-color streams when presentation rate was fast; however, this phenomenon was not evident for slower presentation rates. This implies that a fast presentation rate of visual information, which challenges the spatiotemporal resolution of the visual system, is a critical factor for

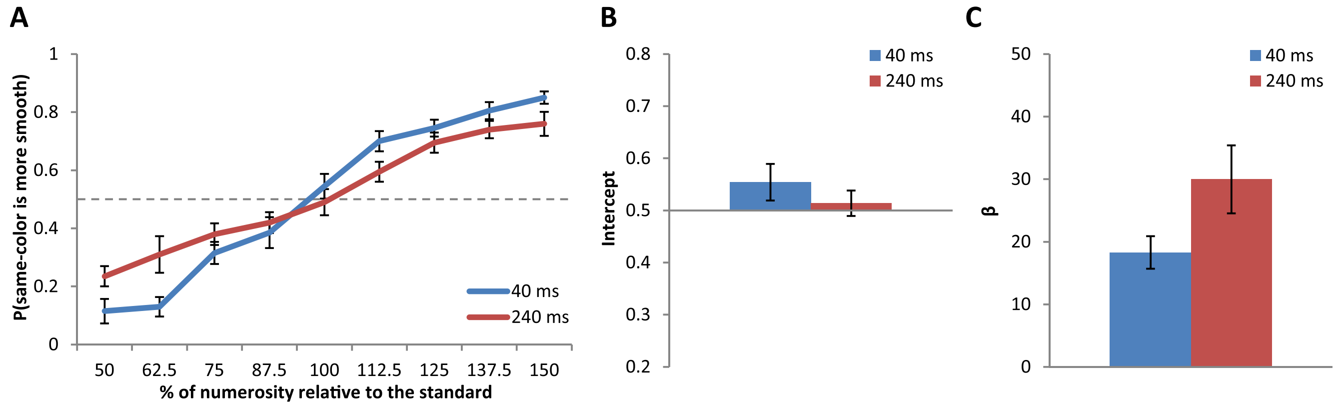

Figure 3. The proportion of trials the observers responded that the same-color stream showed more smooth motion than the different-color stream in the first control experiment (Panel A); average estimated proportion value (i.e., intercept) at $x = 100\%$ (Panel B); average β of the fitted logistic function (Panel C); error bars showing the standard error of the mean.

eliciting the underestimation effect. For the β of the fitting functions, the two conditions did not show any significant difference, $t(11) = 0.620$, $p = 0.548$.

Flickering dots can induce illusory motion in some cases. In a control experiment using the identical set up as in Experiment 1, 10 observers were asked to judge which of the two stimulus streams showed more smooth motion. Results revealed that the observers did not perceive the same-color stream and the different-color stream differently in terms of motion smoothness, for both the fast-rate, $t(9) = 1.550$, $p = 0.156$, and the slow-rate, $t(9) = 0.563$, $p = 0.587$, conditions (testing the intercept of the curve fitted with logistic function; Figure 3). In another control experiment, also using the identical setup as in Experiment 1, 10 observers were asked to judge which of the stimulus streams flickered at a higher frequency (although the two movie streams were always at the same flickering rate). Observers did not perceive different flickering frequencies in the same-color and the different-color streams for both the fast-rate, $t(9) = -1.385$, $p = 0.199$, and the slow-rate, $t(9) = -1.340$, $p = 0.213$, conditions (Figure 4). Thus, the control experiments showed that the underestimation effect in the same-color dot streams relative to different-color dot streams was not likely to be due to differently perceived motion smoothness or flickering rate.

Experiment 2

The aim of Experiment 2 was to investigate whether the numerosity underestimation effect observed in Experiment 1 only occurs in dynamic displays or whether it generalizes to static displays as well. Observers performed a numerosity judgment task similar to Experiment 1, but the dot stimuli were presented on the screen only once for variable durations.

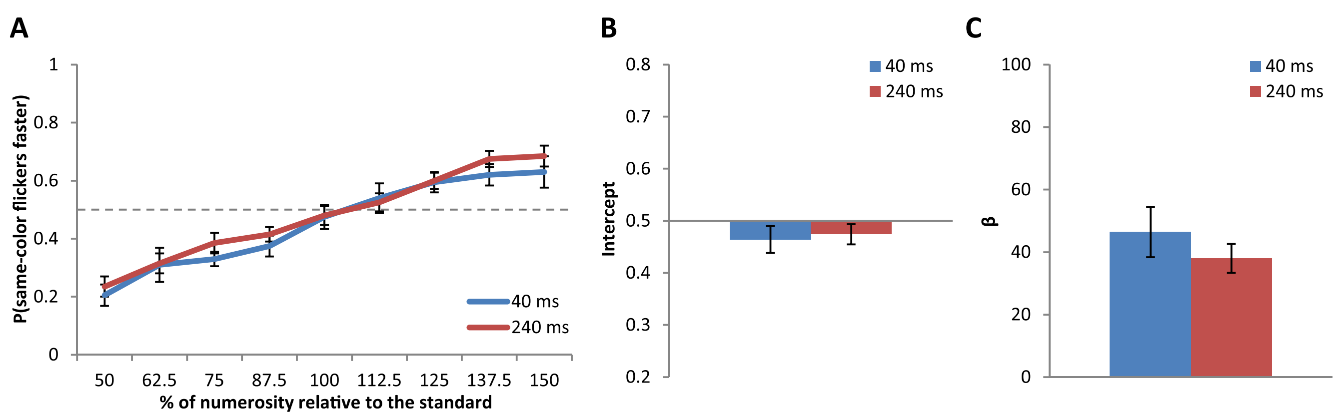

Figure 4. The proportion of trials the observers responded that the same-color stream flickered at a higher frequency than the different-color stream in the second control experiment (Panel A); average estimated proportion value (i.e., intercept) at $x = 100\%$ (Panel B); average β of the fitted logistic function (Panel C); error bars showing the standard error of the mean.

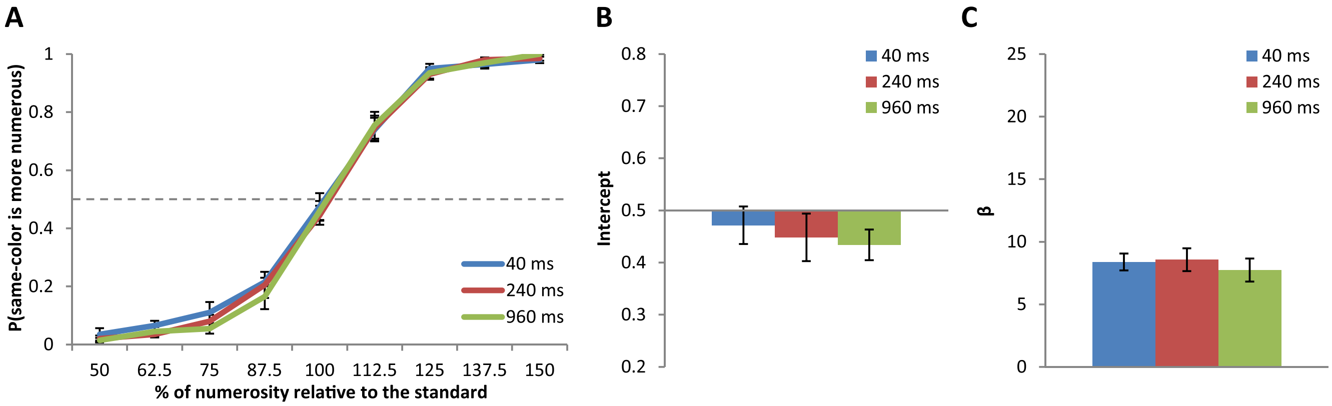

Figure 5. The proportion of trials the observers responded that the same-color stimulus contained a greater number of dots than the different-color stimulus in Experiment 2 (Panel A); average estimated proportion value (i.e., intercept) at $x = 100\%$ (Panel B); average β of the fitted logistic function (Panel C); error bars showing the standard error of the mean.

Methods

Observers

Ten new naïve observers were recruited to participate in the experiment. All had normal or corrected-to-normal vision and gave informed consent before the experiment.

Stimuli and procedure

The experimental flow and stimuli were generally the same as Experiment 1, but instead of presenting two 960-ms streams of dynamic stimuli, two static frames were presented sequentially for short (40 ms), medium (240 ms), or long (960 ms) durations, separated by a 500-ms blank interval. In the different-color stimulus, a total of 48 dots (half red and half green) were presented in random positions selected from the 72 possible positions. In the same-color stimulus, a total of 24, 30, 36, 42, 48, 54, 60, 66, or 72 dots (either all red or all green) were presented in random positions. There were three different blocks in the experiment for each of the three different presentation durations. Observers performed 180 trials in each of the three blocks (order counterbalanced among observers) separated by 5-min breaks.

Results and discussion

As in the dynamic display session (Experiment 1), the proportion of trials where observers responded that the same-color stimulus contained more dots than the different-color stimulus was computed separately for the short (40 ms), medium (240 ms), and long (960 ms) duration blocks (Figure 5A).

Following the analysis in Experiment 1, an omnibus $2 \times 3 \times 9$ (Order $\times$ Presentation Duration $\times$ Numerosity) ANOVA was performed on the data. The main effects of Order, $F(1, 9) = 25.988$, $p < 0.001$, and Numerosity, $F(8, 72) = 594.673$, $p < 0.001$, and the Order $\times$ Numerosity interaction, $F(8, 72) = 11.149$, $p < 0.001$, were found to be significant; while the Presentation Duration main effect, $F(2, 18) = 0.683$, $p = 0.518$, the Order $\times$ Presentation Duration, $F(2, 18) = 0.152$, $p = 0.860$, the Presentation Duration $\times$ Numerosity, $F(16, 144) = 0.390$, $p = 0.983$, and the Order $\times$ Presentation Duration $\times$ Numerosity interactions, $F(16, 144) = 1.094$, $p = 0.366$, were not.

Logistic functions were fitted to each observer's data separately for the three blocks (Figures 5B, 5C). One-sample t tests showed that the intercepts in the 40-ms, 240-ms, and 960-ms conditions at the relative numerosity of 100% (48 dots, where the number of the dots was equal in the same-color and different-color streams) were not significantly different from 0.5, 40 ms: $t(9) = -0.791$, $p = 0.449$; 240 ms: $t(9) = -1.132$, $p = 0.287$; 960 ms: $t(9) = -2.260$, $p = 0.050$. Furthermore, a one-way repeated measures ANOVA revealed no significant difference in estimated intercept between the three conditions, $F(2, 18) = 0.252$, $p = 0.780$. These results indicated that, in static displays, the observers did not perceive a smaller number of dots in same-color displays compared to different-color displays. For the estimated β of the fitted functions, there was no significant difference between the three conditions, $F(2, 18) = 0.461$, $p = 0.638$.

Numerosity underestimation for same-color displays occurred only with the fast-rate dynamic displays. This may be relevant to the fact that the visual system has limited spatiotemporal resolution for processing incoming information. At a slow presentation speed, the visual system is able to accurately identify and localize each individual object. If the incoming information enters the visual system at a speed beyond its processing capacity, some information may go into the system without being fully processed and without ultimately reaching conscious perception. This can be related to the object substitution account of visual masking proposed by Enns and Di Lollo (1997; see also Di Lollo et al., 2000), according to which objects can go unnoticed by conscious perception (i.e., objects are "masked") if new information is rapidly fed into the processing system before the previous information has been fully processed. In such a case, the representation of the scene at the present moment overwrites and substitutes the previous representation. In Experiment 1, when the dots were presented at a fast rate in the dynamic condition, processing of some dots might not have been completed; spatiotemporal proximity (Allik & Tuulmets, 1993) and proximity in feature space (i.e., similarity) of same-color dots might have created the illusion of smaller numerosity due to masking among the same-color dots or registration of multiple dots as a single object by the processing system.

Experiment 3

Studies have suggested that the processes of counting and subitizing are different, in that counting requires spatial attention, while subitizing relies on a limited-capacity process that occurs before attention and after pre-attentive feature detection and grouping operations (Trick & Pylyshyn, 1993, 1994). However, recent studies that combined numerosity judgments with attentional blink or inattentional blindness tasks to manipulate the amount of available attentional resources have suggested that subitizing also requires attention. Specifically, these studies have demonstrated remarkable decreases in subitizing accuracy when observers are engaged in an attentional demanding primary task, but numerosity estimation is not affected by additional attentional demands (Burr et al., 2010; Egeth et al., 2008; Olivers & Watson, 2008; Railo et al., 2008; Vetter et al., 2008). Experiment 3 aimed to examine whether exploiting the availability of attentional resources by having observers simultaneously engage in an additional task would influence the underestimation effect in same-color dynamic displays.

Methods

Observers

Ten new naïve observers were recruited to participate in Experiment 3. All had normal or corrected-to-normal vision and gave informed consent before the experiment.

Stimuli and procedure

The experimental stimuli were created based on those used in Experiment 1. Only the fast-rate (40-ms frames) stimuli were used. Additionally, instead of presenting the whole range of the nine conditions as in Experiment 1, only five conditions were presented: 50% or 100% of frames showing one dot, 50% or 100% of frames showing three dots, and the 0% condition where two dots were presented in all frames.

In addition to the dot stimuli, four white digits (height = 0.68°, width = 0.54°) randomly selected from one to nine were presented sequentially at the center of the screen instead of the fixation cross during each stream in each trial. There were three attentional load conditions performed in separate blocks: no-load, low-load, and high-load. In the no-load condition, observers were instructed to keep fixating at the changing digit while viewing the two dot streams and to judge which of the streams appeared to contain more dots, as in Experiment 1. In the low-load condition, one of the four white digits (randomly selected) in each stream was shown in blue. Observers were asked to report the two blue digits (one for each stream) with the number pad first and then make their numerosity judgment. In the high-load condition, two of the four digits were displayed in blue and observers had to report the four blue digits (two for each stream) before making the numerosity judgment. For the low-load and high-load dual-task conditions, observers were instructed to give priority to the digit task over the numerosity task, and to ensure that all digits were input correctly in each trial. There was a short 5-min break between blocks. The experiment lasted approximately 30 min.

Results and discussion

The observers reported the blue digits with an average accuracy of 94% in the low-load condition and 88.1% in the high-load condition. The proportion of trials where observers judged that the same-color stream contained more dots was computed for each of the three attentional load conditions (only including trials where all the blue digits were correctly reported). The results are shown in Figure 6A.

An omnibus 2 × 3 × 5 (Order × Attentional Load × Numerosity) ANOVA was performed, and the results revealed a significant main effect of Numerosity, $F(4, 36) = 103.456$, $p < 0.001$, Order × Attentional Load interaction, $F(2, 18) = 7.761$, $p = 0.004$, and Attentional Load × Numerosity interaction, $F(8, 72) = 8.384$, $p < 0.001$. The main effects of Order, $F(1, 9) = 3.751$, $p = 0.085$, and Attentional Load, $F(2, 18) = 0.121$, $p = 0.887$, the Order × Numerosity interaction, $F(4, 36) = 2.132$, $p = 0.097$, and the Order × Attentional Load × Numerosity interaction, $F(8, 72) = 0.456$, $p = 0.883$,

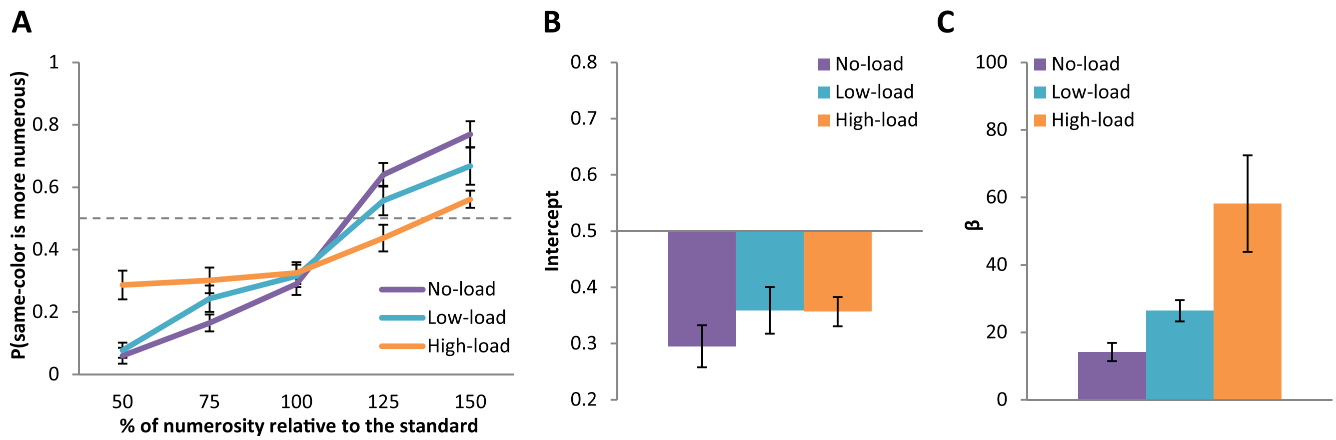

Figure 6. The proportion of trials the observers responded that the same-color stream contained a greater number of dots than the different-color stream in Experiment 3 (Panel A); average estimated proportion value (i.e., intercept) at $x = 100\%$ (Panel B); average β of the fitted logistic function (Panel C); error bars showing the standard error of the mean.

were not significant. Simple main effect analyses revealed a significant main effect of Order in the high-load condition only ($p < 0.05$), suggesting that the presentation order might have influence on observers' response when the task became difficult in the high-load condition. In addition, the main effect of Numerosity was significant in all the three load conditions (all ps < 0.001).

Since Experiment 1 established the underestimation effect for the same-color stream, we performed the same logistic fitting and employed one-tailed hypothesis tests to examine whether the intercepts were smaller than 0.5 (Figure 6B, 6C). One-sample one-tailed t tests revealed that the intercepts were significantly smaller than 0.5 in the no-load, low-load, and high-load conditions, no-load: $t(9) = -5.465$, $p < 0.001$; low-load: $t(9) = -3.401$, $p = 0.004$; high-load: $t(9) = -5.521$, $p < 0.001$. A one-way repeated measures ANOVA revealed no significant difference in intercept between the three conditions, $F(2, 18) = 0.771$, $p = 0.477$. For the β of the fitted functions, the three load conditions were found to differ significantly, $F(2, 18) = 6.840$, $p = 0.006$, with the steepest slope in the no-load condition and the shallowest slope in the high-load condition. The β of the no-load versus high-load pair and low-load versus high-load pair were significantly different from each other ($p < 0.05$ corrected for multiple comparisons), while the no-load versus low-load pair was not.

The comparable intercepts in all three conditions (smaller than 0.5 and not significantly different from each other) suggest that comparable numerosity underestimation effects were observed in all conditions. In contrast, the significant difference in slopes between the conditions suggest that as the attentional load (i.e., the task difficulty) increased, the precision of numerosity judgments decreased, with the shallowest slope indicating the lowest precision in the high-load condition.

Egeth et al. (2008) used the rapid serial visual presentation paradigm, where streams of letters containing a target letter (in a specified color) were presented, to study the effect of attention on numerosity judgments. Observers had to correctly report the target letter and at the same time judge the number of dots presented in the peripheral region of the screen. When the dots were presented soon after the presentation of the target letter (i.e., the period of the "attentional blink"), performance on the numerosity task was markedly reduced, even when the number of objects was clearly within the subitizing range (e.g., two or three objects). These results were also supported by an event-related potential (ERP) study (Xu & Liu, 2008). In Olivers and Watson (2008), the lag between the letter identification task and the numerosity judgment task influenced performance, providing a powerful demonstration that attention is involved in numerosity judgments in the subitizing range. In addition to using attentional blink tasks, dual-task experiments that control the amount of available attentional resources by engaging observers in a spatial attention task (Burr et al., 2010, 2011) also suggest that an attention-dependent mechanism is responsible for subitizing but not estimation of larger numbers. The numerosity judgments in our task with dynamic displays was in the range of estimation (not subitizing), which has been previously found in studies using static displays to be unaffected by attentional load (Burr et al., 2010). Burr et al. (2010, 2011) suggested that an increase in attentional load decreased precision (i.e., increased variability in responses) of numerosity judgments but accuracy (i.e., the mean perceived numerosity) was not affected. This is consistent with the present results showing that the underestimation effect (based on mean perceived numerosity) was observed regardless of attentional load condition, while attentional load increased the variability in responses

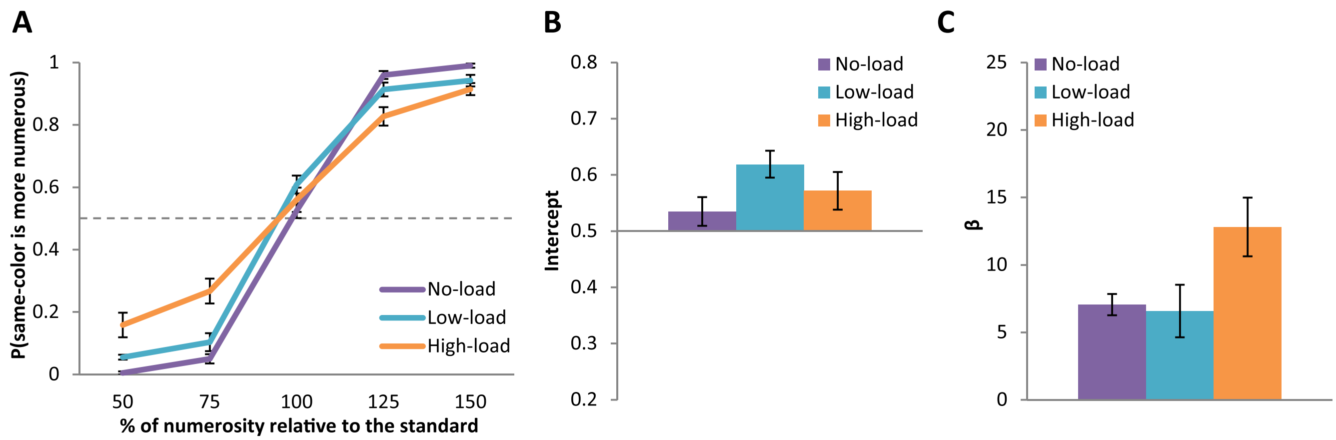

Figure 7. The proportion of trials the observers responded that the same-color stream contained a greater number of dots than the different-color stream in Experiment 4 (Panel A); average estimated proportion value (i.e., intercept) at $x = 100\%$ (Panel B); average β of the fitted logistic function (Panel C); error bars showing the standard error of the mean.

(resulting in the shallower slopes of the response curves in the low- and high-load conditions).

Experiment 4

Experiment 4 was conducted as a control experiment and was a static display version of Experiment 3. That is, we examined whether attentional demands would interact with numerosity estimation for static displays.

Methods

Observers

Ten new naïve observers were recruited to participate in Experiment 4. All had normal or corrected-to-normal vision and gave informed consent before the experiment.

Stimuli and procedure

The stimuli and procedure were based on Experiment 2, but only the 960-ms stimuli were used. Furthermore, instead of presenting the whole range of number of dots conditions (the nine levels from 24 to 72 dots), only the 24, 36, 48, 60, and 72 dots conditions were presented. The three attentional load conditions and the dual-task design were the same as Experiment 3. For each trial in the low-load and high-load blocks, observers first reported the digits in blue using the number pad and then judged whether the first or second display contained more dots. The order of the three attentional load blocks (no-load, low-load, and high-load) was counterbalanced, and there was a 5-min break between blocks.

Results and discussion

The observers reported the digits with an average accuracy of 93.5% in the low-load condition and 84.4% in the high-load condition. The numerosity judgment data are shown in Figure 7A.

An omnibus 2 × 3 × 5 (Order × Attentional Load × Numerosity) ANOVA showed significant main effects of Order, $F(1, 9) = 20.995$, $p = 0.001$, and Numerosity, $F(4, 36) = 861.909$, $p < 0.001$, Order × Numerosity interaction, $F(4, 36) = 6.878$, $p < 0.001$, and Attentional Load × Numerosity interaction, $F(8, 72) = 10.307$, $p < 0.001$. The Attentional Load main effect, $F(2, 18) = 3.245$, $p = 0.063$, the Order × Attentional Load interaction, $F(2, 18) = 2.024$, $p = 0.161$, and the Order × Attentional Load × Numerosity interaction, $F(8, 72) = 0.412$, $p = 0.910$, were found to be not significant. Simple main effect analyses found that the effect of Numerosity was significant across all the three load conditions (all $ps < 0.001$).

To test whether the intercepts estimated at the numerosity of 48 dots in each load condition were significantly smaller than 0.5, one-sample t tests were conducted. The analysis showed that the intercepts for the three conditions were not significantly smaller than 0.5, no-load: $t(9) = 1.368$, $p = 0.898$; low-load: $t(9) = 4.997$, $p = 0.999$; high-load: $t(9) = 2.134$, $p = 0.969$. A one-way repeated measures ANOVA revealed no significant difference in intercepts across the three load conditions, $F(2, 18) = 2.969$, $p = 0.077$. The data is therefore in line with Experiment 2 and confirm that the underestimation effect does not occur in static displays. There was a marginally significant difference in β of the fitted function between the three conditions, $F(2, 18) = 3.436$, $p = 0.054$, with the shallowest slope in the high-load condition. This again might reflect the effect of attentional load on numerosity judgment precision.

General discussion

The present study demonstrated numerosity underestimation when a large number of same-color objects, compared with different-color objects in a similar configuration, are displayed in a dynamic stream. This underestimation effect was only observed in fast-rate dynamic streams, but not in slow-rate streams or static displays. We speculate that this might be due to object substitution (Di Lollo et al., 2000; Enns & Di Lollo, 1997) during fast stimulus presentations among same-color objects that are spatiotemporally and featurally proximal.

Applying the idea of object substitution (as described in the Introduction) to the context of the present study, in the fast-rate same-color streams, as new dots keep appearing at different locations around the screen, new incoming information continuously congests the input layer of the processing system. Information about a dot at a previous moment is quickly replaced by new dots at new positions without being adequately processed to enter conscious perception. If the position of a dot is close to a subsequently presented dot, the new object is more likely to substitute the old one, preventing the old object from reaching consciousness, and leading to an overall perception of smaller numerosity. This is consistent with the prediction made by Allik and Tuulmets (1993) that perceived numerosity decreases with increased spatial and temporal proximity between items in the scene. Furthermore, object substitution (or object updating) is eliminated when the salience of the critical event is high (Moore & Enns, 2004)—the flash-lag effect is eliminated when the tracked moving object is made salient by abruptly changing its appearance at the moment of the flash. This might explain the difference in perceived numerosity between the different-color and same-color streams. As the different-color streams are composed of a mix of different dots and have greater color contrast than the same-color streams, the different color dots might appear more salient to the observer, which may strengthen the signals at the pattern layer of the system. The stronger signal allows the information to remain in the pattern layer for a longer period before decaying completely. This increases the chance that the information will be fully processed by the system and reach conscious perception. As a result, object substitution in the different-color condition does not occur as strongly as in the same-color condition.

Other models of perceived numerosity, such as the occupancy model proposed by Allik and Tuulmets (1991), may also offer relevant explanations to the effects observed here. According to the occupancy model, in a field containing a large number of dots, each dot occupies a circular territory around itself, and the visual system judges numerosity based on the total size of the area occupied by all the dots. When the territories of two dots overlap, the total area occupied by all the dots is smaller. In this case, observers tend to perceive smaller numerosity compared to the case where dot territories do not overlap (i.e., dots that are distributed more evenly and sparsely). In other words, two dots that are closer together have a smaller total impact on perceived numerosity than two dots that are further away from each other; a dot can be masked by another nearby dot (Allik & Tuulmets, 1991). In the dynamic display condition in Experiment 1, the different-color stream may have greater perceived contrast than the same-color stream, and therefore may appear more salient to the observer. Such increased salience might minimize the effect of masking between dots at neighboring positions, leading the visual system to clearly regard them as two distinct identities. Given this advantage, dot pairs of different colors can survive even when the territories they occupy overlap. This might represent an interaction between distinctive object features of the items (physical) and the occupancy of territories of each item (psychological).

He et al. (2009) showed that numerosity judgment in a cloud of dots depends on connectedness among the elements. They claimed that their observation of numerosity underestimation in displays with connected dots reflects processing of perceptual organization from which representations of distinct objects are formed. Specifically, the configuration of two dots joined by a line is represented as a single object, whereas two dots in the same configuration that are not connected by a line are represented as two distinct objects. As a result, a display with lines connecting the dots leads to the illusion of smaller numerosity than one without connecting lines. Such perceptual organization of "single-objectness" might also be relevant to the numerosity underestimation effect in dynamic displays with elements of the same color in the present study. This is also relevant to one of the classical Gestalt principles—the similarity principle, concerning perceptual organization (Koffka, 1935; Köhler, 1928; Wertheimer, 1938). According to the similarity principle, elements that are similar to each other tend to be grouped together. This increases the chance that the visual system processes the groups as single objects. In the same-color dynamic displays in Experiment 1, there may be perceptual interactions among the dots of the same color. Although two dots might be spatially distant from each other, such interactions might lead the visual system to organize the pair of dots into a single object when the time to register perceptual information is very brief (e.g., 40 ms), as in the fast-rate condition. In contrast, in different-color streams, the strength of this kind of organization would be weaker as the different colors inhibit the tendency of the processing system in pairing up the dots. Consequently, displays with dots of the same color are perceived as less numerous than displays with dots of different colors in the fast-rate dynamic condition. Another possibility is that apparent motion (phi phenomenon; Wertheimer, 1912) might have occurred more frequently among dots with the same color than dots with different colors, due to similarity in appearance. In such case, the total number of objects registered in the same-color stream by the visual system would be smaller than that in the different-color stream, leading to the underestimation effect observed in the present results. However, the results of our control experiments (no difference in perceived motion smoothness and flickering rate) point to otherwise.

For the data in Experiments 1 and 2, the asymptotes of the psychometric functions in all the static conditions (Figure 5A) approximately reached one at the right end, but it was not the case for the dynamic conditions (both the fast-rate and slow-rate conditions in Figure 2A). Conceptually, an asymptote with a value below one at the right end would mean that the psychometric curve shift downward vertically, which implies an underestimation effect. In the slow-rate dynamic condition, a weak underestimation effect might exist of which the magnitude was just marginal. We suspect that the presentation frequency of 240 ms per stimulus might be approaching to the threshold of the temporal frequency for object substitution to occur in our experiment. This is consistent with previous studies that object substitution masking starts to emerge when the stimulus-onset asynchrony between the target and the mask approaches 150 ms (Enns, 2004). Another possibility also exist; the task of numerosity judgment in the dynamic conditions was difficult, such that observers' performance could not reach 100% accuracy even in the extreme conditions.

Regarding the technique of dynamic stimulus presentation used in Experiments 1 and 3 in the present study, one may argue that it relies upon the presentation of numbers of dots (i.e., one, two, or three dots) which can be readily subitized; thus there are possibilities that the visual system processes the dynamic stream by rapidly subitizing the numerosity presented in each frame and combines the subitized estimates at the end of the stream to come up with an overall estimate. If this were the case, the underestimation effect observed in the dynamic stream conditions in the present study (at least in the slow-rate 240-ms condition) might be actually related to a memory failure rather than perception. Further examinations are needed to clarify the role of memory in the integration of individual numerosities in dynamic visual stimulus.

In the dynamic streams used in the present study, there were always two dots presented in each frame of the different-color stream, while there were a proportion of frames in the same-color stream presented with either one dot or three dots, with all the rest containing two. One may argue that observers might be able to determine readily whether the same-color stream is more numerous or not simply by detecting the existence of a non-two dot frame among the dynamic same-color stream, and the task is more like detection for a non-two dot frame rather than judgment based on the perception of overall numerosity. Additionally, the shallower slope in the response curve of the high-load condition in Experiment 3 might be representing the impairment of detection performance by the introduction of a secondary task. However, we conducted a control experiment that presented varied proportions of all the one-dot, two-dot, and three-dot frame types in all the same-color streams (at 40 ms/frame), and the results showed similar underestimation effect as in the fast-rate dynamic condition of Experiment 1. Therefore, it is not likely that the observers in the present study performed the task by detecting a non-two dot frame instead of judging based on the overall numerosity of the dynamic streams.

The results of Experiment 3 suggest that the numerosity underestimation effect in dynamic displays is not affected by attentional manipulations. Although precision in the numerosity judgment task dropped remarkably as attentional load increased (indicated by the shallower slope in the higher load conditions), the underestimation effect remained evident in all attentional load conditions. This is consistent with previous work showing that while the process of subitizing requires attention (Egeth et al., 2008; Olivers & Watson, 2008; Railo et al., 2008; Xu & Liu, 2008), the processing of estimating large numerosities does not (Burr, et al., 2010; Vetter et al., 2008). Future studies should explore the involvement of attention in numerosity judgments for dynamic stimuli in the subitizing range.

Conclusions

The present study demonstrated that numerosity underestimation occurs for same-color objects, compared to different-color objects, in dynamic visual displays with fast presentation rate. This effect was not evident when a slow presentation rate or static displays were used. In a dual-task paradigm where the attentional resources available for numerosity comparison were manipulated, we found that a high

attentional load reduced numerosity judgment precision; however, the underestimation effect in the same-color stream was still evident. These findings are consistent with object substitution, where objects that are identical in appearance and that are spatiotemporally proximal may mask each other, leading to an overall misperception of fewer items in dynamic streams.

Keywords: attentional load; dynamic presentation; numerosity; same-color objects

Acknowledgments

This work was supported by Grant-in-Aid for Scientific Research (KAKENHI) #23240034 from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan; Japan Society for the Promotion of Science (JSPS) #11J00287.

Commercial relationships: none.
Corresponding author: Ricky K. C. Au.
Email: ricky@fennel.rcast.u-tokyo.ac.jp.
Address: Research Center for Advanced Science and Technology, The University of Tokyo, Meguro-ku, Tokyo, Japan.

Footnote

[1] Note that in this paper we define *underestimation effect* as the underestimation of numerosity in the same-color stimulus relative to the different-color stimulus but not absolute underestimation or overestimation.

References

Afraz, S.-R., Kiani, R., Vaziri-Pashkam, M., & Esteky, H. (2004). Motion-induced overestimation of the number of items in a display. *Perception, 33,* 915–925.

Allik, J., Helsper, E., & Vos, P. (1991). Numerosity in textures with different spatial distribution of elements. *Perception, 20,* 87.

Allik, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception and Psychophysics, 49,* 303–314.

Allik, J., & Tuulmets, T. (1993). Perceived numerosity of spatiotemporal events. *Perception and Psychophysics, 53,* 450–459.

Allik, J., Tuulmets, T., & Vos, P. G. (1991). Size invariance in visual number discrimination. *Psychological Research, 53,* 290–295.

Ansari, D., Lyons, I. M., van Eimeren, L., & Xu, F. (2007). Linking visual attention and number processing in the brain: The role of the temporo-parietal junction in small and large symbolic and nonsymbolic number comparison. *Journal of Cognitive Neuroscience, 19,* 1845–1853.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Bullier, J., McCourt, M. E., & Henry, G. H. (1988). Physiological studies on the feedback connection to the striate cortex from cortical areas 18 and 19 of the cat. *Experimental Brain Research, 70,* 90–98.

Burgess, A., & Barlow, H. B. (1983). The precision of numerosity discrimination in arrays of random dots. *Vision Research, 23,* 811–820.

Burr, D., & Ross, J. (2008). A visual sense of number. *Current Biology, 18,* 425–428.

Burr, D. C., Anobile, G., & Turi, M. (2011). Adaptation affects both high and low (subitized) numbers under conditions of high attentional load. *Seeing and Perceiving, 24,* 141–150.

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

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44,* 1–42.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129,* 481–507.

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.

Enns, J. T. (2004). Object substitution and its relation to other forms of visual masking. *Vision Research, 44,* 1321–1331.

Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science, 8,* 135–139.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences, 4,* 345–352.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1,* 1–47.

Fründ, I., Haenel, N. V., & Wichmann, F. A. (2011). Inference for psychometric functions in the presence of nonstationary behavior. *Journal of Vision*, *11*(6):16, 1–19, http://www.journalofvision.org/content/11/6/16, doi:10.1167/11.6.16. [PubMed] [Article]

Gallace, A., Tan, H. Z., & Spence, C. (2007). Multisensory numerosity judgments for visual and tactile stimuli. *Perception and Psychophysics, 69,* 487–501.

Gebuis, T., & Reynvoet, B. (2012). The role of visual information in numerosity estimation. *PLoS ONE, 7*(5), e37426, 1–5, doi:10.1371/journal.pone.0037426.

Ginsburg, N. (1991). Numerosity estimation as a function of stimulus organization. *Perception, 20,* 681–686.

Ginsburg, N., & Nicholls, A. (1988). Perceived numerosity as a function of item size. *Perceptual and Motor Skills, 67,* 656–658.

Goldstone, R. L. (1993). Feature distribution and biased estimation of visual displays. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 564–579.

He, L., Zhang, J., Zhou, T., & Chen, L. (2009). Connectedness affects dot numerosity judgment: Implications for configural processing. *Psychonomic Bulletin and Review, 16,* 509–517.

Hollingsworth, W. 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.

Hyde, D. C., & Spelke, E. S. (2011). Spatiotemporal dynamics of processing nonsymbolic number: An event-related potential source localization study. *Human Brain Mapping, 33,* 2189–2203.

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

Koffka, K. (1935). *Principles of Gestalt psychology*. New York, NY: Harcourt, Bruce, and World.

Köhler, W. (1928). An aspect of Gestalt psychology. In C. Murchison (Ed.), *Psychologies of 1925* (pp. 163–195). Worcester, MA: Clark University Press.

Krueger, L. E. (1972). Perceived numerosity. *Perception and Psychophysics, 11,* 5–9.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science, 14,* 396–401.

Lipton, J. S., & Spelke, E. S. (2004). Discrimination of large and small numerosities by human infants. *Infancy, 5,* 271–290.

Moore, C. M., & Enns, J. T. (2004). Object updating and the flash-lag effect. *Psychological Science, 15,* 866–871.

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

Palomares, M., & Egeth, H. (2010). How element visibility affects visual enumeration. *Vision Research, 50,* 2000–2007.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Plaisier, M. A., Tiest, W. M. B., & Kappers, A. M. L. (2010). Range dependent processing of visual numerosity: Similarities across vision and haptics. *Experimental Brain Research, 204,* 525–537.

Railo, H., Koivisto, M., Revonsuo, A., & Hannula, M. M. (2008). The role of attention in subitizing. *Cognition, 107,* 82–104.

Raphael, S., Dillenburger, B., & Morgan, M. (2013). Computation of relative numerosity of circular dot textures. *Journal of Vision*, *13*(2):17, 1–11, http://www.journalofvision.org/content/13/2/17, doi:10.1167/13.2.17. [PubMed] [Article]

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

Ross, J., & Burr, D. C. (2010). Vision senses number directly. *Journal of Vision*, *10*(2):10, 1–8, http://journalofvision.org/10/2/10/, doi:10.1167/10.2.10. [PubMed] [Article]

Shuman, M., & Spelke, E. (2006). Area and element size bias numerosity perception. *Journal of Vision*, *6*(6):777, http://www.journalofvision.org/content/6/6/777, doi:10.1167/6.6.777. [Abstract]

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

Tokita, M., & Ishiguchi, A. (2010). How might the discrepancy in the effects of perceptual variables on numerosity judgment be reconciled? *Attention, Perception and Psychophysics, 72,* 1839–1853.

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,* 80–102.

Vetter, P., Butterworth, B., & Bahrami, B. (2008).

Modulating attentional load affects numerosity estimation: Evidence against a pre-attentive subitizing mechanism. *PLoS ONE, 3*(9), e3269, 1–6, doi:10.1371/journal.pone.0003269.

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

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. [Translation: Experimental studies on the seeing of motion] *Zeitschrift für Psychologie, 61,* 161–265.

Wertheimer, M. (1938). Laws of organisation in perceptual forms. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 71–88). New York, NY: Harcourt, Brace.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception and Psychophysics, 63,* 1293–1313.

Xu, X., & Liu, C. (2008). Can subitizing survive the attentional blink? An ERP study. *Neuroscience Letters, 440,* 140–144.