in visual search

Divided attention effects are caused by objects not by space

Cathleen M. Moore¹
James Pai²
John Palmer²

¹University of Iowa
²University of Washington

Intended for: *Journal of Vision*

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

Short title: Divided attention is limited by objects

Abstract Word count: 349 (must be less or equal to 250 words)

Word count excluding title, references, author note, figure captions: ~12,900 (for JEP:G)

Last revised: 19 July 2021
Abstract

Will rewrite later…
Studies of visual perception have shown that for some tasks, increasing the number of stimuli that must be processed reduces performance. These divided attention effects have been explained in terms of the amount of information that can be processed per unit time (or trial), or in other words, the capacity of information processing (Broadbent, 1958). Using this terminology, conditions that yield no divided attention effects indicate entirely unlimited-capacity processing, whereas conditions that yield divided attention effects indicate limited-capacity processing of some kind. In some cases, conditions can yield divided attention effects that are consistent with processing being limited to a constant amount of information being processed across stimuli within a given unit of time. This is referred to as fixed-capacity processing (Popovkina, Palmer, Moore & Boynton, 2021; Scharff, Palmer & Moore, 2011a; Shaw, 1980).

Object-Based versus Space-Based Divided Attention

According to object-based theories of divided attention, capacity limits are imposed by having to process information from multiple objects rather than a single object (e.g., Duncan, 1984). These theories predict that it is easier to process multiple features of a single object than multiple features of different objects. Pure object-based theories predict that processing multiple features of a single object will yield no cost relative to processing a single feature from a single object (i.e., no divided attention effect), whereas processing multiple features from multiple objects will yield a significant cost compared to processing a single feature from a single object (i.e., a divided attention effect). In contrast, according to space-based theories of divided attention, capacity limits are imposed by having to process information from multiple locations rather than from a single location (e.g., Hoffman & Nelson, 1981; Posner, 1980). These theories predict that processing multiple stimuli at a single location will be easier than
processing the same stimuli across multiple locations. A metaphor that is commonly used to illustrate space-based theories is a single spotlight that can point to only one location at a time. The spotlight illuminates a region of space without regard to what is in it. For example, imagine an actor, part of the stage floor, one end of a table, and part of the curtain behind the actor all being illuminated by a focused spotlight. To illuminate multiple locations using that single spotlight, it must be shifted from one location to another, limiting the number of locations that can be illuminated within a given period of time. A comparable metaphor for object-based theories would be if an actor were highlighted because he/she is wearing a glow-in-the-dark hat. To highlight multiple actors with that hat, it must be passed from one actor to another, limiting the number of actors that can be illuminated within a given period of time.

Object-based and space-based theories of divided attention propose distinctly different sources of capacity limitation. However, it is challenging to discriminate between them because the number of locations is often confounded with the number of objects. Two stimuli that yield two object representations, for example, usually appear in two different locations as well. In this case, the two theories make the same predictions and therefore cannot be distinguished. There have been several approaches to addressing this problem. One has been to equate the spatial extent of relevant stimuli across single- and multiple-object tasks (e.g., Baylis & Driver, 1993; Lavie & Driver, 1996; Watson & Kramer, 1999). Another has been to use motion to present a small number of objects at many different locations over time (e.g., Kahneman, Treisman, & Gibbs, 1992; Pylyshyn & Storm, 1988), thereby dissociating the number of objects (few) from the number of locations (many). Finally, another approach has been to present multiple objects at a single location by having them overlap in the image (e.g., Duncan, 1984; Rock and Gutman, 1981), thereby dissociating the number of objects (multiple) from the number of locations (one).
The current study sought to test whether divided attention effects are caused by having to select multiple objects, multiple locations, or both. As will be reviewed, without additional considerations, the evidence to date is inconsistent with all three possibilities. Because we used the overlapping-objects approach to distinguishing between object-based and space-based divided attention effects, we offer a brief review of its development and application.

Rock and Gutman (1981; see also Neisser & Becklen, 1975) used displays of overlapping stimuli to distinguish between object-based and space-based selective attention (see Fig 1a). Specifically, they asked whether observers could parse the stimuli and selectively process information from one of the two overlapping stimuli, and they found that they could. Duncan (1984) extended this strategy to distinguish between object-based and space-based divided attention. He used displays with an outline rectangle and an overlapping tilted line (see Fig 1b), and asked whether there would be a cost for having to report information from two objects compared to one, even though the two objects were at a single location. 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 strategy. 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); that is, there was a divided attention effect for between-object judgments (limited capacity). However, dual-task performance was no worse than single-task performance when the two judgments were about a single object (e.g., side of gap and size of box); that is there was no
divided attention effect for within-object judgments (unlimited capacity). Thus, in Duncan's study, performance depended on the number of relevant objects even when the number of locations was held constant at one, which is consistent with object-based theories of divided attention and inconsistent with strong space-based theories of divided attention (i.e., ones that posit there are no other sources of capacity limitation).

Another strategy for presenting multiple objects at a single location, which was introduced by Valdes-Sosa and colleagues (Valdes-Sosa, Bobes, Rodriguez, & Pinilla, 1998; Valdes-Sosa, Cobo, & Pinilla, 1998, 2000), is to create displays of multiple overlapping transparent surfaces. In a variation of this approach, Ernst, Palmer, and Boynton (2012) used a dual-task paradigm with two overlapping transparent surfaces that were defined by random-dot kinematograms that appeared as two transparent surfaces moving past each other and extending across the same region of visual space (i.e., same location). Observers detected changes in motion and/or changes in luminance from the same surface or different surfaces. The results were consistent with object-based theories of divided attention in that judgments between two surfaces showed a divided attention effect, whereas judgments within a surface showed little or no divided attention effect.

The results of the studies reviewed so far provided evidence consistent with object-based theories of divided attention. Specifically, they have shown costs associated with processing information across multiple objects, even when controlling for location. Other studies, however, have yielded evidence of unlimited capacity processing for multiple objects (and locations) when the judgments concerned simple features, which is inconsistent with simple object-based theories. Han, Dosher, and Lu (2003; see also Liu, Dosher, & Lu, 2009), for example, showed subjects displays with two Gabor patches that varied in their orientation and phase, and asked
them to make judgments about two different features that could either be from the two Gabor patches (e.g., the orientation of one and the phase of the other) or about a single feature within a single Gabor patch. They found that performance was better for within-object judgments than for between-object 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 for judging both Gabor patches (i.e., the orientations of both or the phase of both) than for a single Gabor patch. This is evidence of unlimited capacity processing across both objects and locations, and 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, there are many studies using visual search that have yielded evidence consistent with unlimited-capacity processing of simple features 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 simple feature is consistent with, and has been offered as evidence of, unlimited-capacity processing of feature information (e.g., Treisman & Gelade, 1980; c.f., Guided search). Moreover, studies using a search task that manipulated whether stimuli were presented all at once simultaneously or in sequential subsets (i.e., the simultaneous-sequential method), a method 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, 2013; Shiffrin &
Gardner, 1972). 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 location-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 task-relevant attributes were simple feature contrasts.

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 does incur costs (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, Palmer, & Moore, 2011a, 2013; Shiffrin & Gardner, 1972; Treisman & Gelade, 1980), that there is little or no cost to processing multiple objects at multiple locations when the judgment is based on a simple feature contrast. No theory that asserts a single pathway of processing that involves object-based limited capacity (e.g., Duncan, 1984) or location-based limited capacity (e.g., Posner, 1980) or both (e.g., Vecera & Farah, 1994; 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, like others have before (e.g., Treisman & Gelade, 1980) that simple feature contrasts are processed within unlimited-capacity pathways, whereas attributes like global shape, a property of an object that cannot be represented on the basis of simple feature-contrasts, must be processed within a limited-capacity pathway. This general framework, which is common to multiple theories of visual processing (e.g., Hoffman, 1979; Treisman & Gelade, 1980; Wolfe,
Divided Attention Is Limited by Objects

1994; 2022), predicts that tasks requiring only judgments based on feature contrast will yield no divided attention effects (unlimited capacity), whereas tasks that depend on an attribute like global shape will yield divided attention effects (limited capacity). Assuming that those predictions hold, the question remains as to whether those divided attention effects are due to object-based or space-based processing limitation, and moreover, whether it is a fixed-capacity processing limitation, such as a serial process would impose.

Overview of Experiments

We addressed these questions in a series of four dual-task experiments that used an extended version of the simultaneous-sequential paradigm used in previous studies to assess capacity limitations (e.g., Scharff et al., 2011a). Across experiments, we varied the task and whether stimuli were spatially separate or overlapping. To preview the results, when the task was a global shape judgment (Experiments 1-3), we found large divided attention effects (limited capacity). When the task required only the judgment of a single feature contrast (size) across multiple objects at multiple locations, there were no divided attention effects (unlimited capacity). Finally, the magnitude of the divided attention effect for global shape judgments was the same for multiple objects at a single location (Experiment 4) as for multiple objects at separate locations (Experiment 2), indicating that the capacity limitation was object based not space based. In no case was there evidence of a space-based capacity limitation.

Together these results are consistent with a multi-pathway processing model with at least one unlimited capacity processing pathway for feature contrast, and at least one limited-capacity processing pathway for which the limit is object based, and not space based. The results are inconsistent with all single-pathway models, including pure space-based theories (e.g., Posner, 1980), pure object-based theories (e.g., Duncan...), and hybrid theories (Vecera).
Experiment 1: Shape Judgments of Multiple Objects at Separate Locations

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

Method

Stimuli. The stimuli were three sets of novel black outline shapes presented on a gray background with 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. 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 each of the four orientations. All images were 100 x 100 pixels (~4 degrees 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.

Gaussian pixel noise was superimposed on the outline shapes during every frame to limit performance. The standard deviation of the noise luminance was 50% of the mean luminance and was used to define its contrast (2 observers had 60% noise contrast). Target contrast and stimulus duration were manipulated to adjust task difficulty for each observer. The target contrast ranged from 35% to 50% with a mean of 44%. The stimulus duration ranged from 0.15 to 0.25 s with a mean of 0.18 s.

1 A fourth set of quadrilaterals with one curved side (not shown) was included for two of the observers, but was dropped for the other observers due to being more difficult to distinguish than the other stimulus sets.
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 ($0^\circ$, $90^\circ$, $180^\circ$, and $270^\circ$) of each exemplar. This resulting in 20 images per set and thus 60 images overall.
**Task and Design.** An extended version of the simultaneous-sequential paradigm was used to measure the effects of divided attention. The task was to report the location of a single pre-specified 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 3 illustrates the three conditions. All trials began with a study display (2 s) that defined the target object, which was followed by a blank display (0.5 s), a fixation display (0.5 s), one or two stimulus displays (0.2 s 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 s). 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 s each), separated by a blank interval (1.8 s). 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 s each), separated by a blank interval (1.8 s). The displays were identical except for the noise.
### Figure 3.
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 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.
Logic. The logic of the extended simultaneous-sequential method is illustrated in Figure 4. 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 will receive 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, for example, can be processed within the time of the display duration, then two stimuli could be processed in the simultaneous condition, whereas four could 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).
<table>
<thead>
<tr>
<th>Simultaneous condition</th>
<th>Unlimited-capacity model</th>
<th>Simple Serial model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st Display</td>
<td>2nd Display</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sequential condition</th>
<th>Unlimited-capacity model</th>
<th>Simple Serial model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st Display</td>
<td>2nd Display</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Repeated condition</th>
<th>Unlimited-capacity model</th>
<th>Simple Serial model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st Display</td>
<td>2nd Display</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Illustration of the logic of the extended simultaneous-sequential paradigm. 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 5 summarizes this logic by showing 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 performance in the overall sequential condition which was fixed at 75%, and an assumed effect of 8 percentage points.

Figure 5. Example predictions of three models are illustrated relative to 75% performance in the sequential condition. 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.
**Observers.** Observers were volunteers with normal or corrected-to-normal visual acuity, some were paid in compensation for their time. To determine the appropriate sample size of observers, we examined the results of four similar experiments in Scharff, Palmer and Moore (2013, excluding Experiment 3 with only 2 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 standard deviation of the sequential advantage was 1.7, 1.8, 3.4 and 4.8% with a mean of 2.9%. Using these 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 .05 and the power of .95 (beta of 0.05). 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 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 appropriate.

**Apparatus.** Experiment 1 was conducted with two apparati. 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 x 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/m², and a black level of 4.1 cd/m², mostly due to 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). A chin rest with an adjustable chair ensured a fixed distance from the display between observers.

Procedure. Observers completed several practice sessions in which the stimuli were high contrast and long duration in order to learn the task. In Experimental sessions, stimulus duration ranged from 0.1 to 0.25 seconds with a mean of 0.18 seconds. Experimental sessions consisted of 12 single-condition blocks 12 trials each (4 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. 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.

Analysis. All statistical tests were 2 tailed. Alpha was fixed at .05. All error bars are standard errors of the mean.

Results

The results of Experiment 1 are shown in Figure 6. 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% ($t(5) = 6.84, p = .001$). There was no reliable difference between the sequential and the repeated conditions ($0.5\% ± 2.0\%, t(5)$
Divided Attention Is Limited by Objects

\[ \theta F = 0.3 \]

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\% \pm 1.1\% (t(5) = 1.42, p > .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 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, 2011b, 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.
Figure 6. 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.
Experiment 2: Shape Judgments of Multiple Objects at Separate Locations 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 period. This cued design, first used in Scharff, et al. (2011b, Experiment 2), controls for potential differences in stimulus interactions, separate from processing limitations, which will be 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 two-alternative forced choice, target present or absent.

Method

Observers. Six observers with normal or corrected-to-normal vision were volunteers. Some were paid in compensation for their time.

Design. Figure 6 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 s) that defined the target, which was followed by a brief noise display (0.2 s), a blank display (0.5 s), the first cue display (0.5 s), another blank display (0.5), and then the first stimulus display (~0.2 s), which contained four shapes: two on each side of fixation. This was followed by another blank display (0.5 s), the second cue display (0.5 s), a blank display (0.5 s), and finally the second stimulus display (~0.2 s). 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 s. 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 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, but 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 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. The noise used with the shapes to limit performance was different than Experiment 1. Dynamic salt-and-pepper 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 .40 to .65 with a mean of .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 would be important for when we conducted an overlapping stimuli version (Experiment 4). Highly visible stimuli (low noise, long duration) were used for training. The entire monitor had a displayable area of about $33^\circ$ by $24^\circ$. Of this, the stimulus and noise fields occupied the central $14^\circ$ by $10^\circ$. The noise fields were $6^\circ$ by $10^\circ$ and were centered $4^\circ$ to either side of fixation. This left a $1^\circ$ space between the inner edge of the noise field and fixation. The squares were jittered by $\pm 0.5^\circ$ around a mean location that was horizontally centered within the noise field and were centered vertically $2^\circ$ above and below the horizontal meridian. The stimulus duration was adjusted for each observer (range 0.1 to 0.5 s, mean of 0.24 s).
Figure 7. 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.
**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 x 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/m², and a black level of 3.9 cd/m², mostly due to 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 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 3 main conditions but 5 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 6 blocks (2 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.
Figure 8. Results of Experiment 2 with a yes-no judgment of shape and a cueing paradigm. 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.
Results

The results of Experiment 2 are shown in Figure 8. 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% (t(5) = 3.36, p = .02). There was no reliable deviation of the sequential condition from the repeated condition (2.1% ± 2.1%, t(5) = 1.02, p > .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 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 (0.7% ± 1.5%, t(5) = 0.43, p > .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 (2.3 ± 2.5%, t(5) = 0.95, p > .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 reviewed in the introduction, the results from many studies using feature-contrast tasks are inconsistent with both object-based and space-based theories in their simplest forms. We sought to confirm this finding within the same logical 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 Experiments 1 and 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 paradigm and kept most of the details the same as Experiment 2. Figure 9 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 s) that defined the target followed by brief noise display (0.2 s), a blank display (0.5 s), the first cue display (0.5 s), another blank display (0.5), and then the first stimulus display (0.2 s). 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 s). 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 s. Performance was limited by superimposing each stimulus display with dynamic salt-and-pepper noise which is represented by the shaded region. The presentation of a noise display near the beginning of the trial was to equate the contrast adaptation of the two stimulus displays.

In the cued-simultaneous condition, there were cues that indicated 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 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 to compare the main conditions to, to 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.
Figure 9. 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 8 stimuli were relevant. The presence of dynamic noise is indicated by shading.
**Apparatus.** The apparatus was the same as Experiment 2 and eye position was recorded. 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 s. 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 AM used 3 pixels). They were presented as one-pixel thick, black outlines on a mid-gray background. 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 ($p = .65$) 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.

**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 8 blocks (2 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**

The results of Experiment 3 are shown in Figure 10. The percent area under the ROC is plotted against the four 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: a mean sequential advantage of -1.0% ± 1.5% (95% CI = -1.4; 2.4, BF = 0.4)
(t(5) = 0.65, p > .1). Instead, there was a reliable difference between the cued-sequential
condition and the cued-repeated condition of 6.4 ± 1.7% (t(5) = 3.78, p = .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 (-1.3% ± 2.4%, t(5) =
0.56, p > .1). The cued-simultaneous condition was also examined and showed no reliable
difference (2.0% ± 1.0%, t(5) = 2.07, p > .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. This difference of 3.7 ± 0.9% was reliable (t(5) =
4.23, p = .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.
Figure 10. Results of Experiment 3 with a yes-no judgment of size and a cueing paradigm. Percent area under the ROC is shown for the three 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.
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 demonstrate a clear processing limitation, they are consistent with both object-based and space-based theories 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 paradigm that was introduced in Experiment 2 so that stimuli were identical across conditions. This was especially important for this experiment because it used overlapping stimuli for which stimulus interactions were expected to be substantial. Using the cued version of the method equated visibility across conditions. Figure 11 shows the three conditions: cued-simultaneous, cued-sequential, and cued-repeated. A trial began with a study display (2 s) followed by a brief noise mask (0.2 s) and a blank display (0.5 s). Then, the first cue (1.0 s) followed by a blank (0.5 s) and the first stimulus display (ranged from 0.15-0.25 s with a mean of 0.19 s), 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. In all trials, there was only a single relevant location and at most two relevant objects. 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 display 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. As in Experiment 1, this experiment was run using two apparati each used by three of the observers.

**Apparatus.** Same two apparati as Experiment 1. We did no eye tracking for this experiment with foveal stimuli.
<table>
<thead>
<tr>
<th>Cued simultaneous</th>
<th>Cued sequential</th>
<th>Cued repeated</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
<td><img src="image3" alt="Diagram" /></td>
</tr>
<tr>
<td>Study display</td>
<td>Study display</td>
<td>Study display</td>
</tr>
<tr>
<td>2 s</td>
<td>2 s</td>
<td>2 s</td>
</tr>
<tr>
<td>Noise</td>
<td>Noise</td>
<td>Noise</td>
</tr>
<tr>
<td>0.2 s</td>
<td>0.2 s</td>
<td>0.2 s</td>
</tr>
<tr>
<td>Blank</td>
<td>Blank</td>
<td>Blank</td>
</tr>
<tr>
<td>0.5 s</td>
<td>0.5 s</td>
<td>0.5 s</td>
</tr>
<tr>
<td>First cue</td>
<td>First display in noise</td>
<td>First display in noise</td>
</tr>
<tr>
<td>1.0 s</td>
<td>~0.2 s</td>
<td>~0.2 s</td>
</tr>
<tr>
<td>Blank</td>
<td>Blank</td>
<td>Blank</td>
</tr>
<tr>
<td>1.0 s</td>
<td>0.5 s</td>
<td>0.5 s</td>
</tr>
<tr>
<td>Second cue</td>
<td>Second display in noise</td>
<td>Second display in noise</td>
</tr>
<tr>
<td>1.0 s</td>
<td>~0.2 s</td>
<td>~0.2 s</td>
</tr>
<tr>
<td>Blank</td>
<td>Blank</td>
<td>Blank</td>
</tr>
<tr>
<td>1.0 s</td>
<td>1.0 s</td>
<td>1.0 s</td>
</tr>
</tbody>
</table>

Figure 11. [Color Figure.] 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.
Stimuli. The three sets of shapes from Experiment 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 displayed 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 Experiment 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 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 (\( p = .65 \) for three observers, \( p = .75 \) for the others).

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 s with a mean of 0.19 s and the noise probability ranged from .65 to .75 with a mean of .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.
Figure 12. 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.
Results

The results of Experiment 4 are shown in Figure 12. 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 (10.5% ± 1.8% (t(5) = 5.89, \(BF = 28.4\), \(p = .002\)). In contrast, there was no reliable difference between the cued-sequential and the cued-repeated conditions (0.7% ± 1.4%, \(t(5) = 0.47, p > .1\)). As with Experiment 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 (0.2% ± 1.7% (t(5) = 0.10, \(p > .1\)) or the cued-simultaneous condition (2.7% ± 3.2%, \(t(5) = 0.84, p > .1\)). Thus, we are confident that collapsing over the two intervals is appropriate.

The results of this experiment indicate that the divided attention effects that we have observed in previous experiments 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. We used a shape judgment task, which was assumed to engage object processing, with non-overlapping (Experiments 1 & 2) and overlapping (Experiment 4) stimuli, and found evidence of fixed-capacity processing in all cases. In contrast, a simple feature-contrast task, which was assumed to not require object processing, but did require the processing of stimuli at multiple locations (Experiment 3), yielded evidence of unlimited capacity processing. Together these results are consistent with object-based theories of divided attention, with no evidence of spatial limitations *per se* on divided attention, and with there being a separate processing pathway for tasks that do not require object processing, but only simple feature contrast detection.

Figure 13 summarizes the results more specifically within the context of the logic of the extended simultaneous-sequential method used in this study (summarized in Figure 5). First consider the results of Experiments 1 and 2, both of which investigated the capacity limitations of global shape judgments of stimuli at multiple locations. In both experiments, there was a reliable advantage of sequential over simultaneous presentation (Figure 13, Panel A), and no reliable difference between sequential and repeated presentations (Figure 13, Panel B). This is the signature pattern for fixed-capacity processing (see Figure 5, right panel). To address whether that fixed-capacity reflects object-based or space-based limitations, consider Experiment 4, which also used the shape judgment task but with stimuli presented overlapping at a single location. There was again a reliable sequential advantage with no repeated advantage, consistent with the same fixed-capacity processing limitation that was measured in the Experiments 1 and 2. Finally, consider the results from Experiment 3, which measured processing capacity...
limitations for a simple feature-contrast task. In this case, there was no reliable advantage for sequential presentation over simultaneous presentation (Figure 13, Panel A), but there was a reliable advantage for repeated displays (Figure 13, Panel B). This is the signature pattern for unlimited capacity processing (Figure 5, left panel). Thus, the results showed fixed capacity for object processing, regardless of the number of locations, but unlimited capacity processing for simple feature contrast detection, despite multiple locations.
Figure 13. 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.
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, allowing the attribution of any effects that do occur to limitations on object processing. It is possible that there would be further processing limitations if the stimuli must also be processed at multiple locations as Vecera and Farah (1994) applied this logic to Duncan’s original divided attention study using overlapping stimuli, and found no difference in divided attention effects for overlapping versus separated stimuli, suggesting that there is no additional cost associated with the stimuli appearing in multiple locations (see also Kramer, Weber & Watson, 1997, Experiment 2).

An analogous comparison can be applied to the results of the current study. Specifically, Experiment 4, which used overlapping objects, revealed evidence of fixed capacity processing (1 location), and was similarly fixed capacity in Experiments 1 and 2, which used non-overlapping objects (2 or more locations). This is evidence against there being any additional effect of location over that of objects. Thus, evidence from entirely different tasks and experimental methods for measuring divided attention effects using the overlapping stimuli strategy all suggest the same conclusion; divided attention is limited for objects but not for space.

---

2These experiments have been criticized on grounds that in the overlapping condition, stimuli were at fixation, whereas in the separated case, they were each a bit off of fixation and would, therefore, have been less visible in the separated case (Kramer et al., 1987). However, the eccentricity was quite small (-2°), and so it is not clear how much less visible the separated stimuli were. More important, the finding was that there was no effect of separation on the size of the divided-attention effect. The prediction for a space-based limitation is that the effect would increase with separated stimuli. It’s not clear how decreased visibility would result in countering an increased processing demand. In our opinion, a more significant concern about comparing overlapping stimuli to separated stimuli is that it confounds the number of locations (1 versus 2) with susceptibility to visual crowding (more versus less). This confound was exacerbated in Experiment 1 of Kramer et al (1997) when the overlapping stimuli were presented off of fixation where visual crowding increases.

It is implausible that decreased visibility would have counteracted an effect of separation.
Strategies other than overlapping stimuli have been used to study object-based divided attention. However, few studies were designed to directly assess the possibility of space-based divided attention effects beyond object-based divided attention effects. One general strategy, for example, has been to present objects at multiple locations, but to equate the spatial distance between within-object and between-object to-be-reported attributes of the stimuli (e.g., Baylis & Driver, 1993; Lavie & Driver, 1996; Watson & Kramer, 1999). Object-based divided attention effects have been reported from all of these methods (but see Davis & Holmes, 2005), but only one spatial modulation of an object-based effect has been reported (Lavie & Driver, 1996) and two separate attempts to replicate that finding failed (Abrams & Law, 2002; Lamme & Eggeth, 2000).

Another general strategy for isolating object-based divided attention effects has been to use motion, as in multiple-object tracking (see Meyerhoff, Papenmeier & Huff, 2017 for a review). But again, studies using this strategy have provided little insight into the possibility of space-based divided attention effects as well. There is evidence that when tracking a subset of randomly moving identical objects, even controlling for visual crowding effects, the closer a tracked object gets to another tracked object, the worse performance becomes (Shin, Alvarez & Jiang, 2008). This cost could reflect a role of space (proximity) for defining and maintaining separate object representations (see Bahcall & Kowler, 1999 for a similar effect with static displays), but it is clearly not a cost associated with increasing the number of relevant locations (i.e., a space-based divided attention effect) since performance was worse with less spatial separation not more.

A number of older studies that were focused on measuring the spatial profile of attention, manipulated the spatial separation of individual stimuli in dual-task experiments, and therefore
afforded the possibility of revealing divided attention effects. Effects of distance were found such that performance was better when stimuli were closer than when they were farther (e.g., Hoffman & Nelson, 1981; Hoffman, Nelson & Houck, 1983; Sagi & Julesz, 1986). However again, these studies were not designed to test between object-based and space-based divided attention effects, and cannot rule out the possibility that proximity played a role in how stimuli were organized.

Finally, a study reported by Kim and Vergoese (2014) addressed the question of whether monitoring one versus multiple locations that were on one versus multiple surfaces, and found that performance improved with reduced spatial uncertainty only when the locations were on separate surfaces; little or no spatial uncertainty effect occurred within a single surface. This is consistent with there being an object (surface in this case) based processing limitation and no space-based limitation.

In summary, consistent with the results and conclusions of the current study, in summary, there is no strong evidence of space-based divided attention effects from previous studies either.

Multiple processing pathways

The focus of this study has been to test object-based and space-based theories of divided attention, and the results lead to the conclusion that there are object-based processing limitations but no space-based processing limitations as such. There is, however, an additional aspect of processing limitations that is required to understand the full set of results reported here and within the extant literature. That is that some attributes of stimuli, such as simple feature contrasts, do not in themselves constitute object properties and can drive responses via a separate unlimited-capacity processing pathway. In the current study, we conducted an experiment in
which the task required only that observers detect a target feature contrast, and found evidence of unlimited capacity processing. These results are consistent with previous findings from dual-task studies (e.g., Han et al., 2003; Liu et al., 2009), other simultaneous-successive experiments (Huang & Pashler, 2005; Scharff et al., 2011a, 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 clearly limited-capacity processing pathways revealed in the current study and elsewhere.

Figure 14 illustrates an example multiple pathway model that accommodates the full set of results. Relevant stimuli (three in the example, represented by S1, S2, and S3) are processed initially in terms of feature contrast information. This processing is assumed to have unlimited capacity, which is shown in the figure by separate arrows for each stimulus. Further processing follows two routes, and which route limits performance depends on the task. For tasks that require only feature contrast (e.g., luminance-increment detection), the feature-contrast pathway is sufficient to guide decision and response processes (left-side path in the figure). For these tasks, all of the stimulus processing that is necessary for the task is unlimited capacity, leading to an absence of divided attention effects, just as has been observed here and elsewhere (e.g., Scharff et al., 2011; Experiment 3 of the current study). In contrast, for tasks that depend on properties that are intrinsic to objects, such as global-shape discrimination (e.g., Pizlo, 2008; Wolfe & Bennett, 1997), feature-contrast information alone is insufficient, and object processing is necessary (right-side path in the figure). Under this theory, object processing has limited-capacity, perhaps fixed capacity as our results with global shape have suggested. This is
Figure 14. Illustration of a dual-route theory that includes separate routes for feature contrast and object properties. Relevant stimuli initially are processed in terms of their features and then, depending on the task, take one of two routes to reach decision and response processes. The arrows on the left represent an unlimited-capacity "shortcut" route that uses feature contrast to make a decision. The arrows on the right represent the routes of object processing required for...
establishing object representations. This route has limited capacity which is illustrated by the single arrow that extends to the decision and response processes.

There are many precedents for the general idea of a multiple-pathway model of the sort described here. Hoffman (1979), for example, proposed an early two-stage model of visual search that included an initial parallel stage and a later serial stage. Feature Integration Theory is a classic elaboration of this two-stage model. It was the first to propose that the early stage is not only parallel, but parallel unlimited capacity, and that the function of the early stage is feature representation, while 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, Treisman and Gibbs (1992) expanded this work to encompass a larger theory of perceptual organization and visual memory that asserts that the establishment of object representations, in particular, depends on limited-capacity, perhaps even serial, processing. Wolfe and colleagues’ original Guided Search model also included an initial parallel unlimited capacity processing stage for feature encoding and a later serial stage (Wolfe, 1994; Wolfe, Cave, & Franzel, 1989). A difference between Guided Search and other two-stage models at the time was that according to Guided Search, the feature stage cannot be accessed directly, it can only contribute to guidance of attention to locations of stimuli.

As such, the feature processing stage itself does not constitute a direct alternative pathway to response processes. More recent versions of guided search, however, include a non-selective unlimited-capacity processing pathway that extracts global scene statistics (e.g., Vo & Wolfe, …; Wolfe, 2022). This brief review is not intended to cover all previous examples of multiple pathway models. Rather it is to make clear that the general idea of a multiple pathway processing architecture is neither novel nor controversial. We highlight the idea here because we
believe that a simple model of the kind illustrated in Figure 14 is sufficient to account for a large body of results from a wide range of studies testing object-based versus space-based theories of divided attention. And again, the conclusion is that visual processing is limited by objects, but not by space, and that tasks that require only simple feature contrast information 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 us to be a safe choice to reliably engage 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.

How might a task that depends on feature conjunctions be processed within an architecture like that illustrated in Figure 14? In many cases, because conjunctions depend on more than simple feature contrast, they will require processing along the limited capacity pathway consistent with FIT and other models. Under the right conditions, however, tasks that depend on feature conjunctions might create feature contrast signals that can be detected by the unlimited-capacity processing route emergent features of a sort. Results from a study using the simultaneous-sequential method to measure divided attention effects support this prediction. 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 proposal does not provide specific predictions for all feature conjunction tasks because feature conjunctions constitute a class of stimulus attributes that depending on the specific conditions may sometimes be extractable through the unlimited-capacity feature-contrast pathway, but other times not. Rosenholtz (2001), for example, showed that performance in a wide 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. Increased heterogeneity will reduce the strength of feature contrast, and increase the likelihood that a given task will require processing along the limited capacity pathway.

Finally, work by Dosher and colleagues 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 only on processing within the unlimited-capacity pathway. The second finding is consistent with the hypothesis that associating a particular feature with a particular object, and not another object, which is what must be done to report different features of two objects, 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 FIT, but is more generally about associating specific information with specific object representations (c.f. Kahnman et al., 1992).

...will re-write this later, when we re-write the abstract
We 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: osf.io/fn6x3/.