Divided attention effects in visual search are caused by objects not by space

Cathleen M. Moore Department of Psychological and Brain Sciences, University of Iowa Iowa City, Iowa, USA

James Pai Department of Psychology, University of Washington, Seattle, WA, USA

John Palmer Department of Psychology, University of Washington, Seattle, WA, USA

Divided attention effects have been observed across a variety of stimuli and perceptual tasks, which have given rise to both object-based and space-based theories of divided attention. Object-based theories assert that processing information from multiple objects is limited, whereas space-based theories assert that processing information from multiple locations is limited. Extant results in the literature are collectively inconsistent with both simple object-based theories and simple space-based theories of divided attention. Using a visual search task with the extended simultaneous-sequential method to reveal capacity limitations, we found evidence of limited-capacity processing of object properties and unlimited-capacity processing of feature contrast. We found no evidence of a separate spatial limitation. A multiple pathway processing theory can account for these and a large body of previous results. According to this theory, tasks that require object processing must follow a limited-capacity pathway and therefore incur divided attention effects. Tasks that depend on only feature contrast can follow a separate unlimited-capacity processing pathway and therefore do not incur divided attention effects.

Introduction

Studies of visual perception have shown that for some tasks, increasing the number of relevant stimuli reduces performance, whereas for other tasks, increasing the number of relevant stimuli has no effect on performance. Such *divided attention* effects are assumed to be caused by one or more aspect of processing that is limited, and are explained in terms of the amount of information that can be processed per unit time (e.g., a trial), referred to as *capacity*. Tasks that show no divided attention effects are inferred to engage only unlimited-capacity processes, whereas tasks that show divided attention effects are inferred to engage one or more limited-capacity process. With divided attention effects, we ask what aspect of processing is limited?

The current study tested two theories of divided attention: that the number of objects that can be processed at one time is limited (*object-based theories*) and that the number of locations from which stimuli can be processed at one time is limited (*space-based theories*). Both could be true. According to object-based theories, the effects of divided attention are imposed by having to process multiple objects rather than a single object (e.g., Duncan, 1984). Pure object-based theories (i.e., ones for which object processing is the only source of limitation) maintain that processing multiple attributes of a single object adds no cost relative to processing a single attribute from a single object. In contrast, according to *space-based theories,* the effects of divided attention are imposed by having to process information from multiple *locations* rather than from a single location (e.g., Posner, 1980). Pure space-based theories (i.e., ones for which spatial location is the only source of limitation) maintain that processing attributes from multiple objects at a single location adds no cost relative to processing the same attributes from a single object at one location.

A metaphor that is often used to illustrate space-based theories is a spotlight that can point to only one location at a time (e.g., Posner, Synder, & Davidson, 1980). A spotlight reveals a region of space without regard to what is in it. In a theater setting, the spotlight highlights not only actors who are in the illuminated region, but also the stage, back curtain, and any props that are in the region as well. Revealing actors who are currently outside of the illuminated region requires movement of the spotlight, and after being moved, anything that was previously highlighted no longer is. This metaphor is often used to talk about the process of

Citation: Moore, C. M., Pai, J., & Palmer, J. (2022). Divided attention effects in visual search are caused by objects not by space. *Journal of Vision*, *22*(12):2, 1–25, https://doi.org/10.1167/jov.22.12.2.

https://doi.org/10.1167/jov.22.12.2 Received April 18, 2022; published November 1, 2022 ISSN 1534-7362

selecting information based on spatial location. It also, however, implies a specific (spatial) capacity limitation, which is that by focusing the spotlight on one area, other areas are less well illuminated. This metaphor yokes selective attention (on what basis is information selected) and divided attention (what aspects of processing are limited) in a way that we will argue is unnecessary and misleading, but it is nonetheless a popular metaphor that has driven a lot of theorizing in the attention literature.

A comparable metaphor for object-based theories is a glow-in-the-dark safety vest. In a construction setting, a worker who is wearing the vest is highlighted, but the road or other objects around the worker are not highlighted. A worker with the vest is visible regardless of where they move, whereas workers without the vest are relatively invisible regardless of where they move. Multiple workers could be "grouped" into a single entity if the vest were somehow draped around them, but then they would be revealed as a unit, organized by the vest. Highlighting a currently invisible worker requires that a vest be passed to them, and whoever passed the vest will no longer be visible. This metaphor, like the spotlight, also yokes selective and divided attention. Selection occurs on the basis of who has the vest, but it implies a capacity limit in that whoever does not have the vest is not highlighted. We will argue that which processes are limited and the basis of selection are not yoked as these metaphors lead us to think.

Object-based and space-based theories of divided attention propose distinctly different sources of processing limitation. However, it is challenging to discriminate between them because the number of locations is often confounded with the number of objects. Two objects, for example, usually appear in two different locations. The general approach to testing between the theories has been to hold constant the number of locations in which relevant attributes appear while manipulating through various methods the number of objects.

One strategy of testing between object-based and space-based theories of divided attention is to present stimuli so that they are overlapping each other at a single location in the image, and comparing judgments of one versus two objects from identical displays. This approach was first used by Rock and Gutman (1981; see also Neisser & Becklen, 1975) to study *selective* attention. They presented overlapping shape stimuli (Figure 1a), and tested whether observers could parse them and selectively process information from one of the two stimuli and found that they could. This is consistent with objects, rather than spatial locations, being selected because there was only one location. Duncan (1984) extended Rock's overlapping-stimuli strategy to distinguish between object-based and space-based *divided* attention. He used displays with an outline rectangle and an overlapping tilted line (Figure 1b) and tested whether there was a cost for having to report information from two objects compared to one. The rectangle varied in height and whether it had a small gap on the left or right side. The line varied in orientation and whether its texture was dotted or dashed. Divided attention effects were measured using a dual-task method. Specifically, observers made two judgments of either two features of a single object (e.g., rectangle size and gap location) or two features of different objects (e.g., rectangle size and line orientation), and these were compared to conditions in which only a single judgment was necessary. Dual-task performance was worse than single-task performance when the two judgments were about two different objects (e.g., side of gap and orientation of line) but no worse than single-task performance when the two judgments were about a single object (e.g., side of gap and size of box). Because the stimuli were at a single location, the divided attention effect in this study can be attributed to object processing.

An important aspect of Duncan's (1984) study is that identical stimulus displays were used across all of the conditions. This allows any differences across conditions to be attributed to differences in limitations due to attention, rather than other differences (e.g., sensory discriminability). A related strategy has been to use displays with multiple overlapping transparent surfaces that each have various reportable attributes. Following Valdes-Sosa and colleagues (Valdes-Sosa, Cobo, & Pinilla, 1998, 2000), Ernst, Palmer, and Boynton (2012) measured divided attention effects using a dual-task method with overlapping transparent surfaces that were defined by random-dot kinemategrams (Figure 1c). Observers detected changes in motion and changes in luminance from the same surface or different surfaces. The results were consistent with object-based theories of divided attention in that there were dual-task deficits

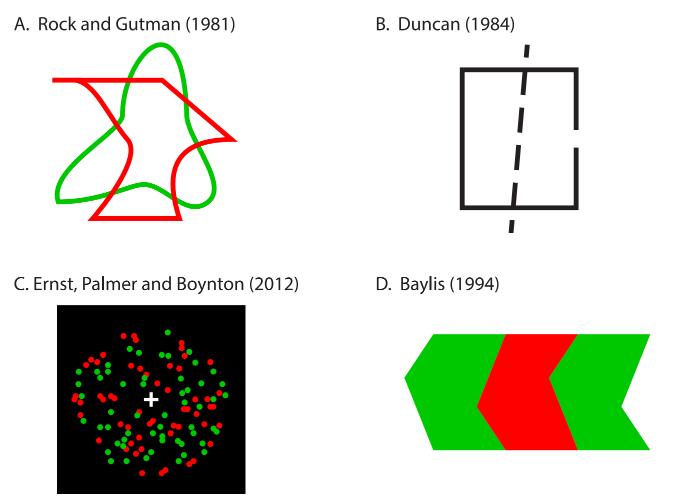

Figure 1. Illustrations of the stimuli used in four studies of divided attention. In all four studies, the same stimulus was presented in the critical conditions. The studies illustrated in panels A,B, and C used the strategy of overlapping two objects. The study illustrated in Panel D manipulated figure-ground relations to vary whether one or two objects were relevant.

for judging attributes on two different surfaces but little or no dual-task deficit for judging attributes on a single surface. The divided attention effects in that study, like Duncan's, can be attributed to object processing in particular because the two surfaces subtended the same region of space, and identical stimulus displays were used across conditions.

Another strategy for using identical displays across divided-attention conditions is to use stimuli with ambiguous figure-ground relations, and manipulate the interpretation. Baylis (1994; see also Baylis & Driver 1993), for example, used stimuli like those shown in Figure 1d, and manipulated whether observers perceived the central region as figure or the two flanking regions as figure. The task was to report the relative position of the two interior vertices between the central and flanking regions. When the central region was perceived as figure, the vertices were parts of one object, whereas when the flanking regions were perceived as figure the vertices were parts of two objects. The results showed that performance was worse when making judgments about vertices of two objects than when making judgments about vertices of one object. This cost can be attributed to limited object processing in particular because the vertices were always in the same locations, and the stimulus displays were identical across conditions; only the perceptual organization of them differed (see Chen, 1998 andChen, 2000 for this strategy applied to selective attention).

The results of the studies reviewed so far are consistent with object-based theories of divided attention. Other studies, however, have yielded evidence that is inconsistent with both simple object-based theories and simple space-based theories of divided attention. Han, Dosher, and Lu (2003; see also Liu, Dosher, & Lu, 2009), for example, showed subjects displays with two Gabor patches at two different locations that varied in their orientation and phase, and asked them to make judgments about two different features that could either be from the two different Gabor patches (e.g., the orientation of one and the phase of the other) or from a single Gabor patch. They found that performance was better for within-Gabor-patch judgments than for between-Gabor-patch judgments, which is consistent with both object-based and space-based theories of divided attention because the Gabor patches were in two different locations. However, when subjects made judgments about a single feature (e.g., orientation or phase), performance was no worse when making that judgment for both Gabor patches (i.e., the orientations of both or the phase of both) than when making it for a single Gabor patch. This lack of a divided attention effect is evidence of unlimited-capacity processing across both objects and locations, because there were two objects and two locations. It is therefore inconsistent with the simplest versions of both object-based and location-based theories of divided attention. Similar results have been found for other simple features and detection-like dual tasks (e.g. Bonnel, Stein & Bertucci, 1992; Graham, Kramer & Haber, 1985).

In addition to dual-task studies such as those of Dosher and colleagues, many studies using visual search have yielded evidence consistent with unlimited-capacity processing of information across *both* multiple objects and multiple locations. Specifically, the pattern of little or no effect of the number of stimuli (i.e., *set size*) on performance when a target is defined by a single feature is consistent with, and is often cited as evidence of, unlimited-capacity processing of feature information across locations (e.g., Treisman & Gelade, 1980; c.f., Wolfe, Cave, & Franzel, 1989). Moreover, studies using a search task that manipulated whether stimuli were presented all at once or in sequential subsets (i.e., the *simultaneous-sequential method*)—a manipulation that controls for the number of decisions and other factors across conditions (e.g., Shiffrin & Gardner, 1972)—also yielded evidence of unlimited-capacity processing of simple features across multiple objects and locations (e.g., Huang & Pashler, 2005; Scharff, Palmer, & Moore, 2011a; Scharff, Palmer, & Moore, 2013). Because stimuli in these experiments were all distinct objects, presented at different locations within the visual field, all of these results are inconsistent with the simplest versions of both object-based and space-based theories of divided attention. The characteristic that they have in common with each other and with the dual-task studies reviewed above (e.g., Han et al, 2003) is that the tasks required only the processing of feature information.

In summary, the collective evidence is inconsistent with the simplest versions of both object-based theories and space-based theories of divided attention. It is clear from studies using overlapping objects, that processing information from multiple objects can incur limited capacity. However, it is also clear, from both dual-task studies (Han et al, 2003; Liu et al., 2009) and visual search studies (e.g., Huang & Pashler, 2005; Scharff et al., 2011a; Scharff et al., 2013; Shiffrin & Gardner, 1972; Treisman & Gelade, 1980), that there is little or no cost to processing information from multiple objects at multiple locations when the judgment is based on some kind of feature information. No theory that asserts a single pathway of processing that has an object-based limit on processing (e.g., Duncan, 1984), a space-based limit on processing (e.g., Posner, 1980) or both (e.g., Vecera, 1997; Baylis & Driver, 1993) can account for all of the results.

As a resolution to this apparent conflict, we propose that information is processed along multiple pathways, and that performance in any given task reveals only the processing limitations that are specific to the information that is relevant to that task. More specifically, we hypothesize, that processing feature-contrast information (e.g., Becker, 2010; Nothdurft, 1991; Nothdurft, 1993; Nothdurft, 2000; Palmer, Verghese, & Pavel, 2000), which need not be attributed to any specific object can have unlimited capacity. Consequently, any task that is based on feature-contrast information can follow an unlimited-capacity processing pathway from stimulus to response. In contrast, we hypothesize that attributes that are intrinsic to objects, such as global shape (e.g., Kahneman, Treisman, & Gibbs, 1992; Popovkina, Palmer, Moore, & Boynton, 2021; Wolfe & Bennett, 1997), must be processed along a limited-capacity pathway. The general idea of two pathways is common to many theories of visual processing (e.g., Hoffman, 1979; Treisman & Gelade, 1980; Wolfe, 1994; Wolfe, 2021; Wolfe et al., 1989). Our version predicts that tasks requiring judgments based on only feature contrast yield no divided attention effects, whereas tasks that depend on object attributes like global shape, yield divided attention effects. It makes no predictions of divided attention effects for multiple locations, per se.

Overview of experiments

To complement previous work using the dual-task method, we used an extended version of the simultaneous-sequential method (e.g., Scharff et al., 2011a). This is a visual-search method that provides a means of distinguishing between unlimited-capacity processing, fixed-capacity processing, and limited but not fixed-capacity processing. Fixed-capacity processing refers to when processing is limited to a constant amount of information (e.g., one object or one location) per unit time (Shaw, 1980). Processing can be limited without being as extremely limited as fixed capacity, and of course it can be unlimited capacity. As reviewed below, the extended simultaneous-sequential method provides a means of discriminating among these alternatives.

Across experiments, we varied the task and whether stimuli were spatially separate or overlapping. To preview the results, we found large divided attention effects that were consistent with fixed capacity in experiments with tasks requiring object processing, regardless of whether the stimuli were in separate locations (Experiments 1 and 2) or overlapping (Experiment 4). In contrast, we found no divided attention effect (i.e., unlimited capacity) in an experiment in which the task depended on feature contrast only and stimuli were presented in separate locations (Experiment 3).

Experiment 1: Shape judgments of multiple objects at separate locations

Experiment 1 used a global-shape discrimination task in the extended simultaneous-sequential method. The stimuli, which were inspired by Rock and Gutman (1981) cannot be discriminated on the basis of simple feature-contrasts (see Figure 2). Stimuli were presented at separate locations.

Method

Stimuli

The stimuli were novel dark gray outline shapes (stroke width was 1 pixel) presented on a mid-gray background. There were three sets of five exemplars each. The sets can be labeled as quadrilaterals, triangles, and curved Rock-like figures (after Rock & Gutman, 1981). The five exemplars from each of the three sets are shown in Figure 2.[1] To create each set, a basic shape was chosen and modified. For example, the set of five quadrilaterals in Figure 1 were created from a base square by perturbing the vertices. Each exemplar was designed to be asymmetrical to provide a unique image at different orientations. All images were 100 × 100 pixels (~4° of visual angle) with the shapes being approximately equal in size within the image bounds. Finally, exemplars could appear at four possible rotations (0°, 90°, 180°, and 270°), for a total of 20 images per set and 60 images for the entire experiment.

A 4° × 4° square of dynamic 1 × 1 pixel noise was combined (contrasts added) with each of the outline shapes. The noise was resampled on every frame of the display. Each pixel of the noise had a luminance drawn from a Gaussian distribution with a standard deviation that was used to define the noise's contrast. In this experiment, the standard deviation of the noise was 50% of the mean luminance (2 observers had 60% noise contrast). Target contrast and stimulus duration were manipulated to adjust task difficulty for each observer. These parameters were adjusted by the experimenter between practice sessions to limit performance. Once

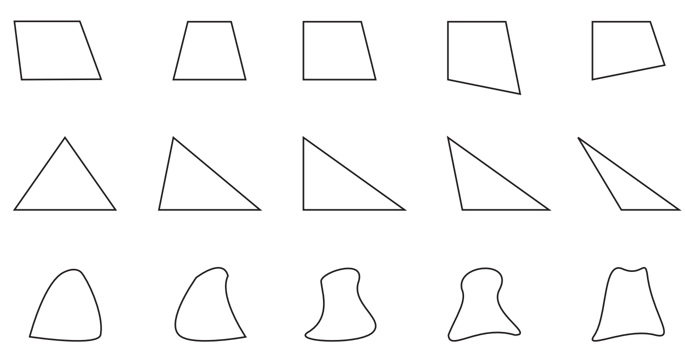

Figure 2. Three stimulus sets with five exemplars each were used in the study. The stimuli were novel outline shapes that included quadrilaterals, triangles, and curved figures inspired by Rock and Gutman (1981). Not shown are the rotated variations (0°, 90°, 180°, and 270°) of each exemplar. This resulted in 20 images per set and thus 60 images overall.

performance stabilized around 80% to 90% correct in the sequential condition, the experiment was begun. No further adjustments were made during the experiment proper. The target contrast ranged from 35% to 50% with a mean of 44%. Stimulus duration ranged from 0.1 to 0.25 second with a mean of 0.18 sencond. A single frame of an example stimulus and noise is shown in panel A of Figure 3. There are different shapes at each location. For this illustration, static noise level is reduced from that used for the dynamic noise in the experiment because static noise has a greater effect than dynamic noise.

Task and design

An extended version of the simultaneous-sequential method was used to measure the effects of divided attention. The task was to report the location of a single prespecified target among three distractors. The critical manipulation for assessing capacity limitations was whether the four stimuli were presented at the same time (*simultaneous*) or in sets of two in separate displays (*sequential*). A third condition in which all four stimuli were presented twice (*repeated*) extended the basic simultaneous-sequential design.

Figure 4 illustrates the three conditions. All trials began with a study display (2 seconds) that defined the target object, which was followed by a blank display (0.5 second), a fixation display (0.5 second), one or two stimulus displays (∼0.2 second each), and a probe display that included a reminder of the target and remained until a response was made. Stimuli within a trial were drawn from a single set (e.g., quadrilaterals). The location of the target was selected randomly for each trial. In order to minimize reliance on simple features to find the target, it was always presented in the stimulus display at a different rotation from that in the study and probe displays. Other than that constraint, stimulus rotation was selected randomly for each stimulus. Dynamic noise was superimposed on stimuli to limit performance (illustrated in the figure by shaded regions).

In the simultaneous condition (left column), all four stimuli were presented in a single stimulus display (∼0.2 second). In the sequential condition, stimuli were presented two at a time (first the upper left and lower right, then the lower left and upper right) across two stimulus displays (∼0.2 second each), separated by a blank interval (1.8 seconds). The target was equally likely to appear in the first or second stimulus display. Finally, in the repeated condition, all four stimuli were presented twice (∼0.2 second each), separated by a blank interval (1.8 seconds). The displays were identical except for the noise.

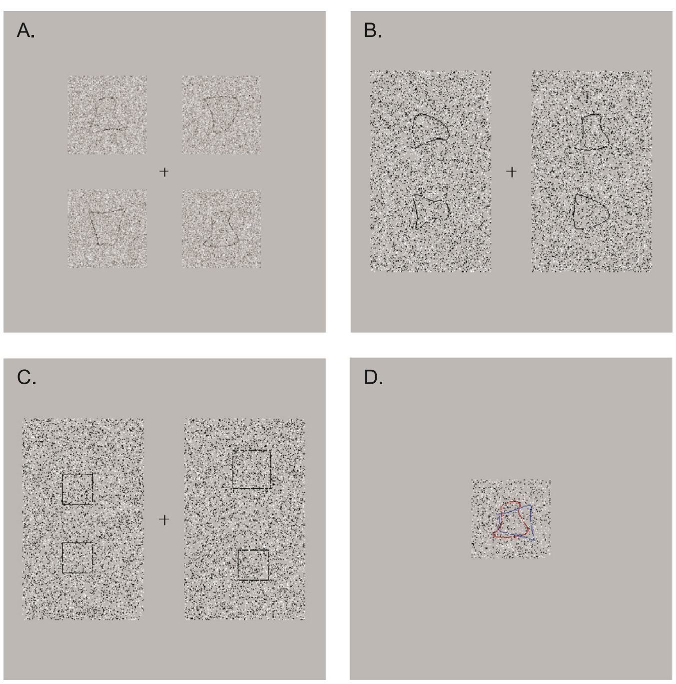

Figure 3. Examples of the stimuli and single frames of the dynamic noise from Experiment 1 (A), Experiment 2 (B), Experiment 3 (C), and Experiment 4 (D). They are all scaled to show the central 16° × 16° of the display. See text for detailed descriptions of each.

Logic

The logic of the extended simultaneous-sequential method is illustrated in Figure 5. It provides a way of discriminating between unlimited-capacity and fixed-capacity models. The numbered circles represent the four stimuli that are presented on each trial, and the gray bars represent the time within the trial each stimulus is present for each of the three conditions. The black arrows inside of the gray bars represent the hypothetical amount of perceptual processing that can be performed on that stimulus, given the constraints of the different models. First consider the unlimited-capacity model. It predicts no difference between the simultaneous and sequential condition, but an advantage for the repeated condition. This is because under an unlimited-capacity model, the amount of processing that is possible for any given stimulus is unaffected by the number of stimuli that are simultaneously present, and therefore the amount of processing that any given stimulus receives is the same in the simultaneous and sequential conditions. However, because the displays are presented twice in the repeated condition, each stimulus receives twice as much processing as it does in either the simultaneous or sequential conditions. Now consider the fixed-capacity model. It predicts an advantage for the sequential condition over the simultaneous condition but no further advantage for the repeated condition. This is because if only two stimuli can be processed within the time of the display duration, then two stimuli could be

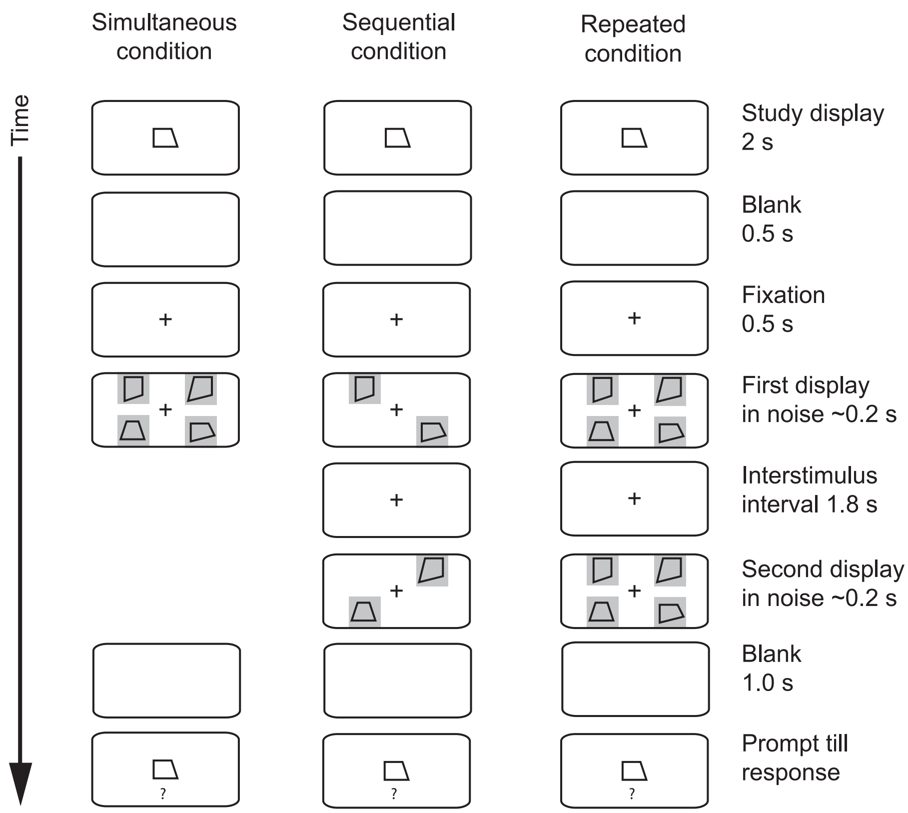

Figure 4. Illustration of the task and design used in Experiment 1. Each trial began with a study display that defined the target shape and ended with a probe display of the same shape. The task was to report in which of four locations the target appeared. In the simultaneous condition (left column), the first and only stimulus display contained four objects, a target (presented at a different rotation from that in the study display) and three distractors. In the sequential condition (middle column), two of the four objects were shown in the first display and the other two were shown in the second display. In the repeated condition (right column), the first stimulus display is like that of the simultaneous condition, but is followed by a second presentation of the same display. Stimuli in the stimulus displays were superimposed with dynamic noise, indicated by the shaded regions in the figure.

processed in the simultaneous condition, whereas four can be processed in the sequential condition, two during the first display and two during the second. Because no more than two stimuli can be processed within a given display duration, however, then repeating displays with all four stimuli present (repeated condition) provides no advantage over just presenting two at a time (sequential condition). Similar arguments can be developed for a fixed-capacity parallel model.

Figure 6 shows the predictions for three benchmark models—unlimited capacity, limited (but not fixed) capacity, and fixed capacity—under a single set of assumptions using a common framework of signal detection theory. These models are formally defined in Scharff et al. (2011a). The predictions here are shown relative to hypothetical performance in the sequential condition fixed at 75%, and a predicted overall effect of 8%. We emphasize the contrasting equality predictions of the unlimited-capacity and fixed-capacity models because they are robust to the details of the model (e.g., distributional assumptions).

Observers

Observers were volunteers with normal or corrected-to-normal visual acuity, some were paid in compensation for their time. All gave informed consent in accord with the Institutional Review Board at the University of Washington in adherence with the Declaration of Helsinki. To determine the appropriate sample size of observers, we examined the results of four similar experiments in Scharff et al. (2013, excluding Experiment 3 with only two observers and Experiment 6 with simple shapes). These experiments all used similar methods to estimate the sequential advantage which is the heart of the current experiments. The observed sequential advantage for each experiment was 10.3%, 9.3%, 8.0%, and 7.0% with a mean of 8.7%. The corresponding sample standard deviation of the sequential advantage was 1.7%, 1.8%, 3.4%, and 4.8% with a mean of 2.9%. Using the mean values, a power analysis was conducted to find the minimum sample size needed to detect a sequential advantage of 8.7% with an alpha of 0.05 and the power of 0.95 (beta of 0.05).

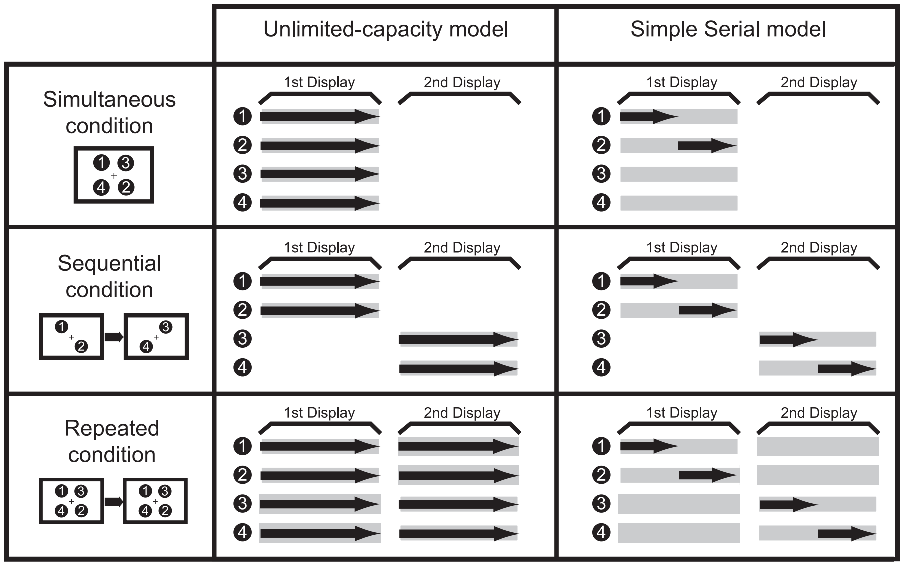

Figure 5. Illustration of the logic of the extended simultaneous-sequential method. The columns correspond to two of the benchmark models and the rows correspond to the three conditions. Each cell illustrates the predictions for a given condition and model. The numbered circles represent the relevant stimuli and the grey bars represent the display time of a stimulus. Within each bar, the black arrows represent the hypothetical processing of the particular stimulus. In the unlimited-capacity model, performance is unaffected by the number of relevant stimuli in a stimulus display. The doubled exposure of each stimulus in the repeated condition yields higher performance. In the fixed-capacity model, the number of relevant stimuli in a stimulus display limits performance. Stimuli can be processed only one after another. In this example, only two stimuli can be processed in a given display. Consequently, two stimuli can be processed in the simultaneous condition and four can be processed in the sequential condition. For this model, the repeated condition gives no advantage over the sequential condition.

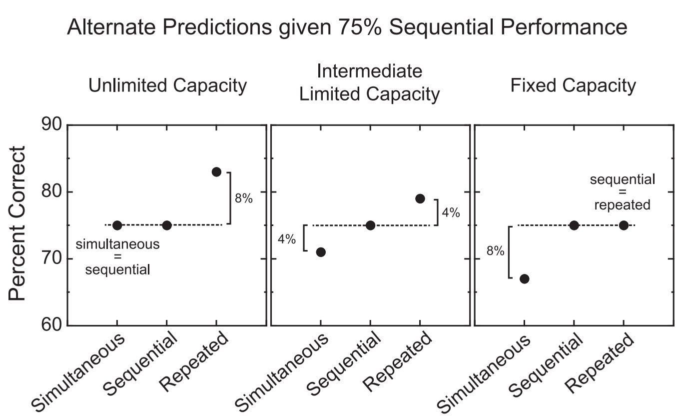

Figure 6. Example predictions of three models are illustrated relative to 75% performance in the sequential condition. See Scharff et al. (2011a) for formal descriptions of the models. As shown in the left panel, unlimited-capacity processing predicts *simultaneous* = *sequential* < *repeated*. As shown on the right panel, fixed-capacity parallel model or a serial model predicts *simultaneous* < *sequential* = *repeated*. And, as shown in the middle panel, an intermediate effect of limited capacity is revealed by the pattern *simultaneous* < *sequential* < *repeated*. The specific 8% difference between simultaneous and repeated is predicted by the further assumption of a yes-no task and Gaussian distributions.

The result was $n = 4$. To be conservative, 6 observers were used as in the Scharff study.

After completing the experiments in this article, we checked how good this strategy was given the results observed here. The sequential advantage in the three global shape experiments were 13.8, 4.9 and 10.5% for a mean of 9.7%. The corresponding sample standard deviations were 5.0, 3.6 and 4.4% for a mean of 4.3%. For these values, the estimated minimum sample size was $n = 5$. Thus our strategy of using a sample size of 6 was reasonable.

Apparatus

Experiment 1 was conducted with two apparatus. Three observers used an original apparatus and three used an updated apparatus. For both, the stimuli were displayed on a flat-screen CRT monitor (19″ ViewSonic PF790) controlled by a Power Mac G4 (Dual 1.0 GHz). The initial apparatus used Mac OS 9.2 and the updated version used Mac OS X 10.4. The stimuli were displayed at a resolution of 832 × 624 pixels, a viewing distance of 60 cm (25.5 pixel/degree at screen center), and a refresh rate of 75 Hz (120 Hz for Mac OS X). The monitor had a peak luminance of 119 cd/m^2, and a black level of 4.1 cd/m^2, mostly because of room illumination. Stimuli were created in Adobe Illustrator CS3 and displayed using Psychophysics Toolbox 2.44 for MATLAB 5.2.1 (Mac OS X: Psychophysics Toolbox 3.0.9 for MATLAB 7.4; Brainard, 1997; Kleiner, Brainard, & Pelli, 2007). A chin rest with an adjustable chair ensured a fixed distance from the display between observers.

Procedure

Observers reported the location of the target (upper-left, upper-right, lower-left, lower-right) using a corresponding key press on a number pad. There was no time limit on the response. Observers completed several practice sessions in which the stimuli were high contrast and long duration to learn the task. Experimental sessions consisted of 12 single-condition blocks 12 trials each (four blocks each of the simultaneous, sequential, and repeated conditions), for a total of 144 trials. Each observer completed 10 experimental sessions resulting in a total of 1440 trials per observer.

Analysis

All statistical tests were two tailed. Alpha was fixed at 0.05. All error bars are the standard error of the mean.

Results and discussion

The results of Experiment 1 are shown in Figure 7. Percent correct is plotted for the three main conditions (simultaneous, sequential, and repeated). Chance performance is 25% correct for this four-alternative localization task. There was a reliable advantage of sequential over the simultaneous presentation with a mean difference of 14% ± 2% (95% confidence interval [CI] = 8.6, 19.0%, $t(5) = 6.84$, $p = 0.001$). There was no reliable difference between the sequential and the repeated conditions, mean differences = 0.5% ± 2.0% (95% CI = −4.8, 5.8%, $t(5) = 0.24$, $p > 0.1$). This pattern of results is consistent with predictions of a fixed-capacity parallel model or a simple serial model and inconsistent with the predictions of an unlimited-capacity parallel. model.

We also tested for temporal effects in the sequential condition. Specifically, there was no reliable difference when the target appeared in the first display versus the second display, with a mean difference of 1.5% ± 1.1% (95% CI = −1.2, 4.2%, $t(5) = 1.42$, $p > 0.1$) in favor of better performance with the target in the second display. This lack of temporal effects supports the appropriateness of collapsing performance over the two sequential displays

The results of this experiment show that visual search for a particular global shape in a display of four shapes has fixed capacity. This contrasts with judgments of simple feature contrasts such as luminance and size, which have yielded evidence of unlimited capacity using

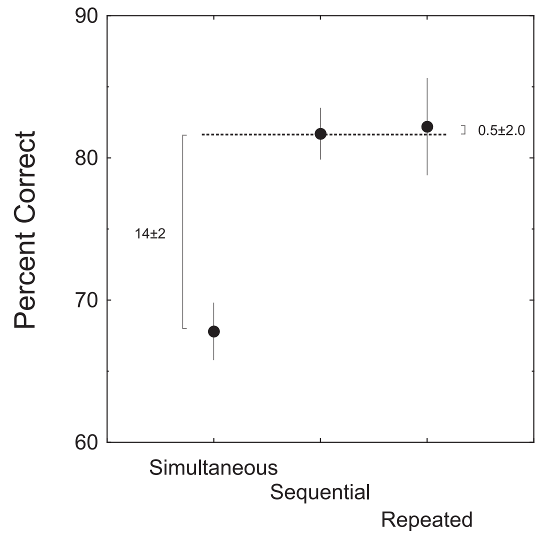

Figure 7. Results of Experiment 1 with a localization judgment of shape. Percent correct is shown for the three conditions. Chance is 25%. There is a sequential advantage for the sequential condition over the simultaneous condition. In addition, there is no reliable difference between the sequential and repeated conditions. This pattern of results is consistent with fixed-capacity parallel or serial models. Error bars are standard error of the mean.

this method (Huang & Pashler, 2005; Scharff et. al., 2011a). Instead, it is similar to the divided attention effects previously seen in other global shape judgments, animal categorization, and word categorization (Harris, Pashler, & Coburn, 2004; Scharff et. al., 2011a; Scharff et. al., 2011b; Scharff et. al., 2013). This experiment is mute with regard to whether the fixed-capacity limit observed here is due to an object-based or space-based processing limitation because the four different objects were presented in four different locations. The method must be adapted to address this further question.

Experiment 2: Shape judgments of objects at separate locations with a cued design

Experiment 2 was logically identical to Experiment 1, but the procedure was refined in anticipation of conducting a version using overlapping stimuli to discriminate between object-based and space-based processing limitations. Specifically, we conducted a version in which the stimuli were identical across all conditions. What differed was how many stimuli were cued as relevant for a given display. This cued design, first used in Scharff, et al. (2011b, Experiment 2), controls for potential differences in stimulus interactions, separate from processing limitations, which is an especially significant concern when overlapping stimuli are used. In addition, the task was changed from a localization task, which cannot be used in a single-location version of the experiment, to a target-present versus target-absent judgment.

Method

Observers

Six observers with normal or corrected-to-normal vision were volunteers. None had participated in Experiment 1. Some were paid in compensation for their time.

Design

Figure 8 illustrates the three conditions: cued simultaneous, cued sequential, and cued repeated (stimuli in the figure are not to scale). All conditions had the following sequence of events. A trial began with a study display (1 second) that defined the target, which was followed by a brief noise display (0.2 second), a blank display (0.5 second), the first cue display (0.5 second), another blank display (0.5 second), and then the first stimulus display (∼0.2 second), which contained four shapes: two on each side of fixation. This was followed by another blank display (0.5 second), the second cue display (0.5 second), a blank display (0.5 second), and finally the second stimulus display (∼0.2 second). Cues indicated which stimuli were relevant for the upcoming stimulus display. On each trial, there was a 50% chance that one of the four relevant stimuli was the target, all other stimuli were distractors. The trial ended with a probe display (not shown) that contained a reminder of the target and remained until response. Feedback was provided following errors, and the intertrial interval was 1.5 seconds. Performance was limited by superimposing each stimulus display with dynamic salt-and-pepper noise, which is represented in the figure by the shaded region. The purpose of the initial noise display near the beginning of the trial was to equate the contrast adaptation of the two stimulus displays (both were preceded by a noise display).

Conditions differed in which stimuli were cued as relevant. In the cued-simultaneous condition, cues indicated that the stimuli on both sides of either the first stimulus display (half the blocks) or the second stimulus display (half the blocks) and no stimuli were cued as relevant for the other stimulus display. In the example shown in the figure, this simultaneous cue was the two lines to either side of fixation in the first cue interval. In contrast, the second cue interval contained the fixation cross alone. This sequence of cues indicated

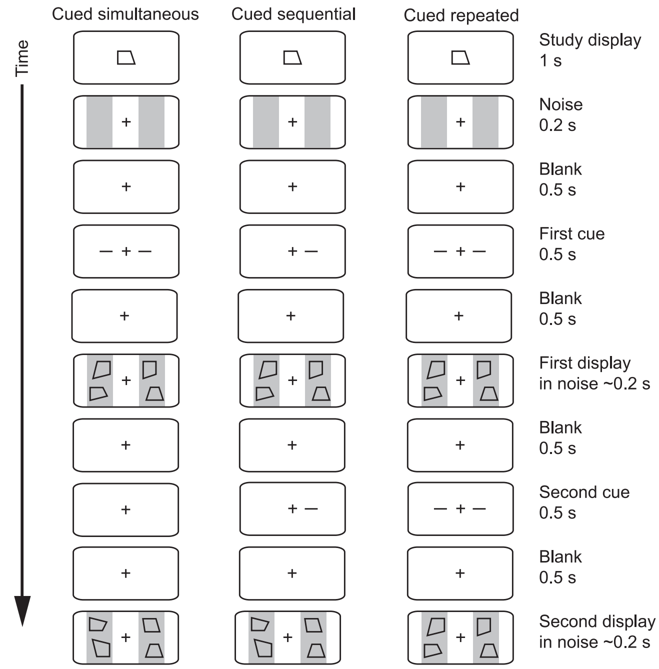

Figure 8. Illustration of the cued procedure used with shape judgments in Experiment 2. As before, each trial consisted of a study display followed by stimulus displays and a probe display (not shown). In each stimulus display, four objects were shown. The cue displays used lines to the left and right of fixation to indicate the relevant stimuli in the following display. In this example of the cued-simultaneous condition, the first cue had lines to the left and right indicating that objects on both the left and right were relevant in the first stimulus display. The second cue had no lines, indicating no objects were relevant in the second stimulus display. In the cued-sequential condition, the first and second cues were lines to the right, indicating the objects on the right were relevant in the first and second stimulus display. In the cued-repeated condition, both lines were present in both cue displays, indicating all objects were relevant in both stimulus displays. The presence of dynamic noise is indicated by shading.

that the first display was relevant to the task and the second display was irrelevant. For other blocks of trials, the second display was cued instead. Targets were presented in only the relevant displays. For the illustrated example, the target was rotated clockwise 90 degrees in the upper right of the first display.

In the cued-sequential condition, one side was cued for both stimulus displays. In the example, the right side was cued. Thus only the stimuli on the right side were relevant to the task. For other blocks of the cued-sequential condition, the left side was cued for the entire block. Targets only appeared on the cued side, and they appeared equally often in the first or second stimulus display.

In the cued-repeated condition, both sides of both displays were cued. The unique feature was that the stimuli in the first display were repeated for the second stimulus display. Specifically, the shapes were identical but the noise and jittered location of the shapes was varied from the first to the second display.

Stimuli

The stimuli were the same three sets of novel outline shapes as used in Experiment 1, but the noise used with the shapes to limit performance was different. Dynamic salt-and-pepper 1×1 pixel noise was applied during every frame of the stimulus display. Each pixel in the display had a probability of being replaced with a randomly chosen black or white pixel. This probability p was adjusted for each observer to maintain an intermediate level of performance (p ranged from 0.40 to 0.65 with an mean of 0.53). This high contrast, salt-and-pepper noise was an effective mask for these outline stimuli and also helped preserve the distinctive colors of the outlines, which was important for when we conducted an overlapping stimuli version (Experiment 4). The entire monitor had a displayable area of about $33° \times 24°$. Of this, the stimulus and noise fields occupied the central $14° \times 10°$. The noise fields were $6° \times 10°$ and were centered 4° to either side of fixation. This left a 1° space between the inner edge of the noise field and fixation. The squares were jittered by $\pm 0.5°$ around a mean location that was horizontally centered within the noise field and were centered vertically 2° above and below the horizontal meridian. An example display of four shapes and one frame of noise is shown in Panel B of Figure 3. Highly visible stimuli (low noise, long duration) were used for initial training. As training progressed, stimulus duration was adjusted for each observer to obtain performance of 80% to 90% correct as in Experiment 1 (range 0.1 to 0.5 second, mean of 0.24 second).

Apparatus

This experiment used a further updated apparatus relative to Experiment 1. The stimuli were displayed on the same flat-screen CRT monitor (19″ ViewSonic PF790) but now it was controlled by a Mac Mini (2.66 GHz Intel Core 2 Duo) using Mac OS X 10.6.8. The display still had a resolution of 832×624 pixels, a viewing distance of 60 cm (25.5 pixel/degree at screen center). But now it had a refresh rate of 120 Hz, a peak luminance of 104 cd/m^2, and a black level of 3.9 cd/m^2, mostly because of room illumination. Other details were the same as Experiment 1.

Unlike Experiment 1, eye position was recorded on all trials using an EyeLink II, 2.11 with 250 Hz sampling (SR Research, ON). The EyeLink II is a head-mounted binocular video system and was controlled by software using the EyeLink Developers Kit for the Mac 1.11.1 and the EyeLink Toolbox 3.0.11 (Cornelissen, Peters, & Palmer, 2002). The position of the right eye was recorded for all trials, and trials were included in the analysis only if fixation within a 2° window was confirmed. When fixation failed, five consecutive high frequency tones were sounded, and the trial was aborted. The percentage of aborted trials for each observer ranged from 2.2% to 7.7%, with an overall mean of 4.8% ± 0.9%. Thus the observers maintained fixation on almost all trials and none of the analyses included trials with blinks or saccades to the stimuli.

Procedure

Observers performed a modified yes-no task rather than the localization task of Experiment 1. They had to determine whether the target was present in a given trial using a rating scale with four possible responses: likely-no, guess-no, guess-yes, and likely-yes. The ratings allowed us to perform a receiver operating characteristic (ROC) analysis to control for bias. Specifically, we calculated the percent of the area under the ROC curve. This measure is an estimate of the unbiased percent correct. To encourage accuracy, there was no time limit on the responses. There were three main conditions but five kinds of blocks. This is because the cued-simultaneous condition had blocks with the first display relevant or the second display relevant. Similarly, the cued-sequential condition had blocks with the left side relevant and others with the right side relevant. Text instructions at the beginning of each block also specified which condition was to be presented (e.g., cued-sequential, left side). Each condition was presented in a block of 16 trials. A single experimental session consisted of six blocks (two from each main condition), for a total of 96 trials. After practice sessions, each observer ran 16 sessions resulting in a total of 1536 trials per observer.

Results and discussion

The results of Experiment 2 are shown in Figure 9. The percent area under the ROC is plotted against the three main conditions. Chance performance is 50% correct for this yes-no task. As with Experiment 1, there was a reliable sequential advantage of the cued-sequential condition over the cued-simultaneous condition: a mean difference of 4.9% ± 1.5% (95% CI = 1.2, 8.6%, $t(5) = 3.36$, $p = 0.02$). There was no reliable deviation of the sequential condition from the repeated condition, mean difference = 2.1% ± 2.1% (95% CI = −3.2, 7.4, $t(5) = 1.02$, $p > 0.1$). This pattern of results is consistent with predictions of a fixed-capacity parallel model or a simple serial model and inconsistent with predictions of an unlimited-capacity parallel model. One might wonder why the sequential advantage is

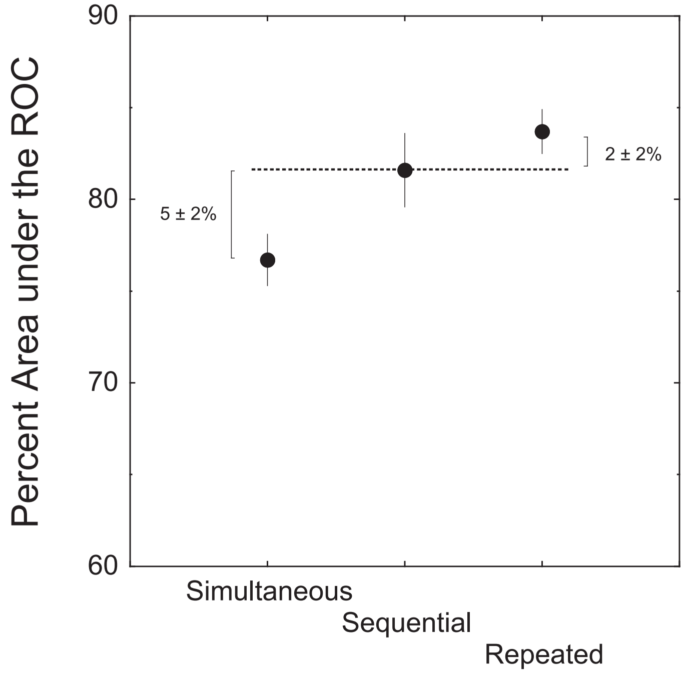

Figure 9. Results of Experiment 2 with a yes-no judgment of shape and a cueing method. Percent area under the ROC is shown for the three conditions. Chance is 50%. There is a sequential advantage for the sequential condition over the simultaneous condition. In addition, there is no reliable difference between the sequential and repeated conditions. This pattern of results replicates Experiment 1 and is consistent with fixed capacity.

only 5% here when it was 14% in Experiment 1. At least part of the reduction in the size of the effect is due to using a yes-no task (50% chance) instead of a four-choice localization task (25% chance). See Busey and Palmer (2008) for a detailed comparison of models for localization and yes-no.

To test for any temporal effects in the cued-sequential condition, we compared performance for the target in the first and second displays and found no reliable difference, mean difference = 0.7% ± 1.5% (95% CI = −3.3, 4.6%, $t(5) = 0.43$, $p > 0.1$). With this cued design, one can also compare performance for targets in the first and second display of the cued-simultaneous condition and again there was no reliable difference, mean difference = 2.3 ± 2.5% (95% CI = −4.0, 8.7%, $t(5) = 0.95$, $p > 0.1$). Thus no sign of temporal effects.

Experiment 3: Size judgments of cued stimuli

Before testing whether the capacity limitations observed in Experiments 1 and 2 reflect object-based or space-based limitations (or both), we sought to confirm that using the same general logic, feature-contrast judgments would reflect unlimited-capacity processing. As briefly reviewed in the introduction, the results from studies using feature-contrast tasks are inconsistent with both object-based and space-based theories of divided attention in their simplest forms. We sought to confirm this finding within the same context as that which we are using to test for object-based and space-based limitations. To this end, Experiment 3 used a size-judgment task in an experiment that was logically identical to Experiment 2. Size is considered a relatively simple feature and a good candidate for processing by the feature-contrast pathway. As such, we expected to find evidence of unlimited-capacity processing.

Method

Design

We used the cued version of the simultaneous-sequential method and kept most of the details the same as Experiment 2. Figure 10 illustrates the four conditions presented in separate blocks: cued simultaneous, cued sequential, cued repeated and cued all. Three of these conditions were the same as Experiment 2 but the cued-all condition was new. All conditions had the following sequence of events. A trial began with a study display (1 second) that defined the target followed by a brief noise display (0.2 second), a blank display (0.5 second), the first cue display (0.5 second), another blank display (0.5 second), and then the first stimulus display (0.2 second). The stimulus display contained four squares: two on each side of fixation in their own noise field. The location of each square was jittered by a small amount as described below in the Stimulus section. On each trial, there was a 50% chance that one of the squares was the target, otherwise they were all distractors. After that there was a similar sequence of blank and cue displays leading up to the second stimulus display (0.2 second). This was followed by a probe display (not shown) that contained a reminder of the target that remained until response. After the response there was feedback for errors and an intertrial interval of 1.5 seconds. Performance was limited by superimposing each stimulus display with dynamic salt-and-pepper noise which is represented by the shaded region.

In the cued-simultaneous condition, there were cues that indicated that the relevant stimuli were on both sides of the first display or both sides of the second display. In the example shown in the figure, this simultaneous cue was the two lines to either side of fixation in the first cue interval. In contrast, the second cue interval contained the fixation cross alone. This sequence of cues indicates that the first display was relevant to the task and the second display was irrelevant. For other blocks of trials, the second display

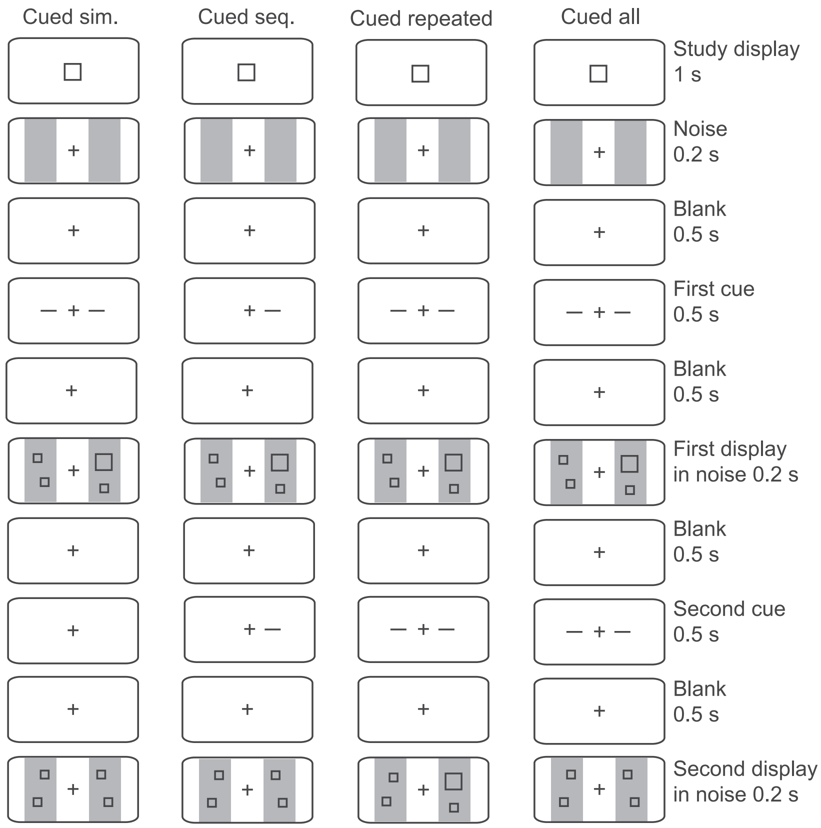

Figure 10. Illustration of the cued procedure used with size judgments in Experiment 3. As before, each trial consisted of a study display followed by two stimulus displays and a probe display (not shown). In each stimulus display, four objects were shown. The cue displays used lines to the left and right of fixation to indicate the relevant stimuli in the following display. In the cued-simultaneous condition, the first cue had lines to the left and right indicating that objects on both the left and right were relevant in the first stimulus display. The second cue had no lines, indicating no objects were relevant in the second stimulus display. In the cued-sequential condition, the first and second cues were lines to the right, indicating the objects on the right were relevant in the first and second stimulus display. In the cued-repeated condition, both lines were present in both cue displays, indicating all objects were relevant in both stimulus displays. In the cue-all condition, all eight stimuli were relevant. The presence of dynamic noise is indicated by shading.

was cued instead. Targets were presented only in the relevant displays. For the illustrated example, the target is the larger square in the upper right of the first display.

In the cued-sequential condition, one side was cued in both stimulus displays. In the example, the right side was cued. Thus only the stimuli on the right side were relevant to the task. For other blocks of trials, the left side was cued for the entire block.

In the cued-repeated condition, both sides of both displays were cued. The unique feature was that the stimuli in the first display were repeated for the second stimulus display. Specifically, the squares are identical but the noise and location of the squares was varied from the first to the second display. Performance in such redundant displays provided a benchmark to compare effects of sequential versus simultaneous displays.

We also included a new condition, the cued-all condition, which had cues for both sides of both displays, making all of the stimuli relevant. This provided a larger relevant set size for comparison to the main conditions. Finding a difference in performance as a function of relevant set-size would confirm that the cues were being used by the observers. This was a particular concern for this experiment because we expected little or no effects of divided attention.

Observers

Six observers were paid or unpaid volunteers with normal or corrected-to-normal vision. Five of them had also participated in Experiment 2.

Apparatus

The apparatus was the same as Experiment 2 and eye position was recorded in the same way. The percentage of aborted trials for each observer ranged from 2.4% to 6.3% with a mean of 4.2% ± 0.5%. Thus the observers maintained fixation on almost all trials and none of the analyses included trials with blinks or saccades to the stimuli.

Stimuli

In this experiment we turned to judgments of size instead of global shape. The stimuli were outline squares presented for 0.2 second. The distractor width was 1.5°, and the targets were incremented by variable amounts (e.g., 3 pixels which was about 0.12°) to achieve overall performance around 80% correct. In fact, all but one observer used an increment of 5 pixels which was about 0.2° (Observer A.M. used 3 pixels). They were presented as one-pixel thick, black outlines on a mid-gray background. As in Experiment 2, dynamic salt-and-pepper noise was applied during every frame of the stimulus display to limit performance. Each pixel in the display had a probability ($p = 0.65$) of being replaced with a randomly chosen black or white pixel. This high-contrast, salt-and-pepper noise was an effective mask for these outline stimuli and also helped preserve the distinctive colors of the outlines for the overlapping stimuli of later experiments. Other details of the stimuli were the same as Experiment 2. An example display of four squares and one frame of noise is shown in panel C of Figure 3.

Procedure

As in Experiment 2, observers performed the modified yes-no task that was used in Experiment 2. Each condition was presented in a block of 16 trials. A single experimental session consisted of eight blocks (two from each main condition), for a total of 128 trials. After practice sessions, each observer ran 16 sessions resulting in a total of 2048 trials per observer.

Results and discussion

The results of Experiment 3 are shown in Figure 11. The percent area under the ROC is plotted against the three main conditions. As with the previous experiment, chance performance is 50% correct for this yes-no task. Performance in the cued-sequential condition had no advantage over the cued-simultaneous condition, mean difference = −1.0% ± 1.5% (95% CI = −4.7, 2.8%, $t(5) = 0.65$, $p > 0.1$). Instead, there was a reliable difference between the cued-sequential condition and the cued-repeated condition of 6.4 ± 1.7% (95% CI = 2.0, 10.8%, $t(5) = 3.73$, $p = 0.01$). This pattern of results is consistent with the predictions of an unlimited-capacity model and is inconsistent with predictions of a fixed-capacity model.

To test for any temporal effects in the cued-sequential condition, we compared presenting the target in the first and second displays and found no reliable difference, mean difference = −1.3% ± 2.4% (95% CI = −7.4,4.7%, $t(5) = 0.56$, $p > 0.1$). The cued-simultaneous condition was also examined and showed no reliable

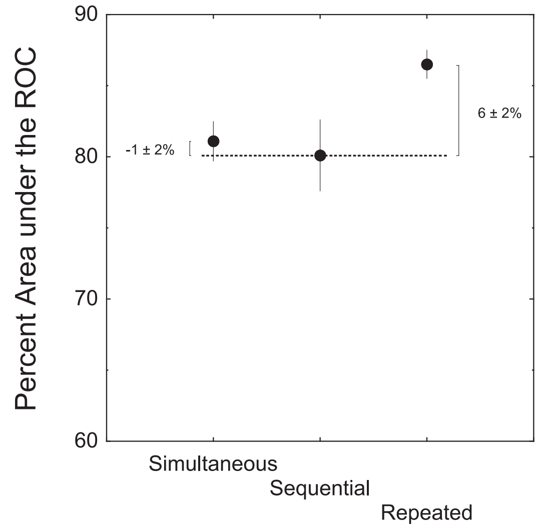

Figure 11. Results of Experiment 3 with a yes-no judgment of size and a cueing method. Percent area under the ROC is shown for the three main conditions. Chance is 50%. There is no reliable sequential advantage for the sequential condition over the simultaneous condition. In addition, there is instead a reliable difference between the sequential and repeated conditions. This pattern of results contrasts with the other experiments and is consistent with unlimited capacity.

difference, 2.0% ± 1.0% (95% CI = −0.5, 4.5%, $t(5) = 2.07$, $p > 0.05$).

Not shown in the figure is the additional condition in which all of the stimuli were cued. This cue-all condition was to test whether the observers were using the cues and in this condition the mean performance was 76.9% ± 2.0%. To test for the use of the cues, we calculated the difference in performance between the average of the two cued conditions (cued-simultaneous and cued-sequential) and the cue-all condition, which was reliable, mean difference = 3.7% ± 0.9% (95% CI = 1.5, 6.0%, $t(5) = 4.23$, $p = 0.008$). Thus the observers were able to use the cues to improve their performance. But despite this, they were not able to use the sequential condition to improve their performance over the simultaneous condition. This equivalence of the sequential and simultaneous conditions is the hallmark of unlimited-capacity processing.

Experiment 4: Shape judgments of overlapping stimuli

Experiments 1 and 2 showed evidence of fixed capacity for shape judgments of outline objects presented at multiple locations. While these results

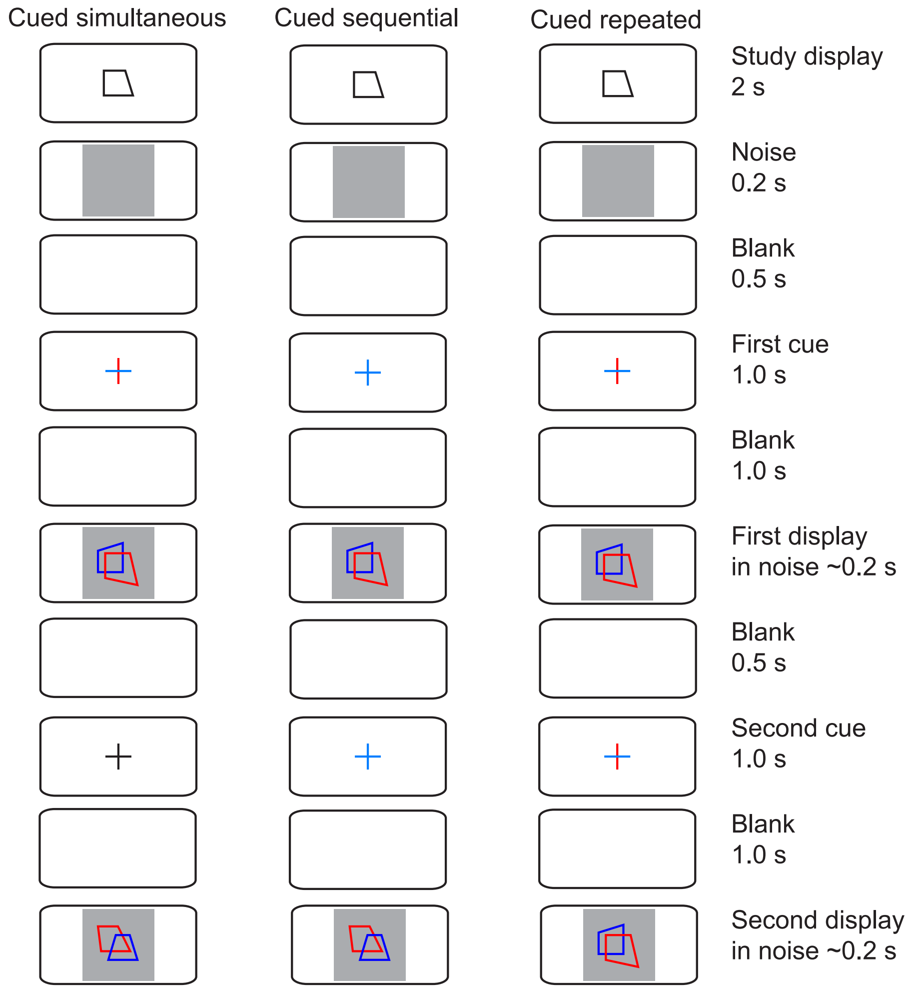

Figure 12. Illustration of the cued procedure used with overlapping objects in Experiment 4. As in previous experiments, each trial consisted of a study display followed by two stimulus displays and a probe display (not shown). In each stimulus display, there was a blue outline and a red outline object. They overlapped at a single location with a slight offset from one another. The cue displays used colored crosses that indicating the relevant stimuli in the following display. In this example of the cued-simultaneous condition, the first cue was blue and red, indicating that both objects were relevant in the first stimulus display. The second cue was black, indicating no objects were relevant in the second stimulus display. In the cued-sequential condition, the first and second cues was one color (e.g., blue), indicating the color of the relevant object in the first and second stimulus display. In the cued-repeated condition, both the first and second cues were red and blue, indicating both objects were relevant in both stimulus displays. The presence of dynamic noise is indicated by shading.

demonstrate a clear processing limitation, they are consistent with both object-based and space-based theories of divided attention because the limit could reflect having to process multiple objects or having to process stimuli at multiple locations. Experiment 4 used the overlapping-stimuli strategy to test between these two alternatives.

Method

Design

Experiment 4 used the cued version of the extended simultaneous-sequential method that was introduced in Experiment 2 so that stimuli were identical across conditions. This was especially important for this experiment because stimulus interactions were expected to be substantial for overlapping stimuli. Using the cued version of the method equated visibility across conditions. Figure 12 shows the three conditions: cued-simultaneous, cued-sequential, and cued-repeated. A trial began with a study display (2 seconds) followed by a brief noise mask (0.2 second) and a blank display (0.5 second). Then, the first cue (1.0 second) followed by a blank (0.5 second) and the first stimulus display (ranged from 0.15–0.25 second with a mean of 0.19 second), then a second pair of cue and stimulus displays. The cues were colored crosses at fixation that varied in color (red, blue, or black) to indicate the color of the relevant stimulus in the following display. The superimposed dynamic pixel noise is shown by the shaded regions.

In all conditions, two pairs of different colored overlapping objects were briefly displayed in the two stimulus displays. In the cued-simultaneous condition (left column), one of the two pairs was cued as relevant by the fixation cross changing to red and blue. A black fixation cross preceded the other pair indicating that it contained no relevant objects. In the example, the first pair is cued. In other blocks, the second pair was cued. Text instructions at the beginning of each block also specified which stimulus display (first or second) contained the relevant objects.

In the cued-sequential condition, a given color (blue in the example) was cued as relevant by the fixation cross changing to that color. Objects of that color were relevant in both displays, and the objects of the other color were irrelevant in both displays. Thus, only one object was relevant in each stimulus display. The relevant color was the same for all trials within the block and was also specified by text instructions at the beginning of the block. In the example, the blue cue indicated the blue outlines were relevant in both displays.

Finally, in the cued-repeated condition, the same pair of colored overlapping objects was briefly displayed in the two stimulus displays. A multicolored cue preceded each stimulus display, indicating that both objects were relevant, as in the simultaneous condition. The difference is that the repeated condition consists of two stimulus displays that contained the same objects.

Observers

Six observers participated in Experiment 4. Five of them also participated in Experiment 1. As in Experiment 1, this experiment was run using two apparatus, each used by three of the observers.

Apparatus

The same two apparatus as in Experiment 1. We did no eye tracking for this experiment with foveal stimuli.

Stimuli

The three sets of shapes from Experiments 1 and 2 were also used in Experiment 4. The fourth set from Experiment 1 was used by three of the observers in the early version of the experiment but not by the other observers. Color was used to help distinguish the overlapping stimuli. The stimuli were red or blue outlines with a 1-pixel thickness. To maximize saturation, the two colors were the primaries of the video monitor. Each stimulus display in the experiment consisted of a red-blue pair of shapes. The red and blue colors of the stimuli were matched in luminance using the maximum luminance available for the blue primary. Additionally, a slight horizontal and vertical offset of 5 pixels (~0.2°) was applied to one shape in an overlapping pair. As illustrated in the examples of the figure, this shift reduced the overlap of the outlines themselves. As in Experiments 1 and 2, the target shape was always a different rotation from the study and probe displays.

Following Experiments 2 and 3, dynamic salt-and-pepper pixel noise was applied during every frame to limit task performance. Each pixel in the display had a probability p of being replaced with a randomly chosen black or white pixel ($p = 0.65$ for three observers, $p = 0.75$ for the others). In this experiment, the noise was a single $4° \times 4°$ square combined with two overlapping shapes. An example pair of shape and one frame of noise are shown in panel D of Figure 3.

Procedure

We used the modified yes-no task that was used in Experiments 2 and 3. To maintain task difficulty, stimulus duration and noise probability was adjusted for each observer based on their performance. Duration ranged from 0.15 to 0.3 second with a mean of 0.19 second, and the noise probability ranged from 0.65 to 0.75 with a mean of 0.7. Each session consisted of 12 blocks for a total of 144 trials per session. Each observer ran 10 sessions for a total of 1440 trials per observer.

Results and discussion

The results of Experiment 4 are shown in Figure 13. Percent area under the ROC curve is plotted for the three main conditions: cued-simultaneous, cued-sequential, and cued-repeated. Chance performance for this yes-no task is 50%. There was a reliable sequential advantage for the cued-sequential condition over the cued-simultaneous condition, mean difference = 10.5% ± 1.8% (95% CI = 5.9%, 15.1%, $t(5) = 5.89$, $p = 0.002$). In contrast, there was no reliable difference between the cued-sequential and the cued-repeated conditions, mean difference = 0.7% ± 1.4% (95% CI = −2.9%, 4.3%, $t(5) = 0.47$, $p > 0.1$). As with Experiments 1 and 2, this pattern of results is consistent with a fixed-capacity model and is inconsistent with an unlimited-capacity parallel model.

There was no reliable difference for trials in which the target was in the first display or the second display in either the cued-sequential condition, mean difference = 0.2% ± 1.7% (95% CI = −4.3, 4.6%, $t(5) = 0.10$, $p > 0.1$), or the cued-simultaneous condition, mean

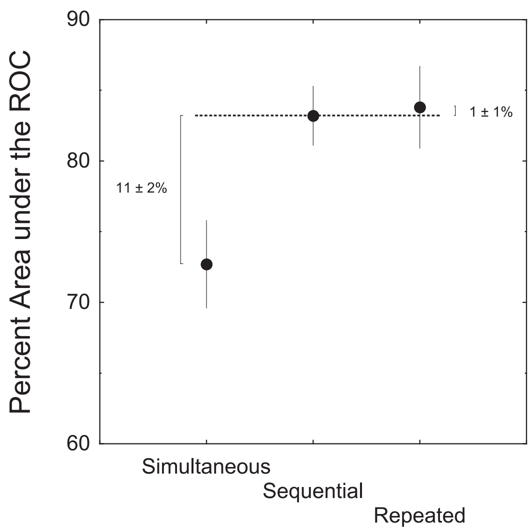

Figure 13. Results of Experiment 4 with a yes-no judgment of the shape of an overlapping object. Percent area under the ROC is shown for the three conditions. Chance is 50%. There is a sequential advantage for the sequential condition over the simultaneous condition. In addition, there is no reliable difference between the sequential and repeated conditions. This pattern of results generalizes prior results with shape judgments to overlapping displays. As with the other shape judgments, the pattern of results is consistent with fixed capacity.

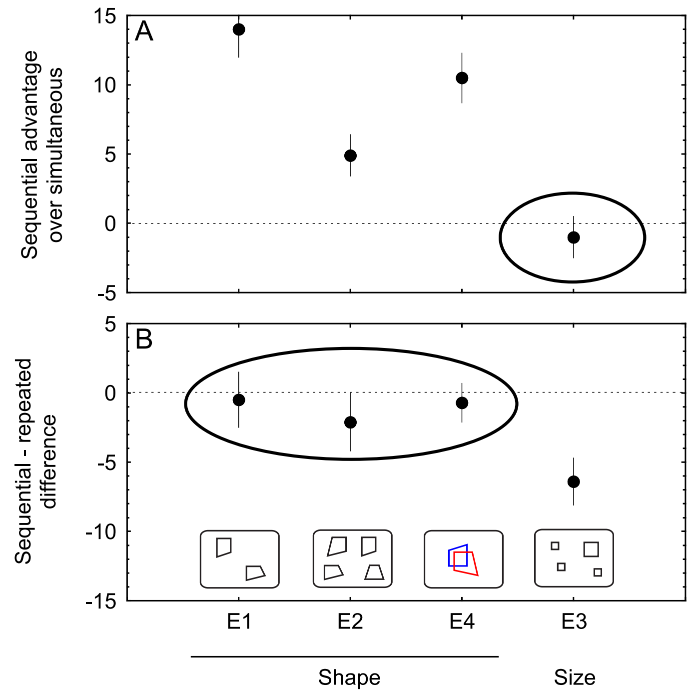

Figure 14. Summary of results from all experiments. Each experiment is summarized by two difference measures. (A) The sequential advantage (sequential-simultaneous) is plotted for all experiments. It is expected to be 0 for unlimited-capacity processing and positive for fixed-capacity processing. (B) The sequential-repeated difference (sequential-repeated) is shown for all experiments. It is expected to be 0 for fixed-capacity processing and negative for unlimited-capacity processing. The results for global shape judgments are consistent with fixed capacity (Experiments 1, 2, and 4), and the results for size judgment are consistent with unlimited capacity (Experiment 3). The points circled for emphasis are those that satisfy the equality predictions.

difference = 2.7% ± 3.2% (95% CI = −5.5%, 10.8%, $t(5) = 0.84$, $p > 0.1$). Thus we are confident that collapsing over the two intervals is appropriate.

The divided attention effects in this experiment cannot be attributed to location-based processing limitations. Observers never had to process stimuli from multiple locations in this experiment, and yet the results were again consistent with fixed-capacity processing. This is consistent with the object-based theories of divided attention.

General discussion

We used the extended simultaneous-sequential method to distinguish between object-based and space-based theories of divided attention. This method complements previous studies that used dual-task methods. In particular, we used a global shape judgment task, which was assumed to engage object processing, with non-overlapping (Experiments 1 and 2) and overlapping (Experiment 4) stimuli and found evidence of fixed-capacity processing in all cases. A feature contrast task, however, which was assumed to not require object processing but did require processing stimuli at multiple locations (Experiment 3), yielded evidence of unlimited-capacity processing. Together the results are consistent with object-based theories of divided attention, with a separate processing path that can be followed for tasks that depend on only feature contrast. There was no evidence of a spatial limitation per se.

Figure 14 summarizes the results using the logic of the extended simultaneous-sequential method (logic illustrated in Figure 6). Panel A shows the sequential advantage for the four experiments, and Panel B shows the sequential-repeated difference for the four experiments. First consider the results of Experiments 1 and 2, both of which investigated the capacity limitations of global shape judgments using stimuli at multiple locations. In both experiments, there was a reliable advantage of sequential over simultaneous presentation (Figure 14, Panel A), and no reliable difference between sequential and repeated presentations (Figure 14, Panel B). This is the signature pattern for fixed-capacity processing (see Figure 6, right panel). To address whether that fixed-capacity reflects object-based or space-based limitations, consider Experiment 4, which also used the global shape judgment task but with stimuli presented overlapping each other at a single location in space. There was again a reliable sequential advantage with no repeated advantage, consistent with the same fixed-capacity processing limitation that was observed in Experiments 1 and 2. Finally on the right side of the figure, consider the results from Experiment 3, which measured processing capacity limitations for a feature-contrast task. In this case, there was no reliable advantage for sequential presentation over simultaneous presentation (Figure 14, Panel A), but there was a reliable advantage for repeated displays (Figure 14, Panel B). This is the signature pattern for unlimited-capacity processing (Figure 6, left panel). Thus the results showed fixed capacity for object processing, regardless of the number of locations, but unlimited capacity for feature contrast, despite having to process stimuli at multiple locations.

One final comment about this summary figure. We emphasize the equality predictions for the unlimited-capacity and fixed-capacity models, because these predictions are robust to the details of the models, such as distributional assumptions. The results of the global shape experiments all satisfy the predicted equality between sequential and repeated conditions (see circled points in Panel B). In contrast the results of the size experiment satisfy the predicted equality between sequential and simultaneous conditions (see circled points in Panel A).

The current results converge with results from earlier studies that found support for object-based accounts of divided attention using dual-task methods (e.g., Duncan, 1984) and manipulations of perceptual organization of stimuli (e.g., Baylis, 1994). The use of the simultaneous-sequential method complements these earlier studies, and provides evidence of a fixed-capacity model of object processing, in particular.

Is there other evidence of space-based processing limitations?

Finding evidence of limited-capacity object processing, does not in itself rule out the possibility that there could also be space-based limitations. The overlapping-stimulus strategy of distinguishing between object-based and space-based divided attention effects does so by excluding the possibility of spatial effects, and using identical displays in the within- and between-objects conditions, thereby allowing the attribution of any effects that do occur to limitations on object processing. It is possible, however, that there are further processing limitations when objects must also be processed at multiple locations. We next consider evidence from two studies that addressed this possibility.

Vecera and Farah (1994) extended Duncan's (1984) method to test whether there are limitations associated with processing information from multiple locations in addition to processing information from multiple objects. They compared within-object versus between-object judgments when the stimuli were overlapping (i.e., in a single location) to when the objects were separated in space (i.e., multiple locations). They found no difference in the cost of processing information from multiple objects across the two conditions. These results are consistent with there being no additional limitation created by having to process information from two objects when they also must be processed from two different locations.

An analogous comparison using the same logic can be made within the current study. Specifically, Experiment 4, which used overlapping objects (i.e., one location), and Experiments 1 and 2, which used objects at multiple locations, all yielded evidence of fixed capacity. If multiple locations added an additional limitation, then our Experiments 1 and 2 should have had capacity limitations that were more extreme than our Experiment 4. For example, the multiple location experiments might have yielded evidence of fixed capacity and the overlapping object experiment evidence of more intermediate limits on capacity (see middle panel of Figure 6). This did not happen. Thus within our own study, there was no sign of an effect of space on divided attention.

A concern regarding the Vecera and Farah (1994) study is that the displays for the different location conditions (i.e., overlapping and separated) are not identical. Specifically, the overlapping objects (single location) condition presented stimuli at fixation, whereas the separated (multiple locations) condition presented stimuli off of fixation. The two conditions therefore differed in the eccentricity of the relevant stimuli (Kramer, Weber, & Watson, 1997). There is a similar issue for addressing this question in the current study, as well in that the same-location objects were presented at fixation (Experiment 4), whereas the different-location objects were presented off of fixation (Experiments 1 and 2).

Although we have emphasized that identical displays are ideal to control for non-attentional contributions to divided-attention effects, given the specifics of the results of the Vecera and Farah study (1994), and those of the current study, which are analogous to theirs, the differences in eccentricity might not be problematic. Consider the Vecera and Farah study. The eccentricity of stimuli in the separated condition was quite small (1.9°), and there was no reliable difference in overall accuracy between the overlapping objects (at fixation) and the separated objects (off of fixation) conditions, suggesting that at least by that measure, visibility was similar for those conditions. Similarly, in the current study, overall performance was tracked to similar levels of performance across the two conditions. More important, however, the critical result from both studies was that the object-level divided attention effect (i.e., between-object cost) was no greater for the separated condition than for the overlapping condition. It seems implausible that an effect of eccentricity would have exactly counteracted an additional divided attention cost. Although implausible, it is not impossible.

Kramer, Weber, and Watson (1997) sought to resolve the problem of differences in eccentricity by adapting the Vecera and Farah (1994) design. When the stimuli overlapped, they presented them to the left or right of fixation in one of the two locations in which the individual stimuli were presented in the separated condition, adding a place-holder stimulus in the opposite location to balance the display. In addition, they added a probe detection task (Experiment 1). On 25% of the trials, immediately following the stimulus display, a small red probe dot appeared in one of the two locations to the left or right of fixation. When this happened, the additional task was to press a button as quickly as possible before reporting the properties of the relevant object(s). Differences in probe reaction time (RT) provided a second way to assess attentional allocation to different locations.

With this adapted design, Kramer et al (1997) found that overall accuracy was lower when stimuli were superimposed than when they were separated, reflecting increased visual crowding due to stimuli being presented off of fixation (e.g., Levi, 2014). In addition, however, and contrary to Vecera and Farah (1994), the between-object cost was greater when stimuli were separated than when they were overlapping. There are two possible explanations for this finding. One is that the increased interference between stimuli made distinguishing between objects (e.g., segmenting them from each other) relatively difficult, thereby reducing the difference between the within- and between-object conditions in the overlapping stimuli condition. The second is that there is an additional cost to processing information from multiple locations beyond the cost of processing information from multiple objects. In a second experiment, Kramer et al. (1997) presented the superimposed stimuli at fixation and the separated stimuli 1.9° off of fixation as in the Vecera and Farah (1994) study. Under these conditions, they found no additional cost for separated stimuli than superimposed stimuli. It seems unlikely that there would be an additional cost for processing information from multiple locations when compared to superimposed stimuli at 1.9° eccentricity, but not compared to superimposed stimuli at fixation. Nonetheless, logically, the results are ambiguous.

Turning to the probe data, in Experiment 1 of Kramer et al. (1997), which was the experiment in which the divided attention effect was larger for stimuli at multiple locations than for stimuli at a single location, probe RT was faster when the probe appeared at the location in which relevant stimuli appeared compared to locations in which no relevant stimulus appeared (see also Kim & Cave, 1999; 2001; Weber, Kramer, & Miller, 1997). The authors concluded that selection is spatially mediated, and object representations consist of grouped arrays of locations (Vecera & Farah, 1994; Vecera, 1994; Vecera, 1997). When a grouped array is selected, stimuli that appear at or near the locations in that array will be selectively processed by virtue of being in the selected region. Notice that this is a conclusion about selective attention. It concerns the representational basis on which selection occurs.

The probe RT strategy used by Kramer et al (1997) and others (Kim & Cave, 1999; Kim & Cave, 2001) is similar to a number of older studies that were focused on measuring the spatial profile of attention by manipulating the spatial separation of individual stimuli in dual-task experiments (Hoffman & Nelson, 1981; Hoffman, Nelson & Houck, 1983; Sagi & Julesz, 1986). Those studies generally found reduced dual-task costs with smaller separations (but see Bahcall & Kowler, 1999 who found increased dual-task costs with smaller separations). Although they were not conceptualized as such at the time, these results are consistent with the grouped-array hypothesis favored by Kramer and colleagues. Processing is better for stimuli that are nearer to a selected grouped array than for stimuli that are farther. Again, though, this is a conclusion about selective attention.

Thus there are clearly spatial attention effects documented throughout the literature, but they are not effects that unambiguously imply a spatial limit of divided attention. The spatial effects in probe RT, dual-task, and other experiments can be due to grouping the probe with the other stimuli. Proximity of stimuli is a powerful organizing principle for the formation of object representations (e.g., Wagemans, Elder, Kubovy, Palmer, Peterson, Singh, von der Heydt, 2012). Stimuli that are presented nearer in space (and time) to other stimuli are more likely to be represented as part of the same grouped representation than are stimuli presented further in space and time. And stimuli that are grouped together are more likely to be selected together than are stimuli that are not grouped (e.g., Harms & Bundesen, 1983; Kahneman & Henik, 1981; Treisman, Kahneman, & Burkell, 1983; c.f. Moore, He, Zeng, & Mordkoff, 2021). Relative spatial proximity alone therefore is an ineffective manipulation for assessing whether a limitation is attributable to processing information from more than one object, more than one location, or both. Results are necessarily ambiguous.

What, in principle, would constitute evidence of a spatial limitation of divided attention? An ideal critical experiment would require reports of object properties, and compare conditions in which the number of relevant objects and the number of relevant locations were both manipulated, while using identical displays. We have emphasized the value of identical displays in asking about processing limitations in particular because any differences in displays across conditions raises the possibility of sensory-level interactions (e.g., visual crowding) affecting performance separate from any attentional effects. This was the basis of the criticism of the original Vecera and Farah (1994) study raised by Kramer and colleagues, and it applies to our study as well when we compare separate stimuli condition (Experiments 1 and 2) to our overlapping stimuli condition (Experiment 4). We have suggested that given the results from those studies (see also Experiment 2 of Kramer et al., 1997), the difference in displays might not be problematic. They would have all had to have exactly compensated for any underlying spatial divided attention effect, which seems implausible. Nonetheless, it is possible.

A study reported by Kim and Verghese (2014) is the closest that we are aware of to a critical experiment for isolating object-based and space-based divided attention effects. They found evidence of an object-based limitation but no evidence of a space-based limitation. In their study, observers monitored one or multiple locations along texture-defined surfaces for contrast increments. In other words, they manipulated spatial uncertainty. The surfaces were either single (unsegmented) surfaces or multiple (segmented) surfaces. Thus, although the locations were identical across conditions, they were on either one or multiple surfaces. The results showed that spatial uncertainty affected performance only when the locations were on separate surfaces. There was little or no spatial uncertainty effect for locations within a single surface. This is consistent with there being an object-based processing limitation but no spatial limitation. Although there was a small difference in displays (i.e., a small gap to define the segments in the segmented displays), this study seems to be the best test yet of possible spatial limitations above and beyond object limitations, and no evidence of a spatial limitation was found.

In summary, we argue that effects of spatial manipulations in most prior studies of divided attention are ambiguous results with regard to whether they reflect space-based or object-based processing limitations. Vercera and Farah (1994) introduced a method to distinguish these possibilities, and found evidence consistent with an object-based limit but no space-based limit. Our findings are analogous. Unfortunately, the displays used in these experiments were not matched across critical conditions, and follow-up experiments by Kramer et al. (1997) yielded similarly ambiguous results. A study by Kim and Verghese (2014) came closer to achieving matched-display conditions, and also found evidence of an object-based limitation but no space-based limitation. We suggest that the balance of evidence weighs in favor of there being a limit to the number of objects that can be processed but no spatial limit in and of itself. Importantly, this conclusion regarding divided attention says nothing about object-based versus space-based selective attention (i.e., the representational basis of selection). We return to that topic at the end of the General Discussion.

Multiple processing pathways

To understand the full set of results reported here and in the broader literature, an aspect of processing separate from objects versus space must be considered. Specifically, we suggest that some attributes of stimuli, such as feature contrast, do not in themselves constitute object properties, or otherwise engage limited-capacity processes, and can instead drive responses via a separate unlimited-capacity processing pathway. In Experiment 3 of the current study, for example, the task required only the detection of size contrast, and it yielded evidence of unlimited-capacity processing even though it required the processing of multiple stimuli at multiple locations. This is consistent with previous findings from dual-task studies (e.g., Han et al., 2003; Liu et al., 2009), simultaneous-successive studies (Huang & Pashler, 2005; Scharff et al., 2011a; Scharff et al., 2013), and many visual search studies in which performance was unaffected by the number of items in a spatial array (e.g, Treisman & Gelade, 1980). Because these experiments all involved multiple stimuli at multiple locations, it must be the case that the information that was needed from them was processed along a pathway separate from the limited-capacity object processing revealed in the current study and elsewhere.

Figure 15 illustrates an example of a multiple pathway theory that accommodates the full set of results. Relevant stimuli (three in the example, represented by S_1, S_2, and S_3) are processed initially in terms of simple feature information, which is assumed to have unlimited capacity. From this start, feature contrast is processed along an unlimited-capacity pathway, shown to the left. If the task depends on only feature contrast, then that pathway is sufficient to determine a response, and therefore there are no effects of divided attention. In Figure 15, this is labeled as the *direct* route from feature-contrast processes to decision and response processes. In contrast, properties that are intrinsic to objects, such as global shape, are processed along a limited-capacity pathway. If the task depends on one or more object property, then there

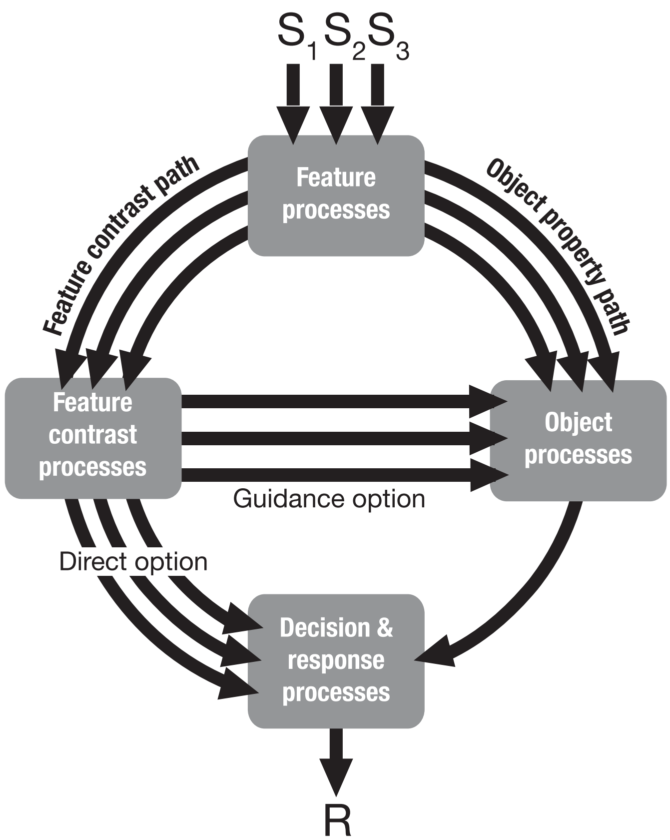

Figure 15. Illustration of a multiple pathway theory that includes separate paths for feature contrast and object properties. Relevant stimuli are initially processed in terms of their simple features and then, depending on the task, take one of two paths to reach decision and response processes. The arrows on the left represent an unlimited-capacity pathway that can be followed when the task depends only on a decision about feature contrast. The arrows on the right represent a limited-capacity pathway that must be followed if the task requires object processing.

will be effects of divided attention. According to this multiple pathway account, feature-contrast information could also serve to guide limited-capacity processes to selectively prioritize some stimuli over others for access to decision and response processes. Such guidance is a focus of specific existing theories such as Guided Search (Wolfe, 1994; Wolfe, 2021; Wolfe et al., 1989). Our data cannot distinguish between direct or guided paths of unlimited-capacity processing.

Our intention in creating this multiple pathway theory is to illustrate the minimum architecture needed to account for both fixed-capacity object-based processing and unlimited capacity processing of feature contrast. It is not intended as a comprehensive theory of visual search. It does not, for example, account for capacity limitations that are more intermediate than the fixed capacity we found evidence for here (see Scharff et al., 2013 Experiment 6 for an example). Such intermediate capacity limitations might depend on some combination of guidance and fixed-capacity processing (e.g., Wolfe, 1994; Wolfe, 2021; Wolfe et al., 1989), or they might dependent on architectural components not included here. Again, this simple multiple-pathway theory highlights the minimal set of components needed to account for the simultaneous fixed-capacity of object processing and unlimited capacity processing of feature contrast evidenced in the current study and elsewhere.

There are many precedents for the ideas behind the multiple pathway theory illustrated in Figure 15. Hoffman (1979), for example, proposed an early two-stage theory of visual search that included an unlimited-capacity parallel stage and a later serial stage. Feature integration theory is a classic elaboration of a two-stage theory of this sort. It assumed that the function of the early unlimited capacity stage is feature processing, whereas the function of the second stage is the binding of features to represent conjunctions via the serial allocation of attention to individual locations within the visual field (Treisman & Gelade, 1980). Kahneman et al. (1992) expanded this work to encompass a larger theory of perceptual organization and visual memory that asserts that that the establishment of object representations, in particular, depends on limited-capacity, perhaps even serial, processing. Wolfe and colleagues' original guided search theory also included an initial parallel unlimited-capacity processing stage for feature encoding and a later serial stage (Wolfe, 1994; Wolfe et al., 1989). A difference between guided search and other two-stage theories at the time is that according to Guided Search, the feature stage cannot be accessed directly, it can only contribute to the guidance of the limited-capacity second-stage processes. More recent versions of guided search, however, include a nonselective (i.e., direct) unlimited-capacity processing pathway based on global scene statistics and gist (e.g., Wolfe, Vo, Evans & Green, 2011; Wolfe, 2021). Finally, other recent two stage theories explicitly incorporate feature contrast, rather than just simple features, into the first stage (e.g., Becker, 2010; Müller, Heller, & Ziegler, 1995; Wolfe, 2021).

This brief review is not intended to cover all previous examples of multiple pathway theories. Rather it is to make clear that the general idea of a multiple pathway processing architecture is neither novel nor controversial. We highlight the general idea here because a theory of the sort illustrated in Figure 15 is sufficient to account for a large body of results from a wide range divided attention studies. This theory elaborates the conclusion that visual processing is limited by objects, but not by space, and that tasks that depend on feature contrast alone can bypass limited-capacity processing along a separate processing pathway.

What is object processing?

In making the general assertion that any task that depends on properties that are intrinsic to the object requires limited-capacity processing, we have not filled in the details regarding the specific processes that incur this dependency. Global shape seemed to be a safe choice for reliably engaging object processes, because it is clearly not available in the local features nor in any simple conjunction of features. Wolfe and Bennett (1997) made a similar point when contrasting global shape with the kinds of simple feature conjunctions that were a focus of the visual search literature at the time (see also Pizlo, 2008).

How might a task that depends on feature conjunctions be processed within an architecture like that illustrated in Figure 15? In many cases, because conjunctions depend on more than feature contrast, they require processing along the limited-capacity pathway consistent with Feature Integration Theory and other theories. Under the right conditions, however, tasks that depend on feature conjunctions could create feature-contrast signals that can be detected by the unlimited-capacity processing path. These can be thought of as emergent features (Pomerantz & Portillo, 2011) . Results from a study using the simultaneous-sequential method to measure divided attention effects support this possibility. Specifically, Huang and Pashler (2005) found that search for targets defined by the conjunction of two features (size and orientation) yielded evidence of unlimited-capacity processing in simultaneous-sequential experiments.

More generally, our multiple pathway theory does not provide specific predictions for feature conjunction tasks because, depending on the specific conditions, conjunction information might be extractable through the unlimited-capacity feature-contrast pathway, but other times not. Rosenholtz (2001), for example, showed that performance in a range of visual search tasks was best accounted for by the degree of heterogeneity of the displays, rather than whether the target was defined on the basis of a single feature or a conjunction of features (see also Duncan & Humphreys, 1989). Increased heterogeneity reduces the usefulness of feature contrast, and increases the likelihood that a given task requires processing along the limited-capacity pathway.

Work by Dosher and Lu (2009) provides a hint about what kind of processing other than global shape representation might constitute "object processing" (Han et al., 2003; Liu et al., 2009). Recall that using a dual-task approach, they found no divided attention effects for multiple judgments of the same feature across two objects (Gabor patches). However, they did find divided attention effects for judgments of two different features across two objects. The first finding is consistent with the feature task being done based on processing only within the unlimited-capacity feature-contrast pathway. The second finding is consistent with the hypothesis that associating a particular feature with a particular object and not another object, cannot be done using feature contrast alone. Instead, it requires object representations with which to associate some specific features and not associate other specific features that are simultaneously present. This hypothesis is similar to the need to bind features at a given location in Feature Integration Theory, but is more generally about associating specific information with specific object representations, or individuation (c.f., Kahneman et al., 1992).

Object-based versus space-based *selective* attention

The focus of this study has been on divided attention. Specifically, we have asked on what basis is processing limited. Is it limited by the number of objects that can be processed simultaneously, the number of locations from which information can be processed simultaneously, or both? A large component of the literature on object-based attention, however, is concerned with selective attention, which addresses the question of what representational basis is information selected: spatial locations, objects, or both. The dominant view within the literature is that selection can be both object based and space based, and that object representations (Kramer et al., 1997; Vecera & Farah, 1994; Vecera, 1994), constrain the spatial profile of selection (e.g., Avrahami, 1999; Egly, Driver, & Rafal, 1994; Kramer et al., 1997; Hollingworth, Maxcey-Richard, & Vecera, 2012; Moore, Yantis, & Vaughan, 1998; Vecera, 1994; Vecera, 1997). We end our discussion by highlighting some examples of studies that addressed the question of object-based and space-based *selection* using designs with identical displays across conditions, similar to those we have emphasized for divided attention.

First as noted in the introduction, before Duncan (1984) used overlapping stimuli to study object-based divided attention, Rock and Gutman (1981) used overlapping stimuli to study selective attention (see also Neisser & Becklen, 1975). With overlapping shape stimuli similar to those used in the current study, they found that observers were able to selectively process one of the two overlapping stimuli. These findings constitute evidence that selection can be object based, but they do not speak to the question of whether selection can also be space based.

A challenge to measuring separate effects of object-based and space-based selective attention, just as with divided attention, is creating designs in which the critical conditions use identical displays so that non-attentional contributions can be ruled out. Chen (1998) was among first to highlight this need with regard to object-based selective attention. She

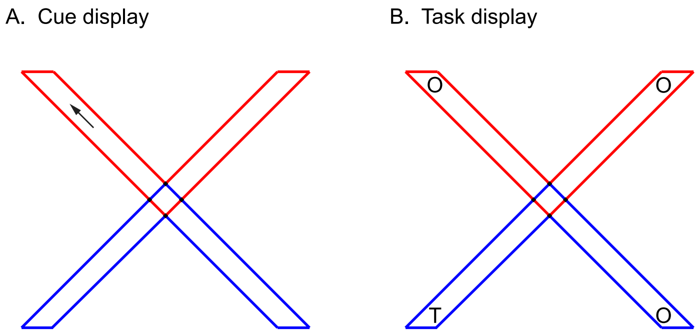

Figure 16. Chen (1998) used displays that could be perceptually organized as two objects (Vs) or one object (X) to dissociate space-based and object-based selective attention effects with identical displays across critical conditions. See text for details.

noted that previous studies that offered evidence of object-based selection (e.g., Egly, Driver & Rafal, 1994; Harms & Bundeson, 1983; Kahneman & Henik, 1981; Kim & Cave, 1999; Moore, Yantis, & Vaughan, 1998) had not used identical displays across same- and different-object conditions and therefore were open to alternative interpretations.

Chen (1998) addressed the concern of non-identical displays by adapting the logic of the two-rectangles method of Egly et al. (1994), which has the feature of measuring both a space-based effect and an object-based effect within the same experiment. The innovation was to manipulate observers' perceptual organization of displays to alter object structure, rather than changing anything about the display. This allowed the number of represented objects to differ across conditions without changing the stimulus itself (c.f., Baylis, 1994; Baylis & Driver, 1993; Chen, 2000). Displays consisted of two Vs, one red and one blue, configured so that they formed a large X (see Figure 16). These displays could be perceptually organized as two Vs (two objects) or one X (one object). Trials began with the end of one V being cued (Panel A of Figure 16), followed by letters at all four ends (Panel B of Figure 16). The task was to find a target T or L among distractor O's and to report which it was. When perceived as two Vs, the standard set of conditions from the two-rectangles method can be defined. Specifically, the target can appear in the cued location (Valid), on the other side of the cued object (Invalid Same Object) or at an equidistant location in the uncued object (Invalid Different Object). ~~In the figure the Invalid labels are written in quotes, because when the displays are organized as a single X, there is no distinction between the two invalid conditions.~~ When observers were told nothing about the displays, there was an effect of both distance from the cue and whether the target appeared in the cued object or not, just as has been found in many studies using the two rectangles method. Specifically, responses were fastest in the valid condition, next fastest in the invalid same-object condition, and slowest in the invalid different-object condition. This is consistent with observers having perceived the displays as two V's. In a critical second experiment, however, observers were told that the stimuli would appear in "an outlined capital X, whose colour could be either red, blue, or a mixture of red and blue." This encouraged them to perceptually organize the displays into a single object, an X, rather than two Vs. In this case, responses were fastest in the valid condition and slower in the invalid conditions, indicating a spatial effect. There was, however, no difference in response time between the two invalid conditions. These results confirm that the same-object versus different-object effect in the first experiment reflected a difference in perceived object structure, rather than stimulus level differences, and that selection can be both object based and space based (see also Zheng & Moore, 2021).

How can it be, as we are suggesting, that divided attention is limited by objects but not space, whereas selective attention is both object based and space based? A general answer to this question is that selection and processing limitations are not yoked, as the spotlight (and related) metaphors imply. The function of selective attention is to control the flow of information processing to compensate for the fact that our senses receive more information than can be processed (e.g., Broadbent, 1958). Given this understanding, two broad questions can be asked: (1) What specific aspects of processing are limited (divided attention)? and (2) On what representational basis can selection occur (selective attention)? Early models of attention focused on the question of when within the stream of processing selection occurs, and as a consequence yoked the two questions. Early models, for example, asserted that stimulus identification processes were limited, and therefore that selection was based on information that could be extracted before stimulus identification such as location. Late models asserted that only processes after stimulus identification were limited, and therefore selection was based on representations of stimulus identity. More recent views maintain that selection occurs at multiple processing loci (e.g., "early" and "late"). By the same token, processing limitations and selection need not be yoked. Consider the simple multiple-pathway theory of divided attention that we proposed in Figure 15. That theory asserts that there are limitations associated with object processing, but not feature contrasts. Where selection occurs is left unstated. It could occur "early," such that stimuli in a selected image location is prioritized over information in other locations (space-based selection), but it could also occur late at the decision stage, such that the input from one batch of object processing is prioritized over that of a different batch of object processing. The point is that theories about what aspects of processing are limited (divided attention) are separable from theories about how (when) selection occurs. They are relevant to each other, to be sure, but they need not be yoked as typical metaphors imply. A more extensive discussion about the relationship between divided attention and selective attention is beyond the scope of the current article and would no doubt overlap with existing discussions for the limitations of the early versus late selection debate. In closing, we re-emphasize that the focus of the current study is divided attention.

Conclusion

This study addressed the distinction between object-based and space-based theories of divided attention. A review of the literature indicated that collectively, the extant evidence was inconsistent with both simple space-based theories (e.g., Posner, 1980) and simple object-based theories (e.g., Duncan, 1984), as well as single-pathway hybrid theories (e.g., Vecera, 1997). Using the extended simultaneous-sequential method (Scharf et al., 2011a), we found evidence that is consistent with fixed-capacity object processing, but no evidence of a limitation due to spatial processing itself. Furthermore, we found evidence of unlimited-capacity processing of feature contrast across both objects and locations. Together the results are consistent with a multiple pathway theory in which tasks that require object processing must follow a fixed-capacity pathway, whereas tasks that depend on only feature contrast can follow an unlimited-capacity processing pathway. In neither pathway is there a processing limitation because of space.

Keywords: divided attention, object-based attention, space-based attention, overlapping objects

Acknowledgments

The authors thank Dina V. Popovkina and Alec Scharff for helpful comments and thoughtful criticism. This article was partially supported by grant R21 EY029432 to C. M. Moore. The data for this study can be found at the OSF depository: https://osf.io/fn6x3/.

Commercial relationships: none.
Corresponding author: Cathleen M. Moore.
Email: cathleen-moore@uiowa.edu.
Address: University of Iowa, G60 Psychological and Brain Sciences Building, 340 Iowa Avenue, Iowa City, IA 52242-1407, USA.

Footnote

A fourth set of quadrilaterals with one curved side (not shown) was included for three of the observers, but was dropped for the other observers due to being more difficult to distinguish than the other stimulus sets.

References

Avrahami, N. (1999). Objects of attention, objects of perception. *Perception & Psychophysics, 61*, 1604–1612.

Bahcall, D. O., & Kowler, E. (1999). Attentional interference at small spatial separations. *Vision Research, 39*, 71–86.

Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 451–470.

Baylis, G. C. (1994). Visual attention and objects: Two-object cost with equal convexity. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 208–212.

Becker, S. I. (2010). The role of target–distractor relationships in guiding attention and the eyes in visual search. *Journal of Experimental Psychology: General, 139*, 247–265.

Bonnel, A. M., Stein, J. F., & Bertucci, P. (1992). Does attention modulate the perception of luminance changes? *Quarterly Journal of Experimental Psychology, 44A*, 601–626.

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

Broadbent, D. E. (1958). *Perception and communication*. New York: Pergamon Press.

Busey, T., & Palmer, J. (2008). Set-size effects for identification versus localization depend on the visual search task. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 790–810.

Chen, Z. (1998). Switching attention within and between objects: The role of subjective organization. *Canadian Journal of Experimental Psychology–Revue Canadienne De Psychologie Experimentale, 52*, 7–17.

Chen, Z. (2000). An object-based cost of visual filtering, *Perception & Psychophysics, 62*, 482–495.

Cornelissen, F., Peters, E. M., & Palmer, J. (2002). The eyelink toolbox: Eye tracking with MATLAB and the psychophysics toolbox. *Behavior Research Methods, Instruments, & Computers, 34*, 613–617.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*, 501–517.

Duncan, J., & Humphreys, G.W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*, 433–458.

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual-attention between objects and locations: Evidence from normal and parietal lesion subjects.

Journal of Experimental Psychology: General, 123, 161–177.

Ernst, Z. R., Palmer, J., & Boynton, G. M. (2012). Dividing attention between two transparent motion surfaces results in a failure of selective attention. *Journal of Vision, 12*, 1–17.

Graham, N., Kramer, P., & Haber, N. (1985). Attending to the spatial frequency and the spatial position of near-threshold visual patterns. In M. I. Posner, & O. S. M. Marin (Eds.), *Attention and Performance XI*. Hillsdale, NJ: Erlbaum.

Han, S., Dosher, B. A., & Lu, Z. L. (2003). Object attention revisited: Identifying mechanisms and boundary conditions. *Psychological Science, 14*, 598–604.

Harms, L., & Bundesen, C. (1983). Color segregation and selective attention in a nonsearch task. *Perception & Psychophysics, 33*, 11–19.

Harris, C. R., Pashler, H. E., & Coburn, P. (2004). Moray revisited: High-priority affective stimuli and visual search. *Quarterly Journal of Experimental Psychology Section A, 57*, 1–31.

Hoffman, J. E. (1979). A two-stage model of visual search. *Perception & Psychophysics, 25*, 319–327.

Hoffman, J. E., & Nelson, B. (1981). Spatial selectivity in visual search. *Perception & Psychophysics, 30*, 283–290.

Hoffman, J. E., Nelson, B., & Houck, M. R. (1983). The role of attentional resources in automatic detection. *Cognitive Psychology, 51*, 379–410.

Hollingworth, A., Maxcey-Richard, A. M., & Vecera, S. (2012). The spatial distribution of attention within and across objects. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 135–151.

Huang, L., & Pashler, H. (2005). Attention capacity and task difficulty in visual search. *Cognition, 94*, B101–B111.

Kahneman, D., & Henik, A. (1981). Perceptual orgnziation and attention. In M. Kubovy, & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 181–211). Hillsdale, NJ: Erlbaum.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219.

Kim, Y. J., & Verghese, P. (2014) The influence of segmentation and uncertainty on target selection. *Journal of Vision, 14*(3):3, 1–11.

Kim, M.-S., & Cave, K. R. (1999). Grouping effects on spatial attention in visual search. *Journal of General Psychology, 126*, 326–352.

Kim, M.-S., & Cave, K. R. (2001). Perceptual grouping via spatial selection in a focused-attention task. *Vision Research, 41*, 611–624.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception, 36*: ECVP Abstract Supplement.

Kramer, A. F., Weber, T. A., & Watson, S. E. (1997). Object-based attentional selection—Grouped arrays or spatially invariant representations? Comment of Vecera and Farah (1994). *Journal of Experimental Psychology: General, 126*, 3–13.

Levi, D. (2014). Visual crowding. In J.S. Werner, & L. M. Chalupa (Eds), *New Visual Neurosciences* (pp 681–694), Cambridge, MA: MIT Press.

Liu, S. H., Dosher, B. A., & Lu, Z. L. (2009). The role of judgment frames and task precision in object attention: Reduced template sharpness limits dual-object performance. *Vision Research, 49*, 1336–1351.

Moore, C. M., He, S., Zheng, Q., & Mordkoff, J. T. (2021). Target-flanker similarity effects reflect image segmentation not perceptual grouping. *Attention, Perception, & Psychophysics, 83*, 658–675.

Moore, C. M., Yantis, S., & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological Science, 9*, 104–110.

Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics, 57*, 1–17.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology, 7*, 480–494.

Nothdurft, H. C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research, 31*, 1073–1078.

Nothdurft, H. C. (1993). The role of features in preattentive vision: comparison of orientation, motion and color cues. *Vision Research, 33*, 1937–1958.

Nothdurft, H. C. (2000). Salience from feature contrast: variations with texture density. *Vision Research, 40*, 3181–3200.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40*, 1227–1268.

Pizlo, Z. (2008). *3D Shape: Its unique place in visual perception*. Cambridge, MA: MIT Press.

Popovkina, D. V., Palmer, J., Moore, C. M., & Boynton, G. M. (2021). Is there a serial bottleneck in visual object recognition? *Journal of Vision, 21*, 1–21.

Pomerantz, J. R., & Portillo, M. C. (2011). Grouping and emergent features in vision: Toward a theory of basic gestalts. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 1331–1349.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and detection of signals. *Journal of Experimental Psychology, 109*, 160–74.

Rock, I., & Gutman, D. (1981). The effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 275–285.

Rosenholtz, R. (2001). Visual search for orientation among heterogeneous distractors: Experimental results and implications for signal-detection theory models of search. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 985–999.

Sagi, D., & Julesz, B. (1986). Enhanced detection in the aperture of focal attention during simple discrimination tasks. *Nature, 321*, 693–695.

Scharff, A., Palmer, J., & Moore, C. M. (2011a). Extending the simultaneous-sequential paradigm to measure perceptual capacity for features and words. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 813–833.

Scharff, A., Palmer, J., & Moore, C. M. (2011b). Evidence of fixed capacity in visual object categorization. *Psychonomic Bulletin & Review, 18*, 713–721.

Scharff, A., Palmer, J., & Moore, C. M. (2013). Divided attention limits perception of 3-D object shapes. *Journal of Vision, 13*:18, 1–24.

Shaw, M. L. (1980). Identifying attentional and decision-making components in information processing. In R. Nickerson (Ed.), *Attention and Performance VIII* (pp. 277–296). Hillsdale, NJ: Erlbaum.

Shiffrin, R. M., & Gardner, G. T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology, 93*, 72–82.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Treisman, A. M., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the cost of filtering. *Perception & Psychophysics*, *33*, 527–532.

Valdes-Sosa, M., Cobo, A., & Pinilla, T. (1998). Transparent motion and object-based attention. *Cognition, 66*, B13–B23.

Valdes-Sosa, M., Cobo, A., & Pinilla, T. (2000). Attention to object files defined by transparent motion. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 488–505.

Vecera, S. P. (1997). Grouped arrays versus object-based representations: Reply to Kramer et al. *Journal of Experimental Psychology: General, 126*, 14–18.

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General, 123*, 146–160.

Vecera, S. P. (1994). Grouped locations and object-based attention: Comment on Egly, Driver, and Rafal (1994).

Wagemans, J.,, Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., . . . von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: Perceptual grouping and figure-ground organization. *Psychological Bulletin, 138*, 1172–1217.

Weber, T. A., Kramer, A. F., & Miller, G. A. (1997). Selective processing of superimposed objects: An electrophysiological analysis of object-based attentional selection. *Biological Psychology, 45*, 159–182.

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1*, 202–238.

Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin and Review, 28*, 1060–1092.

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

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 and Performance, 15*, 419–433.

Wolfe, J. M., Võ, M. L. H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences, 15*, 77–84.

Zheng, Q., & Moore, C. M. (2021). Task-specific engagement of object-based and spaced-based attention with spatiotemporally defined objects. *Attention, Perception, & Psychophysics, 83*, 1479–1490.

Errata for Moore, Pai and Palmer (2022)
Published in Journal of Vision, 22, 1-25
John Palmer
Last revised 5 November 2022

1. Page 22, left column, near the bottom. The following sentence is a leftover from an earlier version of the figure. It should've been deleted.

"In the figure the Invalid labels are written in quotes, because when the displays are organized as a single X, there is no distinction between the two invalid conditions."