Divided attention limits perception of 3-D object shapes

Alec Scharff   Department of Psychology, University of Washington, Seattle, WA, USA ✉

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

Cathleen M. Moore   Department of Psychology, University of Iowa, Iowa City, IA, USA 🏠 ✉

Can one perceive multiple object shapes at once? We tested two benchmark models of object shape perception under divided attention: an unlimited-capacity and a fixed-capacity model. Under unlimited-capacity models, shapes are analyzed independently and in parallel. Under fixed-capacity models, shapes are processed at a fixed rate (as in a serial model). To distinguish these models, we compared conditions in which observers were presented with simultaneous or sequential presentations of a fixed number of objects (The extended simultaneous-sequential method: Scharff, Palmer, & Moore, 2011a, 2011b). We used novel physical objects as stimuli, minimizing the role of semantic categorization in the task. Observers searched for a specific object among similar objects. We ensured that non-shape stimulus properties such as color and texture could not be used to complete the task. Unpredictable viewing angles were used to preclude image-matching strategies. The results rejected unlimited-capacity models for object shape perception and were consistent with the predictions of a fixed-capacity model. In contrast, a task that required observers to recognize 2-D shapes with predictable viewing angles yielded an unlimited capacity result. Further experiments ruled out alternative explanations for the capacity limit, leading us to conclude that there is a fixed-capacity limit on the ability to perceive 3-D object shapes.

Introduction

In perception, divided attention refers to situations in which an observer must monitor multiple stimuli at once. Such demands are common in real-world tasks such as driving a car. Here we investigate how divided attention affects shape-based object recognition. Although many theories posit detrimental effects of divided attention on object shape perception, there has been no definitive empirical study of these effects.

The scarcity of research on this topic may be due to the difficulty of isolating the phenomena involved from other cognitive and perceptual processes. Shape perception can be conflated with non-shape feature discrimination (Evans & Treisman, 2005) or semantic categorization (Warrington & Taylor, 1978). We aim to dissociate the effects of attention on perceptual processes from effects on sensory, memory, and decision processes (Palmer, 1995).

Shape-based object recognition

Shape-based object recognition is achieved when an observer apprehends an object's spatial contours sufficiently to match it to a known exemplar. For the purpose of object recognition, the diagnostic value of a three-dimensional shape is unsurpassed by other visual features. Object shape perception is the result of *shape-constancy* mechanisms that infer three-dimensional forms from a retinal image containing only two-dimensional projections of those forms. Thus, object shape percepts remain constant despite changes in the corresponding retinal images (Pizlo, 2008).

Our goal is to measure divided attention effects on the perception of object shape. To do so, we use a psychophysical task that relies solely on object shape perception. Below we elaborate our approach to minimizing extraneous non-shape information in the task.

Equating non-shape features across objects to prevent feature matching

We attempted to eliminate any diagnostic non-shape features in the stimuli because these features might be

exploited to identify target objects without achieving object shape constancy. For example, an elephant may be identified not only by its shape but also by its gray color, bumpy texture, or large size. We refer to a strategy that exploits these diagnostic non-shape features as feature matching. To prevent feature matching, we equated non-shape features across objects. For example, if an elephant were a stimulus in the experiment, then this approach would require observers to recognize an elephant among other large, gray, bumpy objects.

Rotating objects to preclude template matching

Even when non-shape image features are equated, there remains the possibility that observers could employ a template-matching strategy to identify an object without achieving shape constancy. In an example template strategy, the observer compares luminance value at each pixel in an image with a representation in memory and makes a decision based on the degree of match. More plausibly, the observer could identify an object on the basis of some partial template (e.g., a diagonal edge in the lower-left corner of the image). Template matching strategies are possible when the observer knows the exact form the image will take, so we attempted to prevent such strategies by unpredictably varying the observer's view of the object. In two of our experiments (Experiments 5 and 6), we examined the possibility of template matching by using predictable images.

Use novel objects to minimize role of semantic categorization

Another concern is that performance in an ostensibly perceptual task might also depend on semantic processes. For example, reading words is subject to an attentional-capacity limitation (Scharff, Palmer, & Moore, 2011b), but it is ambiguous whether this limit is attributable to the perceptual demands or the semantic demands of the task. To minimize the role of semantics, we used novel object stimuli that did not have prior semantic associations and required observers to recognize specific physical objects rather than categorize them.

Divided attention and capacity

In perception, effects of divided attention can be characterized in terms of *capacity*, the quantity of information that can pass through a system during a given time interval (Broadbent, 1958; Kahneman, 1973; Townsend, 1974). Two extreme, boundary-defining model classes are *unlimited-capacity models* and *fixed-capacity models*. Here, capacity refers to throughput at the level of the perceptual system. All models assume that we accrue information over time, but they differ in the extent to which dividing attention among multiple objects affects this information accrual.

Unlimited-capacity models posit that divided attention among multiple objects does not limit perception, because there is no system-level limit on perceptual processing. Importantly, unlimited-capacity processing does not necessarily imply fast or accurate object recognition, because perception is still noise-limited at the level of local mechanisms. The definitive property of unlimited-capacity models is that the rate of individual stimulus processing is not degraded by divided attention. The prototype model of this class is the standard parallel model in which multiple analyses occur in parallel, with the rate of each analysis unaffected by the number of objects being analyzed (Gardner, 1973). A less intuitive example model is the super-capacity serial model in which only one analysis can be carried out at a time, but the speed of each analysis increases proportionally with the number of objects of interest so that system-level capacity is unlimited (Townsend, 1974). The standard parallel model and super-capacity serial model can make equivalent predictions for many experiments, making them difficult to distinguish.

In contrast, fixed-capacity models assume an inflexible limit on the overall rate of information accrual. An intuitive example is the standard serial model in which one object process must be completed before another can begin (Davis, Shikano, Peterson, & Michel, 2003; Townsend, 1974). Parallel processing can also have fixed capacity, for example, if multiple concurrent analyses are carried out with efficiency inversely proportional to the number of relevant stimuli (e.g., the parallel sampling model of Shaw, 1980).

Between these two extremes are *intermediate-capacity models* in which there is a limit less restrictive than fixed capacity models propose (e.g., limited resource models, Norman & Bobrow, 1975; crosstalk models, Mozer, 1991). The more common term *limited capacity* encompasses both intermediate-capacity and fixed-capacity models.

Relationship to parallel-serial architecture

Many authors have characterized divided attention effects in terms of processing architecture—whether object processes occur in serial or in parallel. It is difficult to distinguish between parallel and serial models because the predictions of a particular serial model can often be mimicked by an appropriately formulated parallel model and vice versa (Townsend, 1974). For example, the standard serial and parallel sampling models make equivalent predictions for many

experiments, as do the standard parallel and super-capacity serial models. For this reason, we focus on capacity, rather than architecture, as a tractable property that characterizes how perception is affected by divided attention.

Effects on perception versus decision

Attention research is usually concerned with the effects of non-stimulus manipulations on the quality of perception. However, effects on perception can be difficult to distinguish from effects on the decision processes involved in psychophysical performance (Eckstein, Peterson, Pham, & Droll, 2009; Palmer, Verghese, & Pavel, 2000). In the domain of divided attention phenomena, the most common experimental approach is the visual search paradigm, in which an observer searches for a specified target stimulus among some number of distracter stimuli. Response time is often measured as a function of the number of distracter stimuli with the total number of stimuli referred to as *set size*. Large effects of set size on response time are sometimes interpreted as evidence for limited-capacity perception. However, this interpretation neglects the impact of set size on decision processes. If judgments about individual stimuli are subject to error, then increasing the number of stimuli under consideration in a given task increases the overall error rate, as when multiple comparisons are made in statistical decisions (Palmer et al., 2000; Shaw, 1980; Tanner, 1961). Thus, large set-size effects may be observed even in the absence of capacity limitations, as demonstrated by Huang and Pashler (2005).

A second common method of investigating perceptual capacity limitations is the psychophysical dual-task experiment, in which observers make two independent psychophysical judgment concurrently (Bonnel, Stein, & Bertucci, 1992; Sperling & Melchner, 1978). Performance in the dual task can then be compared to single-task controls to determine whether dividing attention in the dual-task condition caused performance decrements. The challenge with this method is to determine whether any performance decrements are due to loss in perceptual sensitivity or to interference in decision and/or response mechanisms because the number of stimuli to be monitored is confounded with the number of decisions and responses to be made.

In the experiments presented here, we use the extended simultaneous-sequential method to measure the effects of divided attention on perception. This method uses a visual search task but keeps the decision structure and response alternatives constant across conditions. Thus, any divided attention effects are attributable solely to perception.

Theories of attention and object recognition

Early accounts described all of perception as capacity limited (Kahneman, 1973; Posner, 1980). This notion has been contradicted by numerous studies that show unlimited-capacity performance in search for image features such as stimulus luminance, orientation, color saturation, and size (Bonnel et al., 1992; Palmer, 1994; Scharff et al., 2011b). Subsequent theories proposed that some aspects of perception have unlimited capacity, while others are subject to capacity limits. Most prominently, Treisman and Gelade (1980) proposed that while image feature perception has unlimited capacity, perceiving the conjunction of two or more features in the same object requires a limited-capacity mechanism. This theory is challenged by studies showing unlimited-capacity performance in conjunction search tasks (Eckstein, 1998; Eckstein, Thomas, Palmer, & Shimozaki, 2000; Huang & Pashler, 2005). Several modern theories describe object perception as a limited-capacity process (Kahneman, Treisman, & Gibbs, 1992; Wolfe, 1994). Others have challenged the view that object perception is limited capacity, proposing that at least some special cases can have unlimited capacity (Allport, 1987; Rousselet, Thorpe, & Fabre-Thrope, 2004; Van der Heijden, 1996).

There is no solid evidence to support either kind of theory for shape perception. Though several studies have addressed the idea or related questions, none provide unambiguous results. Each is rendered ambiguous by a combination of reasons given above. Here we briefly review some representative studies and discuss why we consider them inconclusive.

Biederman, Blickle, Teitelbaum, and Klatsky (1987) used the response time search task with line drawings of familiar objects as stimuli (e.g., car, fire hydrant, filing cabinet). Following a brief display of an array of objects, subjects indicated whether or not a specific object had been present. Both error rate and response time increased with set size, effects that were interpreted as evidence for limited-capacity object perception. However, Biederman et al.'s use of the response-time set-size paradigm makes this interpretation ambiguous, because the set-size effect could be explained by an effect on the decision processes as described above. Furthermore, the semantic categorization required by the task confounds object shape recognition with semantic categorization.

More recent work by Rousselet, Fabre-Thrope, and Thorpe (2002) used search studies in which subjects searched for animals or vehicles among distracter images that did not contain those targets. Rousselet et al. compared speed and accuracy performance to the predictions of unlimited-capacity models. The results were consistent with standard parallel models of object perception, at least in the special case of animal and

vehicle searches. However, as others have pointed out, such categories of image might easily be discriminated by non-shape image features, either those intrinsic to the object (e.g., furriness; Evans & Treisman, 2005) or those induced by the style of photography used in images containing animals (sharp focus and blurred background, Wichmann, Drewes, Rosas, & Gegen-furtner, 2010). Thus, although object recognition tasks yielded unlimited capacity results in these studies, it is not clear if parallel search was accomplished on the basis of object shape recognition or on the basis of non-shape diagnostic features.

Our recent study addresses a closely related question but confounds semantic categorization with object shape perception (Scharff, Palmer, & Moore, 2011a). We tested object categorization under divided attention, using a search task in which observers searched for a particular kind of animal on each trial. In an effort to minimize non-shape features, we used animal images as both targets and distracters. We required observers to indicate the presence of a specific animal type (e.g., squirrel, moose, etc). In this study, we found unambiguous fixed-capacity results. However, it is still unclear whether it was the shape perception or the categorization aspect of this task that caused the fixed-capacity limit.

To summarize, our goal was to characterize divided attention effects on shape-based object recognition. We designed our stimuli and tasks so that objects could be distinguished only on the basis of object shape, not by feature or template matching, and without requiring semantic categorization. To measure divided attention effects, we used the extended simultaneous-sequential paradigm. This method can distinguish unlimited-capacity and fixed-capacity processes while relying on relatively modest assumptions (Scharff et al., 2011b).

General method

The six experiments described here used the extended simultaneous-sequential paradigm to distinguish models of divided attention (Eriksen & Spencer, 1969; Scharff et al., 2011b; Shiffrin & Gardner, 1972). All of the experiments use the same three experimental conditions. These conditions are simultaneous, sequential, and repeated (illustrated in Figure 1). In this section we describe the three conditions and the logic of the predictions. Quantitative predictions are presented in the Appendix.

Localization search task

In all conditions, observers performed a localization task, indicating the location of specified target stimulus in a briefly presented array. Four stimulus locations
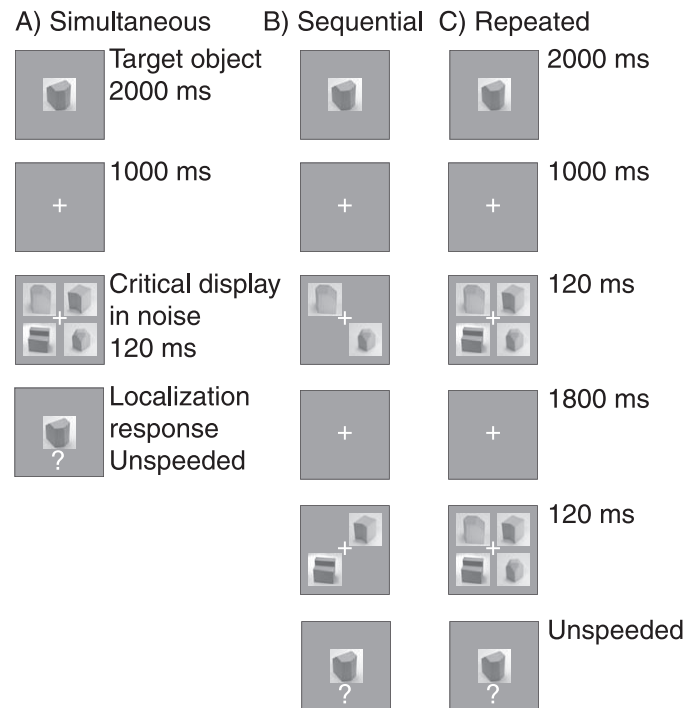

Figure 1. Schematic of conditions used in Experiment 1. Each condition began with a preview of the target object. Next the critical display appeared, depicting the target object in one of four locations. The target object was always shown from a different viewpoint in the critical display than in the preview display. The critical display differed across conditions. In the simultaneous condition, all four stimuli were shown concurrently. In the sequential condition, two stimuli were shown in one display and then the remaining two stimuli were shown in a subsequent display. In the repeated condition, all stimuli were shown twice. Following the critical display, the observer responded by indicating the location where the target object had appeared.

were positioned at the corners of an invisible square surrounding fixation. To begin each trial, observers saw a preview image that indicated the target object for that trial (first row of Figure 1). Target stimuli varied randomly between trials. The preview image subtended 3.9° in the center of the display and persisted for 2000 ms. The preview was then replaced by a fixation cross that persisted for 1000 ms. Next, the critical display began. The critical display was a simultaneous, sequential, or repeated presentation of the four stimuli, as described in the sections that follow. Observers were instructed to maintain fixation throughout the trial (in Experiment 4 fixation was enforced with eye tracking). In all conditions, the display included the target stimulus along with three distracter stimuli. Following the critical display, the preview image reappeared and observers pressed one of four keys to indicate the location where the target had appeared. Observers were instructed to strive for the best possible accuracy

performance and not rush their responses. Conditions alternated in blocks throughout each session, with six trials to a block. Each block was preceded by a warning screen indicating the condition of the following block.

Simultaneous condition

In the simultaneous condition (Panel A of Figure 1), the critical display consisted of a brief, simultaneous presentation of all four stimuli. Each image subtended 3.9° and was centered 3° eccentric from the fixation cross. The center-to-center spacing of neighboring images was 4.2°. The objects were shown with superimposed dynamic noise (these are described in the Methods section for each experiment).

Sequential condition

In the sequential condition (Panel B of Figure 1), the critical display showed the four stimuli sequentially in two subdisplays (henceforth referred to as intervals). Two of the four stimuli were shown in the first interval and then the other two stimuli were shown in a second interval. There was an 1,800 interstimulus interval (blank screen between the two displays). This relatively long interstimulus interval was used to provide ample time for shifts of attention (Duncan, Ward, & Shapiro, 1994). To discourage eye movements, each interval always presented stimuli on opposite sides of fixation (in Experiment 4, the role of eye movements was explicitly tested with eye tracking). The order in which the subsets of objects appeared alternated by block and observers were informed of the order by a warning screen preceding each block.

Repeated condition

In the repeated condition, the critical display had two intervals presented sequentially and all four images appeared in both intervals (Panel C of Figure 1). Timing was as in the sequential condition. Identical stimuli were used in the first and second intervals, but unique noise sequences were generated for each interval.

Predictions

All models predict better accuracy in the repeated than the simultaneous conditions. Fixed-capacity models predict that sequential accuracy will be similar to repeated accuracy, while unlimited capacity models predict similarity between simultaneous and sequential accuracies. We refer to Figure 2 to demonstrate the logic of these predictions. The placement of each gray bar indicates how and when the stimuli appears in a

trial. Locations are listed vertically and numbered one to four; display intervals are listed horizontally and labeled "first" and "second." The black arrows overlying the gray bars represent the observer's analysis of a stimulus.

The predictions of unlimited-capacity models are illustrated with the standard parallel model (left column of Figure 2). Here, the observer analyzes all visible stimuli independently and in parallel. Accuracy is determined by the amount of time available to analyze the stimuli. In both the simultaneous and sequential conditions, each stimulus is displayed for one interval duration; therefore the model predicts equivalent accuracy between these two conditions (Shiffrin & Gardner, 1972). In the repeated condition, the stimuli are visible for a longer time (two interval durations), yielding better accuracy than the other conditions. Overall, the standard parallel model (and other unlimited-capacity models) predicts the pattern simultaneous = sequential < repeated.

The predictions of fixed-capacity models are illustrated with the standard serial model (right column of Figure 2). In this example, the observer has time to scan through just two of the stimuli during each brief interval duration. The observer has no information about the unscanned stimuli. Under this model, accuracy is impaired in the simultaneous condition (in which the observer has no information about two of the stimuli) compared to the sequential and repeated conditions (in which the observer has information about all four stimuli). Thus accuracy is equivalent in the repeated and sequential conditions, and both are superior to the simultaneous condition, yielding the overall predicted pattern simultaneous < sequential = repeated. This prediction can be generalized to several other serial models and to fixed-capacity parallel models (see Appendix).

The predictions are predicated on the assumption that accuracy is limited by stimulus display duration (i.e., accuracy would improve given increased display time). The comparison between the simultaneous and repeated conditions can be used to verify whether this assumption is correct. In some pilot studies, we have found cases that yielded similar accuracy in the simultaneous and repeated conditions even though the repeated conditions present stimuli for twice the duration as the simultaneous condition. For example, brief displays of objects without noise can yield this result in violation of the assumption of duration-dependent accuracy. One explanation for this violation is that the duration of the simultaneous condition exceeds the maximum window of temporal integration, so that additional viewing time provides no benefit. Whatever the explanation, if the repeating the display does not improve accuracy then results from this paradigm are insensitive to the capacity differences of
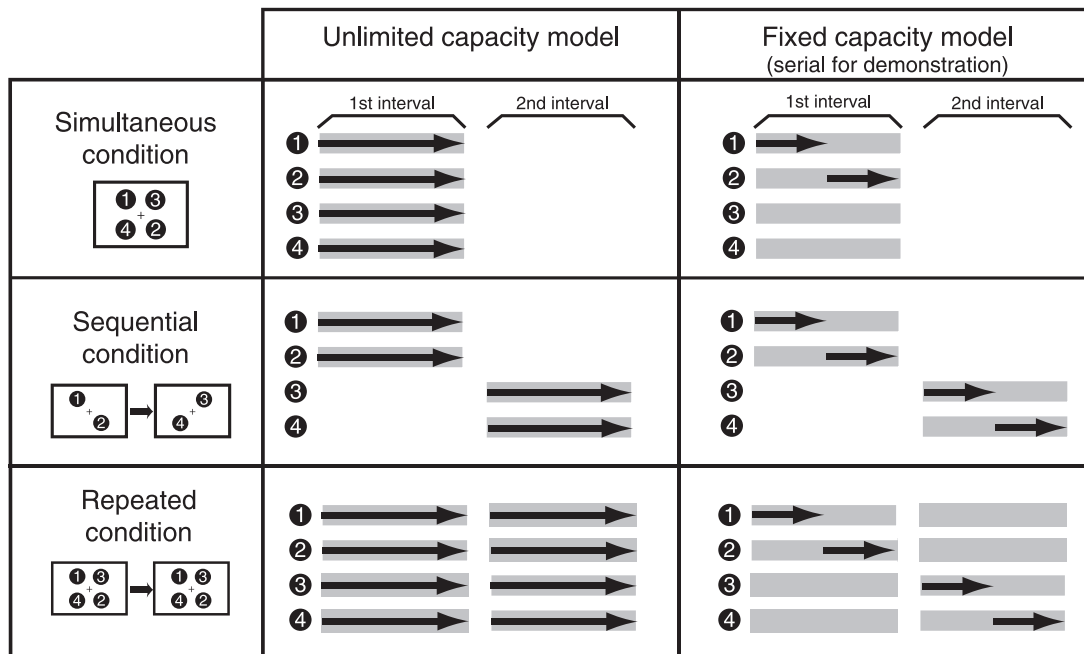
Figure 2. Illustration of the logic for predictions from unlimited-capacity and fixed-capacity models. Gray bars represent the presence of a stimulus and black arrows represent perceptual processing of a stimulus. Under unlimited-capacity models, accuracy improves with longer exposure duration but is not influenced by the number of relevant stimuli (illustrated here with the standard parallel model). Thus, the predicted pattern of results simultaneous = sequential < repeated. Under fixed-capacity models, accuracy improves with exposure duration but is degraded when the number of relevant stimuli is increased (illustrated here with the standard serial model that allows the observer to scan only two stimuli in each brief display). Fixed-capacity models predict the pattern simultaneous < sequential = repeated.

interest. Consequently, simultaneous-sequential experiments without a repeated condition cannot distinguish an unlimited capacity result from an insensitive experiment.

Session length and practice

Experimental sessions included 48 trials in each of the three conditions and lasted approximately 20 minutes. Each observer completed at least two practice sessions before beginning the experiment. If observers were not performing reliably better in the repeated than simultaneous conditions, they were given additional practice and instructions to exploit both intervals of the repeated display. Following practice, each observer completed 10 experimental sessions for a total 480 trials per condition per observer. Experiment 3 instead used 30 sessions with 24 trials per condition per duration.

Overview of experiments

Here we provide a brief summary of the experiments to aid the reader in following their progression. Full methods and results follow.

In Experiment 1, we used the extended simultaneous-sequential method to measure capacity limits in a shape-based recognition task. We used the strategies described earlier to prevent template matching, feature matching, and semantic categorization. Here we found results consistent with fixed capacity. This is the primary finding of the study: A carefully controlled shape-based object recognition task yields fixed-capacity results.

Experiments 2 through 4 ruled out alternative explanations for the fixed-capacity result in Experiment 1. Experiment 2 showed that this result was not due to physical display differences between conditions. Experiment 3 showed that capacity is independent of task difficulty. Experiment 4 showed that the fixed-capacity result was not caused by saccadic eye movements. Through the first four experiments, all results were consistent with the fixed-capacity processing in shape-based object perception.

For Experiments 5 and 6, we sought complementary unlimited-capacity results. Contrasting unlimited-capacity results would help delineate which aspects of perception invoke capacity limits. Experiment 5 was similar to Experiment 1, but we used unrotated object stimuli and encouraged observers to adopt a template-matching strategy. Contrary to our prediction, we found results consistent with fixed capacity. We
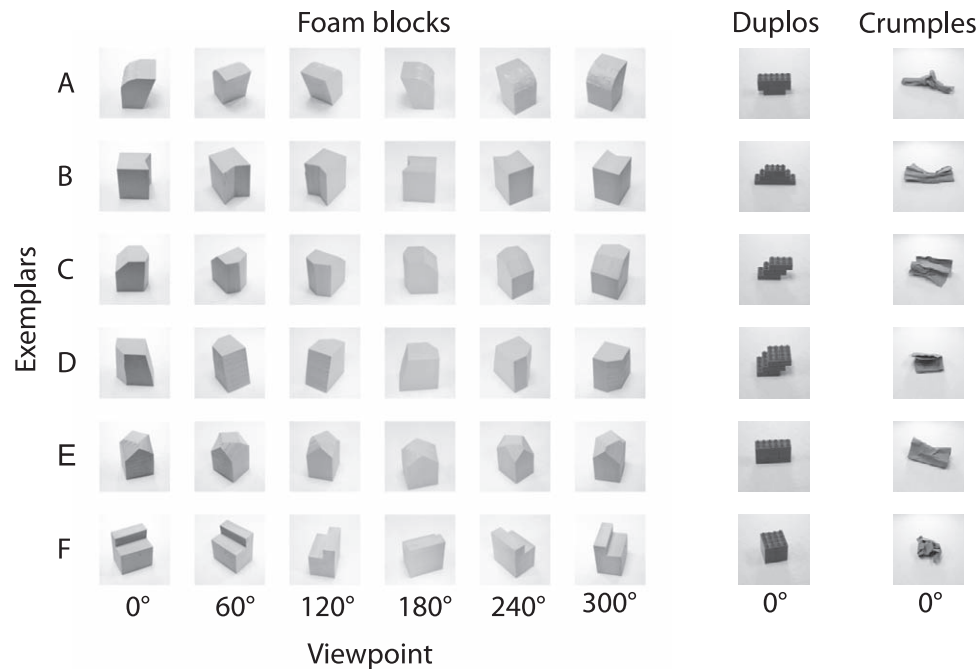
Figure 3. Example stimuli from the study. Stimuli were photographs of experimenter-made objects. For example, six different foam block stimuli can be seen in the six rows. Each object was photographed from six viewpoints (six leftmost columns). The complete set of viewpoint images is shown for one set of objects, foam blocks. The other two sets of objects used were duplos and crumples (two rightmost columns). For these two sets, only one viewpoint image is shown for each object, but the experiments used the full set of viewpoint images.

speculate that the presence of object shape information made it difficult or unintuitive for observers to use a template matching strategy. In Experiment 6, we further encouraged template matching by having observers search for simple, familiar, unrotated two-dimensional shapes. With these shapes we found unlimited-capacity results that were consistent across observers. This is the second main finding of the study, that a simple shape discrimination task can also yield unlimited-capacity results. In the Discussion section of this article, we contrast this simple task with the object shape perception tasks of the earlier experiments to better understand the locus of capacity limitations in object perception.

Experiment 1: Shape-based object recognition

This initial experiment is the central experiment for this study. The goal of Experiment 1 was to create a shape task in which accuracy relied solely on perception of an object's real-world 3-D shape. Observers attempted to locate a specific object in an array of similar objects. Object sets were created to minimize non-shape cues to object recognition. The extended

simultaneous-sequential paradigm was used to measure capacity limitations.

Method

Stimuli

Figure 3 shows example stimuli from the study. The stimulus set included three categories: foam blocks, toy duplos, and paper crumples. Each category comprised photographs of six physical models. To create the stimuli, we photographed each physical model from six evenly spaced viewpoints. The models were placed on a circular table covered in white cloth and illuminated by ambient room light and a stationary desk lamp. The objects occupied the central region of each photograph and the white tablecloth filled the background. We then applied a square crop to each image and converted all photographs to gray scale, 100×100 pixel images.

Superimposed dynamic noise

We found that conventional dynamic pixel noise was not effective in limited accuracy for this task. To create noise that was effective in limiting accuracy, we superimposed movies of scrambled stimulus images with the stimuli. To create scrambled stimulus images, we divided up a central region of each stimulus image

into squares and randomly shuffled the squares within the central region that included the object. The central region was 74 × 74 pixels for the foam block images and 50 × 50 pixels for the duplo and crumple images. Noise movies were superimposed with stimuli from their own categories using the following procedure. First, we subtracted the mean pixel value of the noise image from each individual pixel value in that image, centering the distribution of pixel values at zero. We then multiplied each pixel in both the noise image and the stimulus image by 0.7 (to prevent clipping) and combined them by taking the pixelwise sum. This gave the appearance of the stimulus image superimposed with a transparent lower contrast noise image. Each stimulus display in the experiment lasted 120 ms (nine screen refreshes), during which one noise image persisted for 40 ms (three refreshes) before changing to a new randomly chosen noise image.

Procedure

Details of the procedure are described in the General method section. Each trial consisted of a preview display followed by a critical display. The observer's task was to indicate which of four locations in the critical display contained the object that was shown in the preview and post-critical displays. The object category (foam blocks, duplos, or crumples) varied randomly from trial to trial. Distracters were always drawn from the same category of objects as the target. Importantly, different viewpoint images were used in the preview display and the critical display. Thus, observers knew what object they were searching for but not the viewpoint image that would appear in the critical display. Interval duration was 120 ms in all conditions. The apparatus was the same as in Scharff et al. (2011b).

Observers

The six observers were paid and unpaid volunteers. All had normal or corrected-to-normal vision. Author AS was an observer in this experiment and all others except Experiment 3. In this and all other experiments, each observer completed at least two practice sessions before participating.

Results

The results of Experiment 1 are depicted in Figure 4. Proportion correct is plotted against the three conditions (average of six observers with error bars corresponding to standard error of the mean across observers). The dotted lines plot the predictions of the unlimited-capacity (simultaneous = sequential) and
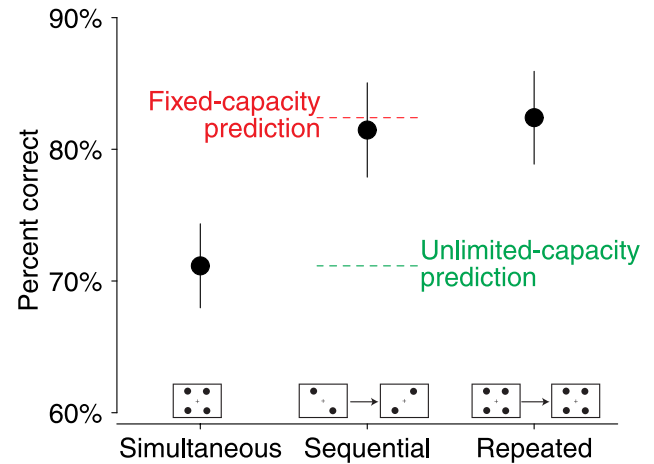

Figure 4. Results of Experiment 1. Percent correct responses in the simultaneous, sequential, and repeated conditions. Error bars are ±1 standard error of the mean over observers. The upper and lower dashed lines represent the sequential condition predictions from fixed-capacity models (sequential = repeated) and unlimited-capacity models (sequential = simultaneous), respectively. The results are consistent with fixed-capacity models and reject unlimited-capacity models.

fixed-capacity models (sequential = repeated). There was a reliable advantage for sequential over simultaneous presentation, mean within-observer difference of 10.3% ± 0.7% standard error of the mean difference, two-tailed paired-samples $t(5) = 15.59$, $p < 0.0001$, Cohen's $d = 1.24$, and for repeated over simultaneous presentation, 11.3% ± 0.3%, $t(5) = 34.8$, $p < 1.0 \times 10^{-6}$, $d = 1.37$. However, there was no reliable difference between the sequential and repeated conditions, 0.9% ± 0.8% in favor of repeated, $t(5) = 1.16$, $p = 0.29$, $d = 0.11$. This pattern of results across the three main conditions is consistent with fixed-capacity models and rejects unlimited-capacity models.

To test whether observers treated the two sequential display intervals similarly, we compared accuracy between the sequential trials with targets in the first versus the second interval. There was no reliable accuracy difference between trials with targets in the first versus the second interval of the sequential condition, 2.4% ± 3.1% in favor of second interval trials, $t(5) = 0.77$, $p = 0.48$, $d = 0.27$. If such a difference had been present, it would undermine a critical assumption of the model.

We tested whether there was an effect of the angular viewpoint difference between the preview image and the target image in the critical display. There was no reliable effect of viewpoint difference on accuracy: For 60° viewpoint differences, mean accuracy was 77% ± 3%; for 120° viewpoint differences, mean accuracy was 76% ± 3%; and for 180° viewpoint differences, mean accuracy was 75% ± 4%, $F(2, 5) = 1.12$, $p = 0.37$.

Discussion

In sum, results of Experiment 1 were consistent with fixed-capacity models and rejected unlimited-capacity models for shape-based object recognition. Secondary analyses indicated that observers treated the two sequential display intervals similarly and that image viewpoints had a negligible effect on accuracy. The central finding of this study is that shape-based object recognition has fixed capacity. The remaining experiments are controls and comparisons needed to more fully interpret this result.

Experiment 2: Identical displays with cueing

In Experiment 1, we observed a fixed capacity result for shape-based object perception. In Experiments 2, 3, and 4, we test a series of alternative explanations for the fixed-capacity limit: sensory limitations, limitations due to task difficulty, and limitations due to eye movements, respectively.

In Experiment 2, we use a cueing design to rule out any sensory limitations as causes of the limited perceptual capacity in observed in Experiment 1. The simultaneous-sequential design equates decision and response demands but not sensory demands. In particular, the simultaneous condition presents a denser array of stimuli (with four stimuli presented at once) compared to the sequential condition. Consequently, accuracy differences may have been stimulus driven (e.g., by mutual crowding of densely presented stimuli, Pelli & Tillman, 2008).

The cueing design used here eliminates physical differences between the simultaneous and sequential conditions; thus all sensory factors are held constant across conditions (e.g., crowding). Any observed effects must be attributed to attentional capacity limitations. For similar experiments see Palmer (1994) and Scharff et al. (2011a).

Method

Experiment 2 included three conditions, as depicted in Figure 5. These conditions all used two display intervals, but they are fundamentally different from the repeated condition from Experiment 1 because stimuli were not repeated. The *cued simultaneous* and *cued sequential* conditions (Panels B and C of Figure 5) replicate the corresponding conditions from Experiment 1 in terms of the number and order of relevant stimuli. The *neutral* condition (Panel A of Figure 5) was a check to ensure observers followed the cueing
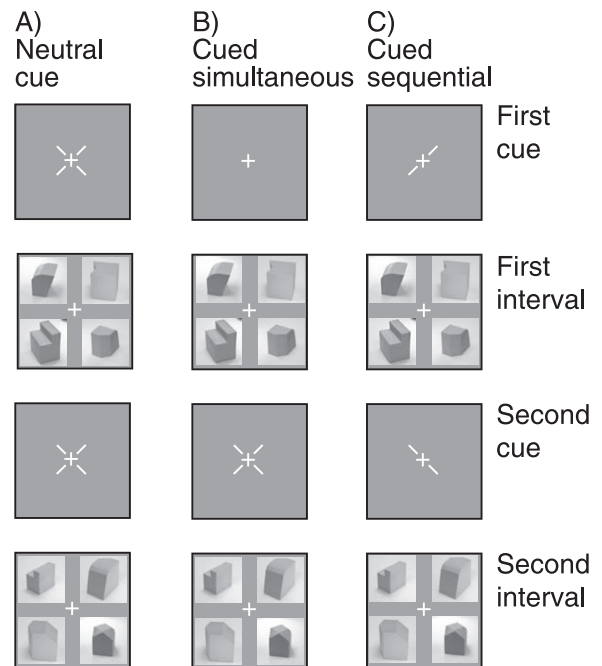

Figure 5. Schematic of conditions used in Experiment 2. This experiment repeated the simultaneous-sequential comparison using identical stimulus displays. Each condition included eight unique images, four in each interval, with only one target among them. Cues preceding each interval indicated possible target positions. In the neutral condition, the target can appear in any location in either interval. In the cued simultaneous condition, the target can appear in any location in one of the intervals (the second interval in the example shown). In the cued sequential condition, the target can occur within the cued subset of locations in either interval.

instructions. We did not include a repeated condition because it would have required inconsistent instructions and stimulus presentations. However, the repeated condition from Experiment 1 can be used as a reference.

The stimuli and display parameters were similar to the repeated condition from Experiment 1. The critical difference was that each trial included eight unique stimuli displayed across the two intervals. Each trial included one target and seven distracters with four shown in the first interval and four shown in the second. Conditions were run in separate blocks. Cueing instructions were given before each block and were reinforced with central precues pointing toward the relevant locations. Only cued locations could contain the target. In the cued simultaneous condition, cues indicated which interval could contain the target. The cued interval (first or second) was consistent within a block of trials. In the cued sequential condition, observers were told which subset of locations would contain the target—two in the first interval and two in the second. The cued locations were always at opposite

corners of the display, and the order in which they were cued was consistent within a block of trials. In the neutral condition, observers were told that the target could appear at any location and in either interval. The precues were white lines extending from near fixation towards relevant locations. Precues started 0.5° from fixation and ended 1° away. Preceding each stimulus display interval, cues appeared for 300 ms and were followed by a 300 ms blank interval.

Observers

Four observers completed the experiment, three of whom had previously completed Experiment 1. The fourth observer practiced Experiment 1 for several hours before participating. Observer AS was an observer in the study.

Predictions

If cues were ineffective (e.g., if observers ignored them), accuracy will be equivalent across all three conditions because the three conditions are identical except for the arrangement of the cues. If the accuracy advantage of the sequential over the simultaneous condition in Experiment 1 was solely attributable to perceptual capacity limits, the advantage will be undiminished in Experiment 2. If the advantage was due to sensory factors, it will be eliminated in Experiment 2. If the advantage was due to a mixture of capacity limits and sensory factors, it will be attenuated in Experiment 2.

Results

The results of Experiment 2 are depicted in Figure 6. There were reliable differences in accuracy across the cuing conditions, indicating that the cues were effective. There was a reliable advantage for cued sequential over cued simultaneous, $9.3\% \pm 0.9\%$, $t(3) = 10.46$, $p = 0.002$, $d = 1.47$. To determine whether the magnitude of the sequential-simultaneous effect was consistent with fixed capacity, we compared it to the results of Experiment 1. First, we compared accuracy in the cued-sequential condition from Experiment 2 with accuracy in the repeated condition from Experiment 1. For the three observers who participated in both experiments, there was no reliable difference between these conditions, $0.9\% \pm 2.4\%$ in favor of the cued sequential condition, $t(2) = 0.37$, $p = 0.74$, $d = 0.11$. Second, we compared the magnitude of the simultaneous-sequential difference for these observers between Experiment 1 $(10.0\% \pm 0.9\%)$ and Experiment 2 $(10.6\% \pm 1.5\%)$ and found them to be similar, $t(2) = 1.09$, $p = 0.39$, $d =$
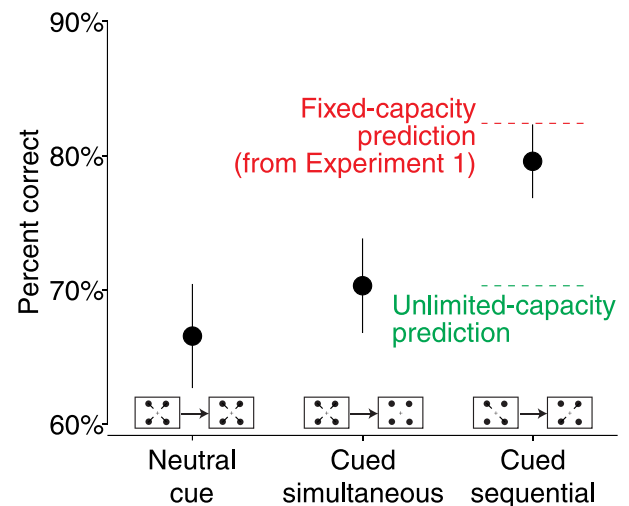

Figure 6. Results of Experiment 2. This experiment repeated the simultaneous and sequential conditions with physically equivalent displays. Cues were used to manipulate attention across conditions. The figure plots the proportion correct in the neutral, cued simultaneous, and cued sequential conditions. The upper dashed lines represent the prediction from fixed-capacity models—that cued sequential will match the repeated condition accuracy from Experiment 1. The lower dashed line represents the prediction from unlimited-capacity models that cued sequential will match cued simultaneous. The neutral condition was included to check effectiveness of the cues; ineffective cues would result in equivalent accuracy across all three conditions. This observed result is consistent with limited-capacity models and comparisons with Experiment 1 indicate that the fixed-capacity limit persists (see text for details of the comparison).

0.34. Both comparisons indicated that the effects in Experiment 2 were consistent with fixed capacity.

Discussion

In Experiment 2, we replicated the simultaneous-sequential difference from Experiment 1, indicating that the capacity limit observed is an attentional phenomenon under observer control. Because no repeated condition was used, there was no direct benchmark available to assess whether accuracy was consistent with fixed capacity. However, comparisons to results from Experiment 1 showed that the observed accuracy is consistent with fixed capacity models. Furthermore, the fact that capacity limits were unchanged by adding several irrelevant stimuli to the display indicates that observers can efficiently select which objects to submit to the limited-capacity mechanism.

Experiment 3: Varying display duration

The goal of Experiment 3 was to investigate any relationship between task difficulty and perceptual capacity. A common intuition is that difficult tasks have limited capacity, while easy tasks have unlimited capacity. However, several prior studies have demonstrated that capacity does not depend on task difficulty (e.g., Busey & Palmer, 2008; Huang & Pashler, 2005; Palmer, 1994; Palmer et al., 2000). We tested this notion by changing the display duration to manipulate difficulty. While Experiment 1 used 120 ms display durations, Experiment 3 included two different display durations, one shorter (40 ms) and one longer (240 ms) than those used in Experiment 1.

Method

Experiment 3 was identical to Experiment 1 except that different display durations were used in the critical display. Instead of 120 ms displays, 40 ms displays and 240 ms displays were used in separate blocks. As in Experiment 1, the integrated noise images were changed every 40 ms, so a single noise image was used in each display in the 40 ms condition and six noise images were used in the 240 ms condition. Experiment 3 comprised 30 sessions with 24 trials per condition at each duration in each session for a total of 720 trials per condition and duration.

Observers

Two observers participated in this experiment, one of whom had completed Experiment 1 (KE). The other observer practiced Experiment 1 for several hours before participating.

Results

The results of Experiment 3 are shown in Figure 7. Changing duration influenced overall accuracy, $F(1, 7) = 417.74$, $p < 1.0 \times 10^{-8}$, but did not alter the pattern of results compared to Experiment 1. For the 40 ms duration conditions, accuracy was better in the sequential than the simultaneous condition, Observer KE: 5.6% ± 2.7%, $t(19) = 2.05$, $p = 0.05$, $d = 0.58$; Observer YK: 5.7% ± 2.4%, $t(19) = 2.37$, $p = 0.02$, $d = 0.53$, but repeated was not reliably better than sequential, KE: 1.3% ± 2.2%, $t(19) = 0.57$, $p = 0.57$, $d = 0.14$; YK: 2.3% ± 2.1%, $t(19) = 1.06$, $p = 0.30$, $d = 0.20$.
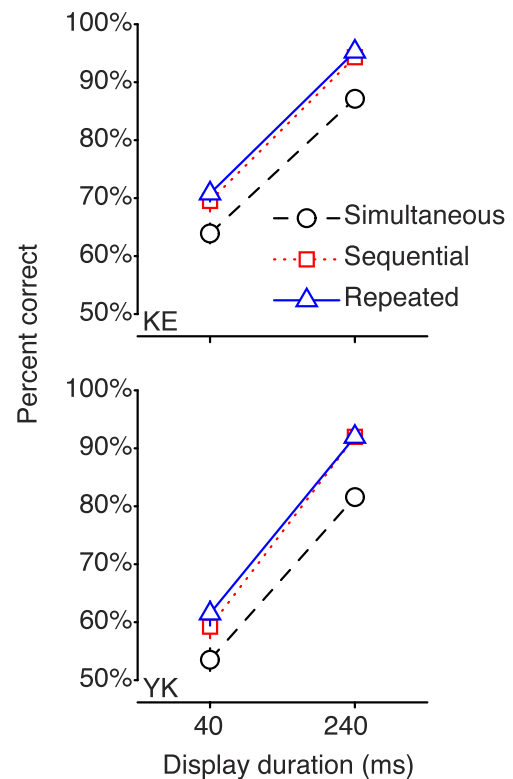

Figure 7. Results of Experiment 3. This experiment replicated Experiment 1 but used longer and shorter display durations. Proportion of correct responses in the simultaneous (circles), sequential (squares), and repeated (triangles) conditions are shown for two different display durations, 40 ms and 240 ms. The two panels depict data from different observers. Error bars signify the standard error of the mean across sessions for each observer. In some cases the error bars are smaller than the data point symbols and are hidden. All results are consistent with fixed-capacity models and reject unlimited-capacity models.

For the 240 ms duration conditions, the pattern of results was the same: Accuracy in sequential was again better than simultaneous, KE: 7.2% ± 1.5%, $t(19) = 4.85$, $p < 1 \times 10^{-4}$, $d = 1.13$; YK: 10.4% ± 1.8%, $t(19) = 5.82$, $p < 1 \times 10^{-5}$, $d = 1.54$, and repeated was not reliably better than sequential, KE: 1.0% ± 1.3%, $t(19) = 0.76$, $p = 0.45$; YK: 0.0% ± 1.5%, $t(19) = 0$ $p = 1.0$, $d = 0$. At both shorter and longer durations, accuracy is consistent with the predictions of fixed-capacity models. This result reinforces the notion that perceptual capacity is independent of task difficulty.

Experiment 4: Eye-movement control

In the other experiments presented here, we instructed observers to maintain eye fixation throughout each trial. Fixation was further encouraged by brief

Observer	Percent aborted trials (%)	Mean horizontal position (°)	Mean vertical position (°)	SD horizontal position (°)	SD vertical position (°)
AS	3.6	N.A.	N.A.	N.A.	N.A.
JS	7.5	0.1	−1.0	1.6	1.5
JW	10.1	−0.2	−0.8	1.3	1.4
RV	7.2	0.1	−1.3	1.5	1.3
ER	4.7	−0.3	0.7	1.2	1.7
MS	4.2	0.1	2.3	−1.6	1.7
Mean	6.2	0.0	0.0	0.8	1.5

Table 1. Fixation data.

display durations and stimuli that presented on opposite sides of fixation. However, if these measures were ineffective in controlling fixation, observers might make eye movements to fixate on different stimuli as they are presented. This might introduce a fixed-capacity limit on perception of objects because foveated objects will be perceived more clearly than peripheral objects and only one object can be fixated at a time. This would make object recognition dependent on a serial mechanism. To investigate whether eye movements caused the fixed-capacity limit, we repeated Experiment 1 while using an eye tracker to enforce fixation. If eye movements were solely responsible for the fixed capacity result, we would expect unlimited-capacity results in this experiment.

Methods

We monitored eye position using a noninvasive video system (EyeLink II, Version 2.11, SR Research, Osgoode, Ontario, Canada) controlled by a separate DOS computer. The EyeLink II is a binocular, head-mounted, infrared video system with 250 Hz sampling. It was controlled by the EyeLink Toolbox extension of MATLAB Version 1.2 (Cornelissen, Peters, & Palmer, 2002). We recorded and analyzed only the right eye position. The system has a resolution of at least 0.1° for detecting saccades and a resolution of about 1.0° for sustained eye position over many trials.

At the beginning of each trial, only the fixation cross was present on screen. Observers were instructed to fixate the cross at the beginning of each trial. Once the observer held a constant eye position for 500 ms, the trial was initiated and the fixated coordinates were used as the baseline position for that trial. The trial-to-trial stability of the eye tracking can be judged by the means and standard deviations of the baseline position coordinates, which are reported in Table 1. During each trial, eye position was measured directly before and after each stimulus presentation. If the eye position was measured as being more than 1.0° from baseline, an error tone played and the trial was excluded from analysis. This resulted in 6% aborted trials overall: 5%

in the simultaneous, 6% in sequential, and 8% in repeated. More aborted trials are expected in the sequential and repeated conditions because they are longer in duration than the simultaneous condition, giving more opportunities for observers to break fixation.

Observers

Six observers participated in the experiment. Three had already completed Experiments 1, 2, and 5 (AS, JS, RV). One had completed Experiment 5 (JW), and one had completed Experiment 2 (ER). In addition, observers JW and ER had each practiced Experiment 1 for several hours.

Results

Results of Experiment 3 are plotted in Figure 8. There was a reliable advantage for sequential over simultaneous, $8.0\% \pm 1.4\%$, $t(5) = 5.80$, $p = 0.002$, $d = 1.02$, and for repeated over simultaneous presentation, $10.0\% \pm 1.4\%$, $t(5) = 7.01$, $p = 0.001$, $d = 1.41$. There was no reliable difference between repeated and sequential, $2.0\% \pm 1.3\%$, $t(5) = 1.57$, $p = 0.16$, $d = 0.30$. This pattern of results rejects unlimited-capacity models and is consistent with fixed-capacity models. In sum, Experiment 4 replicated Experiment 1 with eye-movements controlled, indicating the fixed-capacity limit is not caused by eye movements.

Experiment 5: Recognition with fixed viewpoint

Having established that the 3-D object shape task yields fixed capacity, we sought a complementary case that would yield a contrasting unlimited-capacity result. Our goal was to obtain an unlimited-capacity result while using methods and stimuli similar to those used in the preceding experiments. This approach
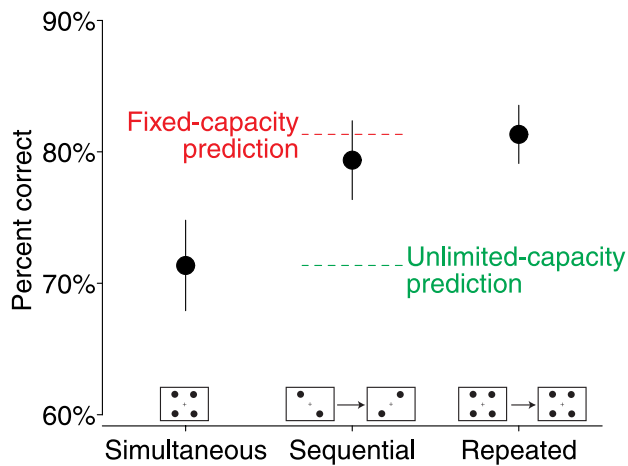
Figure 8. Results of Experiment 4. This experiment replicates Experiment 1 but with eye position monitored. Trials in which observers broke fixation were excluded from analysis. Proportion of correct responses are shown for the simultaneous, sequential, and repeated conditions. Consistent with the earlier experiments, these results support fixed-capacity models and reject unlimited-capacity models.
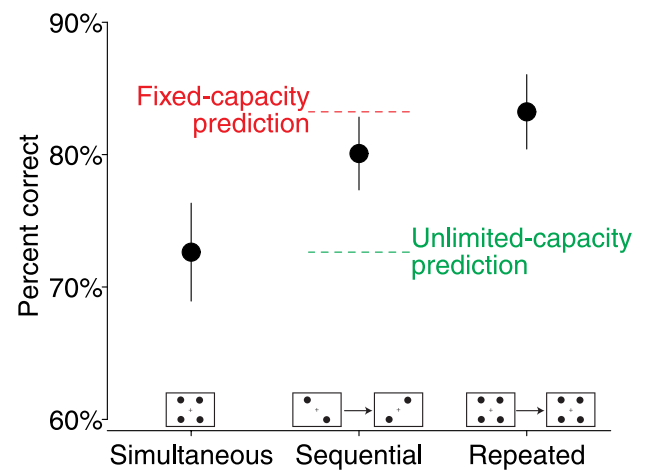

Figure 9. Results of Experiment 5. This experiment replicated Experiment 1 but had fixed viewpoints in the critical display and a shorter 40 ms display duration. Proportion of correct responses are shown for the simultaneous, sequential, and repeated conditions. These results are consistent with fixed-capacity models and reject unlimited-capacity models.

enables identification of the critical factors that causes the capacity limit in object perception.

For Experiment 5, we used the stimuli from Experiment 1 but did not vary the viewpoint of the target object between the preview and critical displays. Observers were encouraged to adopt feature matching or template matching strategies. For example, the observer could search for a particular oriented edge instead of a particular object, effectively performing the task as an orientation search rather than an object search. We speculated that the ability to apply a feature matching or template matching strategy would result in unlimited-capacity processing.

Methods

Experiment 5 replicated Experiment 1, with two differences. First, the preview and critical displays of the target were always the same viewpoint image. Second, critical displays had 40 ms display durations instead of 120 ms display durations. We made this change because in practice sessions with 120 ms display durations, overall accuracy was close to the ceiling. The change in display duration made overall accuracy similar to Experiment 1. Observers were encouraged to exploit image features such as size, contrast, and image shape.

Observers

Six observers completed the experiment. Two observers had previously completed Experiments 1 and

2 (JS, AS), two observers had completed Experiment 2 (RV, ER), and one observer had completed Experiment 1 (AP). A sixth observer had practiced Experiment 1 for several hours (JW).

Results

Figure 9 shows the results of Experiment 5 pooled across observers. Overall, the results were consistent with fixed-capacity models and rejected unlimited-capacity models. Simultaneous accuracy was reliably inferior to both sequential, 7% ± 2%, $t(5) = 4.83$, $p = 0.005$, $d = 0.94$, and repeated accuracy, 11% ± 2%, $t(5) = 5.41$, $p = 0.005$, $d = 1.32$. Repeated accuracy was not reliably different from sequential accuracy, 3% ± 2% in favor of repeated, $t(5) = 1.78$, $p = 0.14$, $d = 0.46$.

Discussion

Using a fixed viewpoint made the task easier than using an unpredictable viewpoint—this we can see by comparing the mean accuracy in Experiment 5 (79%) with the mean accuracy in the 40 ms condition used in Experiment 3 (63%). Despite making the task easier, the fixed viewpoint did not alleviate capacity limits. Results here showed fixed capacity for the typical observer.

This finding contradicts our predicted result, that the presence of predictable image features would allow the observers to circumvent fixed-capacity limits using a template matching strategy. Perhaps observers adopted a shape constancy strategy by default because it is an

ecologically valid interpretation of the stimulus. More generally, this result demonstrates that viewpoint uncertainty is not necessary to obtain a fixed capacity result.

Experiment 6: Image shape recognition

All prior experiments showed fixed capacity results. In Experiment 6, we attempted to obtain a contrasting unlimited-capacity result using an experiment similar to Experiment 1. Achieving an unlimited capacity result in a similar experiment would allow us to rule out any factors that the experiments have in common as causes of the fixed-capacity result in Experiment 1. Furthermore, obtaining both fixed- and unlimited-capacity results using similar procedures validates the procedure as capable of distinguishing the two model classes (Scharff et al., 2011b).

To encourage feature and template matching strategies, we further simplified the stimuli, using simple 2-D shapes. We anticipated that the availability of image features would yield unlimited capacity performance.

Methods

Methods were similar to Experiment 1 except as follows. Instead of photographed object stimuli, we used six simple shapes created with a vector graphics program. These shapes were an arch, circle, crescent, square, star, and triangle (shown in Panel A of Figure 10). The shapes were black on a white background and were approximately equivalent in area. Each shape was originally drawn with a vector graphics program and then converted to a 100×100 raster image. The shapes were presented at reduced contrast to induce an appropriate degree of difficulty (contrast $= 0.05$ for most observers, 0.07 for observer JP). Stimuli were never rotated. Critical displays lasted 120 ms with dynamic noise as described in Experiment 1. Noise images were 4×4 checkerboard scrambles of the complete stimulus images reduced in contrast by a factor of 0.07. We manipulated difficulty by reducing contrast rather than duration because the dynamic noise was less effective at impairing accuracy at short durations.

Observers

Six observers completed the experiment. Most observers had already participated in Experiments 1, 2, 4, and/or 5. Author JP was an observer and had not completed the previous experiments.
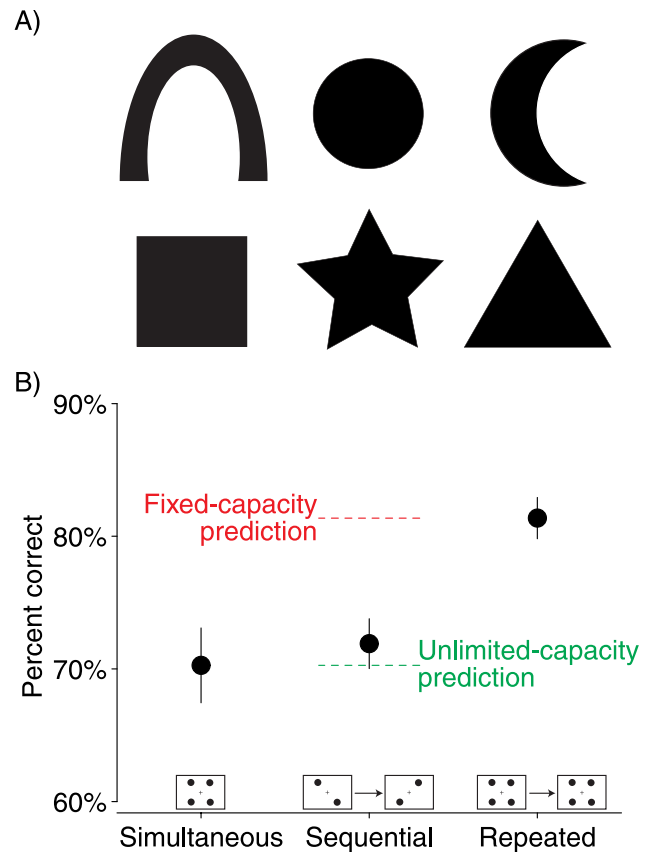

Figure 10. (A) Stimuli from Experiment 6. Six unrotated 2-D shapes were used as stimuli in this experiment. (B) Results of Experiment 6. Proportion of correct responses are shown for the simultaneous, sequential, and repeated conditions. These results are consistent with unlimited-capacity models and reject fixed-capacity models.

Results

The results of Experiment 6 are depicted in Panel B of Figure 10. There was no reliable difference between simultaneous and sequential, $1.6\% \pm 1.4\%$, $t(5) = 1.15$, $p = 0.30$, $d = 0.28$. There was a reliable advantage for repeated over sequential, $9.5\% \pm 0.9\%$, $t(5) = 10.21$, $p < 0.001$, $d = 2.25$, and for repeated over simultaneous, $11.1\% \pm 1.7\%$, $t(5) = 6.55$, $p = 0.001$, $d = 1.99$. There was no reliable accuracy difference between trials with targets in the first versus the second interval of the sequential condition, $2.0\% \pm 2.3\%$ in favor of second interval trials, $t(5) = 0.87$, $p = 0.42$, $d = 0.36$.

Discussion

This pattern of results is consistent with unlimited-capacity models and rejects fixed-capacity models. Obtaining this result with the image shape task

supports the notion that unlimited-capacity performance is possible when diagnostic-image features are readily available. In the Discussion section we consider the differences and similarities between this experiment and the others that produced fixed-capacity results.

Effect-size analyses

The analyses above relied on two key equality predictions: equality of simultaneous and sequential accuracy for unlimited-capacity models and equivalence of sequential and repeated accuracy for fixed-capacity models.

These two equalities have been our main focus because they rely on few assumptions, lending generality to our findings.

Here, we introduce quantitative models that can predict these equalities and compare their specific predictions with our results. Our goal in doing so is to fully justify the equality predictions used above and secondarily to explore the range of models that can make these predictions. These models illustrate how the fixed-capacity equality predictions described above are implemented by both serial and parallel models. Both kinds of models have fixed capacity in that they assume a discrete limit on the rate of information processing. The experimental results we observed are shown to be quantitatively similar to the predictions of both kinds of fixed-capacity models.

Figure 11 summarizes the data in comparison to the predictions of several quantitative models. In Panel A of Figure 11, sequential accuracy is plotted against simultaneous accuracy. In Panel B, repeated is plotted against simultaneous. Results from each experiment are plotted as colored markers. These markers indicate mean accuracy across observers with error bars representing standard error of the mean across observers. Model predictions are plotted as curves on each plot. These are quantitative models that make the equality predictions. Model details are given in the Appendix.

Each model has a difficulty parameter (representing target-distracter discriminability in terms of signal-to-noise ratio) that sweeps out the curve, with the most difficult discriminations in the lower-left corner and the easiest in the upper-rightmost point on each curve. The standard serial model has an additional parameter m, reflecting the number of stimuli that the observer can scan during a brief stimulus display. The models assume that an observer cannot scan more stimuli than are present in a display. Each of these assumes that an observer can scan through m stimuli during the brief
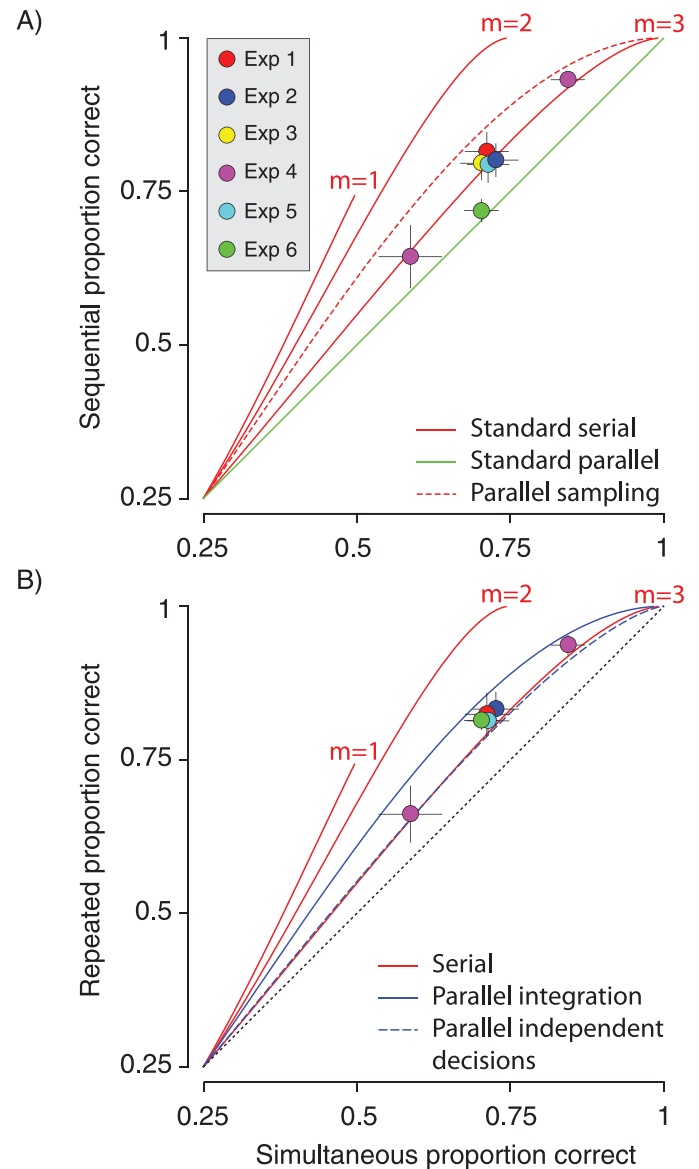

Figure 11. Comparison of experimental data with model predictions. (A) Models and data for simultaneous-sequential difference. Models predictions are curves swept out by discrimination difficulty. Plotted points reflect means and standard errors across observers for the six experiments (E1–E6). Experiment 3 is split into the 40 ms and 240 ms display durations. (B) Models and data for the simultaneous-repeated difference. Experiment 2 is not shown because it did not have a repeated condition.

stimulus display, and has no knowledge of the $4 - m$ unscanned stimuli.

In Panel A of Figure 11, simultaneous-sequential differences are plotted for all six experiments. Experiment 3 is split into the 40 ms and 240 ms display conditions. Fixed-capacity models are shown as red curves. The standard serial models are plotted as solid red curves (with values of $m = 1$, 2, and 3). A parallel sampling model is plotted as a dashed red curve. This

model assumes a statistical sampling framework for the quality of perception, in which perceptual samples are divided evenly among the number of available samples. Finally, a standard parallel is plotted as a green curve. This model assumes that the quality of perception is unaffected by divided attention manipulations.

By assessing the data relative to the model predictions for the simultaneous-sequential predictions in Panel A, it is apparent that Experiments 1, 2, 4, and 5 produced indistinguishable results, each within a standard error of the others (red, yellow, purple, and cyan markers). This is reasonable, because Experiments 1 (shape-based object recognition), 2 (cueing control), and 4 (eye-tracking control) are essentially replications of each other. Experiment 5 (fixed viewpoint) also produced a somewhat similar result. These four experiments produced limited-capacity results consistent with a serial model with $m = 3$, but the other fixed-capacity models could fit the data with the addition of lapse rate or another feature to attenuate effect size (see Appendix). Experiment 6 (simple shape recognition) produced the only result consistent with the standard parallel model. The duration manipulation in Experiment 3 caused diagonal shifts that roughly parallel the model prediction curves. Diagonal shifts are consistent with a change in the target-distracter discriminability with longer display durations yielding easier discriminations. The diagonal shift is inconsistent with the simple serial model in which the m parameter is determined by viewing time, which would predict horizontal shifts towards curves generated by models with different m values.

Panel B of Figure 11 depicts data and predictions for the simultaneous-repeated comparison. The three curves for serial models are generated by assuming the repeated condition yields information about $2 * m$ stimuli (with a maximum value of four for m), where the simultaneous condition yields information about only m stimuli. Two parallel model predictions are plotted as blue curves. Because there is no divided attention manipulation between the simultaneous and repeated conditions, the parallel sampling and standard parallel models make identical predictions. Two types of parallel models are shown, differing in how the observer pools information across the two displays: integration or independent decisions (see Appendix for details). Overall, all data are consistent with the predicted effect sizes from several models. The observed simultaneous-repeated effect sizes could be predicted by the parallel independent decisions model, the standard serial model with $m = 3$, or a modified parallel integration model. Notably, the duration manipulation in Experiment 3 also produced a diagonal shift on this plot consistent with a change in signal-to-noise ratio and inconsistent with a simple serial model

in which display duration determines how many stimuli can be scanned.

This analysis demonstrates that the results are fully compatible with two simple implementations of fixed capacity: a serial version and a parallel version. Notably, the duration manipulations failed to reveal an obvious signature of serial processing.

Discussion

This article reports two main findings. First, we established fixed-capacity attentional limitations in a 3-D object shape perception task, a result that was replicated several times within this study (Experiments 1–4). Second, we altered the task in hopes of finding a contrast case with unlimited-capacity results (Experiments 5 and 6). We did find a contrast case in Experiment 6, in which a simple 2-D shape task yielded unlimited-capacity performance.

Considering alternative explanations for the fixed-capacity limit

Our goal was to assess the impact of divided attention on object shape perception. Based on our findings, we propose that perception of object shape perception is severely limited by divided attention. In the paragraphs that follow, we discuss several alternative explanations that we believe can be ruled out.

Eye movements or sensory interference

In Experiment 2, we ruled out any explanations based on sensory-level interference between stimuli by replicating the fixed-capacity result using physically identical displays. In Experiment 4, we showed that the capacity limit was not caused by eye movements.

Interference during decision or response processes

The design of the extended simultaneous-sequential paradigm rules out any effects on decision or response processes, because the decision and response are equated across conditions.

Limits on semantics or categorization

The design of the stimulus and search task was intended to minimize demand for semantic categorization of stimuli. Observers always searched for a nameless physical object among an array of similar objects.

Memory

The sequential condition requires the observer to retain information over a longer period of time than the simultaneous condition. However, in all cases we found that accuracy in the sequential condition was equivalent between trials with targets in the first interval versus the second interval, indicating that memory demands over this time scale had no reliable effect on accuracy. Memory demands were otherwise consistent across conditions.

Task difficulty

It may be intuitive to think that difficult judgments have limited capacity while easy judgments have unlimited capacity. However, our findings are agree with previous studies in finding that perceptual capacity is independent of task difficulty (Busey & Palmer, 2008; Huang & Pashler, 2005; Palmer, 1994; Palmer et al., 2000). When we manipulated stimulus display durations in Experiment 3, we observed no effect on capacity despite large effects on overall accuracy. Furthermore, Experiments 1 and 6 had similar levels of overall accuracy in spite of having opposite capacity outcomes.

Learning

To determine whether the degree of capacity limitation changes as observers became more familiar with the task, we calculated linear regressions between the number of sessions completed and the simultaneous-minus-sequential difference for each session. For Experiments 1, 2, 3, 4, and 6, we found no reliable relationships between session number and simultaneous-minus-sequential difference. In Experiment 4 we did find a reliable relationship between session number and the simultaneous-minus-sequential difference. The slope of the regression line was −0.007, reflecting a small reduction in the simultaneous-sequential difference with each session of practice. However, this effect is not reliable once a correction for multiple comparisons is applied (one-sample t test on slope of regression line, uncorrected $p = 0.04$). Overall, there is no evidence that learning alleviates capacity limits in these experiments.

Automaticity

The automaticity theory of Schneider and Shiffrin (1977) posits that capacity limits are invoked when varied mapping of stimuli to target-distracter designations are used in a task. Varied mapping refers to a design in which a given stimulus can appear as either a target or a distracter stimulus in contrast with consistent mapping, in which a given stimulus is always a target or always a distracter. All experiments presented here were variable mapping experiments. However, the simple shape task in Experiment 6 also had variable mapping yet yielded an unlimited-capacity result. We conclude that automaticity theory does not adequately explain our findings.

Heterogeneity of distracters

A recent article by Mazyar, van den Berg, and Ma (2012) proposed that the critical factor in determining capacity limits was heterogeneity of distracters. The article supported this hypothesis by appealing to prior literature and new findings in which search tasks with heterogeneous distracters always yield limited capacity results. However, we obtained both unlimited-capacity (Experiment 6) and fixed-capacity results (Experiments 1–4) using heterogeneous distracters. Similarly, earlier studies have also found unlimited-capacity results using search tasks with heterogeneous distracters (Eckstein, 1998; Eckstein et al., 2000; Huang & Pashler, 2005). These data present a challenge for the hypothesis that heterogeneity of stimuli is the determining factor in whether or not search tasks have capacity limits.
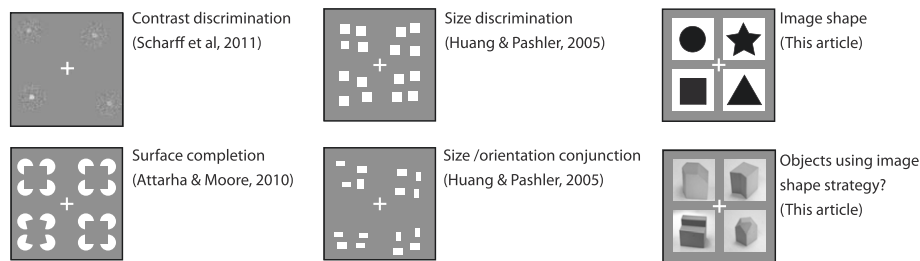
Mental rotation

Another possible explanation for the observed fixed-capacity limit in object perception is that matching stimuli to memory standards requires mental rotation of object percepts (Shepard & Metzler, 1971). Our data do not support this notion. First, in Experiment 5 we found that fixing the viewpoint of the critical display did not alleviate capacity limits even though no mental rotation was required. Second, we found that accuracy in the object task did not differ reliably across the range of angular differences between preview and critical display viewpoints (this analysis is described in the results section of Experiment 1). We conclude that mental rotation did not cause the fixed-capacity limit.

Interpreting the fixed-capacity limit

Above we ruled out many alternatives for the fixed capacity limit in the object shape perception task. If we accept that some aspect of object shape perception is the locus of the observed capacity limit, we are left with the question of what it is about the task that induces this limit. As discussed earlier, we can address this question by contrasting these fixed-capacity tasks with similar tasks that do not yield capacity limits. Here we consider the contrasts between Experiment 1 (shape-based object recognition, fixed capacity) and Experiment 6 (simple image shape recognition, unlimited capacity).

Sensory and segmentation processes have unlimited capacity:

Contrast discrimination (Scharff et al, 2011)

Size discrimination (Huang & Pashler, 2005)

Image shape (This article)

Surface completion (Attarha & Moore, 2010)

Size /orientation conjunction (Huang & Pashler, 2005)

Objects using image shape strategy? (This article)

Object and semantic processes have fixed-capacity:

Object shape (This article)

Categorization (Scharff, Palmer & Moore, 2011b)

Reading (Scharff, Palmer & Moore, 2011a)

Figure 12. When is perception limited by attention? Studies using the simultaneous-sequential method have shown that many perceptual tasks that rely on sensory and segmentation processes are unaffected by divided attention (i.e., have unlimited capacity). Several tasks that demand object and semantic processes yielded results consistent with fixed-capacity models, indicating a restrictive attentional bottleneck on perception of object shape and semantics.

There are at least six differences between these two experiments that could account for the fixed-capacity result in Experiment 6:

- Discriminability of targets and distracters
- Display duration
- Variable versus fixed viewpoint
- 3-D objects versus 2-D objects
- Complex stimuli versus simpler stimuli
- Number of objects and images in a set

The first two ideas can be ruled out as the critical factors for fixed-capacity processing. Prior research has shown that discriminability (Item 1 in the list above) is not a critical factor, as several prior studies have yielded unlimited-capacity results even for stimuli that are very difficult to discriminate (see discussion of task difficulty above). The display duration was not equivalent between these two experiments (Item 2: 120 ms in Experiment 1 vs. 40 ms in Experiment 6), but in Experiment 3 we replicated the fixed-capacity result for shape-based object perception using the same duration 40 ms used in Experiment 6.

The remaining items on the list are all reasonable hypothetical causes of fixed-capacity processing. The mixed results of Experiment 5 tell us that while variable versus fixed viewpoint (Item 3) and 3-D versus 2-D (Item 4) may play a role in inducing the fixed-capacity limit, but the nature of that role is not straightforward and may be subject to observer strategies. Finally, complexity of stimuli (Item 5) and the number of stimuli in the set (Item 6) are worthy candidates for future study.

Results of a prior study further constrain the question of where capacity limits arise in the processing of images into object percepts. Attarha et al. (Attarha & Moore, 2010; Attarha, Moore, Scharff, & Palmer, 2013) used the extended simultaneous-sequential paradigm to show that *surface completion* processes have unlimited capacity (see also He & Nakayama, 1992; Keane, Mettler, Tsoi, & Kellman, 2011). Surface completion refers to the process of recovering object surfaces from images in which those surfaces occlude each other. This processing may be a first necessary step in parsing an image into functional units for object shape perception (Pizlo, 2008). Thus, we can place the bottleneck restricting object perception somewhere between early image segmentation and full object shape perception.

When is perception limited by attention?

By considering the experiments reported here in context with other divided attention studies, we can begin to formulate a general theory of divided attention. In Figure 12, we show example displays from several recently published simultaneous-sequential experiments using a variety of tasks. Five tasks that rely on *sensory and segmentation* processes have yielded results consistent with unlimited-capacity models. We propose that perceptual tasks that require interpretation of image features, as well as some degree of perceptual segmentation of images, can be accomplished without effects of divided attention. In addition, we included the object task with fixed viewpoint in

this category as a reminder that tasks that nominally rely on more complex processes may be completed with unlimited capacity when image features are available. In contrast, three kinds of tasks yielded results consistent with fixed-capacity: object shape recognition, object categorization, and word reading. We propose that the *object and semantic processes* required by these tasks are limited by a fundamental bottleneck in perception. A lingering question is whether semantic demands of the categorization and reading tasks would be sufficient to induce capacity limitations or if the capacity limitations observed in these tasks are caused solely by their object processing demands.

Theories of object recognition

The object shape experiments presented here were intended to test the mechanisms of object shape perception under divided attention. Specifically, we addressed perception of 3-D object shape rather than the perception of 2-D shapes that are directly represented in the retinal image. In this regard, our study is tied to the literature on *shape constancy*, which refers to the fact that veridical 3-D object shape is often perceived in spite of variations in the retinal image.

The mechanisms of shape constancy have been a subject of much research. Alternative theories proposed to explain shape constancy can be broadly classified as either structural description theories or view-based theories. In structural description models, shape perception involves constructing a view-invariant representation of an object from an input image (e.g., Biederman & Gerhardstein, 1993; Pizlo, 1994). In view-based theories, shape perception involves comparing the input image to a discrete set of learned view images stored in memory (e.g., Bülthoff & Edelman, 1992; Tarr & Bülthoff, 1998). Hybrid theories combine aspects of view-based and structural description models (Foster & Gilson, 2002; Hayward, 2003; Hummel, 2001).

The present study was not intended to distinguish between alternative accounts of shape perception. Rather, the goal was to characterize the effects of divided attention on the mechanisms of shape perception, whatever the details of those mechanisms. Thus, the fixed-capacity limit is compatible with any of the above theories. For example, divided attention may limit either the formation of structural descriptions or the comparison of input images with stored view images.

Comparison with selective attention effects

Selective attention refers to situations in which a subset of stimuli are task-relevant and other stimuli are task-irrelevant (or less relevant). Studies of selective attention often have the goal of characterizing the differences between the perception relevant and irrelevant stimuli. Rock and Gutman's (1981) study of shape perception and selective attention complements the present study. They showed that memory for task-irrelevant shape stimuli was obliterated, as no observers reported seeing the irrelevant shapes. In Rock and Gutman's study and in the present study, the effect of attention on shape perception was as large as could reasonably be expected. Selective attention and divided attention are sometimes thought of as two phenomena stemming from one central limitation (Pashler, 1999) and that could certainly be the case in shape perception: Because shape perception has such a limited capacity, task-relevant shapes are prioritized for perception while irrelevant shapes are left unprocessed.

Conclusion

This study demonstrates capacity limits in 3-D object shape perception. More specifically, the degree of limitations is consistent with fixed-capacity models, including the standard serial and parallel sampling models. In contrast, the recognition of simple shapes that had a specific target image had unlimited capacity. Having ruled out several alternative explanations, we argue that the process of recovering generalized 3-D object shape from the 2-D retinal projection is subject to a fixed-capacity bottleneck.

Acknowledgments

Corresponding author: Alec Scharff.
Email: scharff@uw.edu.
Address: Department of Psychology, University of Washington, Seattle, Washington, USA.

References

Allport, D. A. (1987). Selection for action: Some behavioral and neurophysiological considerations

of attention and action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 395–419). Hillsdale, NJ: Erlbaum.

Attarha, M., & Moore, C. M. (2010). Testing capacity limitations of surface completion using the simultaneous-sequential method [Abstract]. *Visual Cognition, 18*(10), 1487–1491.

Attarha, M., Moore, C. M., Scharff, A., & Palmer, J. (2013). Surface completion engages online unlimited-capacity processes. Manuscript submitted for publication.

Biederman, I., Blickle, T. W., Teitelbaum, R. C., & Klatsky, G. J. (1987). Object search in nonscene displays. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 456–467.

Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance, 19*(6), 1162–1182. doi:10.1037/0096-1523.19.6.1162.

Bonnel, A.-M., Stein, J.-F., & Bertucci, P. (1992). Does attention modulate the perception of luminance changes? *The Quarterly Journal of Experimental Psychology, 44*(4), 601–626.

Broadbent, D. E. (1958). *Perception and communication.* Elmsford, NY: Pergamon Press, Inc.

Bülthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences, 89,* 60–64.

Busey, T., & Palmer, J. (2008). Set-size effects for identifications versus localization depend on the visual search task. *Journal of Experimental Psychology: Human Perception and Performance, 34*(4), 790–810.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The eyelink toolbox: Eye tracking with MATLAB and the psychophysics toolbox. *Behavioral Research Methods, 34*(4), 613–617.

Davis, E. T., Shikano, T., Peterson, S. A., & Michel, R. K. (2003). Divided attention and visual search for simple versus complex features. *Vision Research, 43*(21), 2213–2232.

Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature, 369,* 313–315.

Eckstein, M. P. (1998). The lower visual search efficiency for conjuctions is due to noise and not serial attentional processing. *Psychological Science, 9*(2), 111–118, doi:10.1111/1467-9280.00020.

Eckstein, M. P., Peterson, M. F., Pham, B. T., & Droll, J. A. (2009). Statistical decision theory to relate neurons to behavior in the study of cover visual attention. *Vision Research, 49*(10), 1097–1128, doi: 10.1016/j.visres.2008.12.008.

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception and Psychophysics, 62*(3), 425–451.

Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology, 79*(2, Part 2), 1–16, doi:10.1037/h0026873.

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: Is it really attention free? *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1476–1492, doi:10.1037/0096-1523.31.6.1476.

Foster, D. H., & Gilson, S. J. (2002). Recognizing novel three-dimensional objects by summing signals from parts and views. *Proceedings of the Royal Society B: Biological Sciences, 269*(1503), 1939–1947, doi:10.1098/rspb.2002.2119.

Gardner, G. T. (1973). Evidence for independent parallel channels in tachistoscopic perception. *Cognitive Psychology, 4,* 130–155.

Hayward, W. G. (2003). After the viewpoint debate: Where next in object recognition? *Trends in Cognitive Sciences, 7*(10), 425–427, doi:10.1016/j.tics.2003.08.004.

He, J. H., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature, 359,* 231–233.

Huang, L., & Pashler, H. (2005). Attention capacity and task difficulty in visual search. *Cognition, 94*(3), B101–B111, doi:10.1016/j.cognition.2004.06.006.

Hummel, J. E. (2001). Complementary solutions to the binding problem in vision: Implications for shape perception and object recognition. *Visual Cognition, 8,* 489–517.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

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

Keane, B. P., Mettler, E., Tsoi, V., & Kellman, P. J. (2011). Attentional signatures of perception: Multiple object tracking reveals automaticity of contour interpolation. *Journal of Experimental*

Psychology: Human Perception and Performance, 37(3), 685–698.

Mazyar, H., van den Berg, R., & Ma, W. J. (2012). Does precision decrease with set size? *Journal of Vision, 12*(6):10, 1–16, http://www.journalofvision. org/content/12/6/10, doi:10.1167/12.6.10. [PubMed] [Article]

Mozer, M. (1991). *The perception of multiple objects.* Cambridge, MA: MIT Press.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44–64.

Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect. *Current Directions in Psychological Science, 4*(4), 118–123.

Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research, 34*(13), 1703–1721, doi:10.1016/0042-6989(94)90128-7.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40*(10-12), 1227–1268.

Pashler, H. (1999). *The psychology of attention.* London: MIT.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience, 11*(10), 1129–1135. doi:10.1038/nn.2187.

Pizlo, Z. (1994). A theory of shape constancy based on perspective invariants. *Vision Research, 34*(12), 1637–1658, doi:10.1016/0042-6989(94)90123-6.

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

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

Rock, I., & Gutman, D. (1981). The effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance, 7*(2), 275–285, doi:10.1037/0096-1523.7.2.275.

Rousselet, G. A., Fabre-Thrope, M., & Thorpe, S. J. (2002). Parallel processing in high-level categorization of natural images. *Nature Neuroscience, 5*(7), 629–630.

Rousselet, G. A., Thorpe, S. J., & Fabre-Thrope, M. (2004). How parallel is visual processing in the ventral pathway? *Trends in Cognitive Sciences, 8*(8), 363–370.

Scharff, A., Palmer, J., & Moore, C. M. (2011a). Evidence of fixed capacity in visual object categorization. *Psychonomic Bulletin & Review, 18*(4), 713–721, doi:10.3758/s13423-011-0101-1.

Scharff, A., Palmer, J., & Moore, C. M. (2011b). Extending the simultaneous-sequential paradigm to measure perceptual capacity for features and words. *Journal of Experimental Psychology: Human Perception and Performance, 37*(3), 813–833, doi:10.1037/a0021440.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 81*(1), 1–66.

Shaw, M. L. (1980). Identifying attentional and decision-making components in information processing. *Attention and Performance, 8.*

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701–703, doi:10.1126/science.171.3972.701.

Shiffrin, R. M., & Gardner, G. T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology, 93*(1), 72–82, doi:10.1037/h0032453.

Sperling, G., & Melchner, M. J. (1978). The attention operating characteristic: Examples from visual search. *Science, 20*(202), 315–318.

Tanner, W. P. (1961). Physiological implications of psychophysical data. *Annals of the New York Academy of Sciences, 89,* 752–765.

Tarr, M. J., & Bülthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Cognition, 67,* 1–20.

Townsend, J. T. (1974). Issues and models concerning the processing of a finite number of inputs. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* (pp. 133–168). Hillsdale, NJ: Erlbaum.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology, 12*(1), 97–136.

Van der Heijden, A. H. C. (1996). Two stages in visual information processing and visual perception? *Visual Cognition, 3*(4), 325–361.

Warrington, E. K., & Taylor, A. M. (1978). Two categorical stages of object recognition. *Perception, 7,* 695–705.

Wichmann, F. A., Drewes, J., Rosas, P., & Gegenfurtner, K. R. (2010). Animal detection in natural scenes: Critical features revisited. *Journal of Vision, 10*(4):6, 1–27, http://www.journalofvision.org/content/10/4/6, doi:10.1167/10.4.6. [PubMed] [Article]

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1*(2), 202–238.

Appendix

Introduction

This article describes two kinds of predictions: the equality predictions emphasized in the body of this article and the more specific effect size predictions. In this appendix, the first topic is the derivation of effect size predictions for the three common models discussed in the body of the: standard parallel, standard serial, and parallel sampling models. The second topic is a discussion of a wider range of models that make the equality predictions: either sequential = simultaneous for unlimited-capacity models or sequential = repeated for fixed-capacity models. This is a first step toward identifying the larger domain of models that make these equality predictions.

Predicted sequential and repeated effects for the three models

We used models based on signal detection theory to make quantitative predictions for the *sequential effect*, the magnitude of the advantage for the sequential over the simultaneous condition and the *repeated effect*, the magnitude of the advantage for the repeated condition over the simultaneous condition. In this part of the appendix, we consider three specific models: the standard serial, standard parallel, and parallel sampling models.

Notation and common assumptions

Following signal detection theory, these models assume that each stimulus gives rise to a percept that is represented by a single random variable. These variables represent the perceived likelihood that each stimulus is the target. These random variables are assumed to be statistically independent and normally distributed. Throughout, we assume n stimuli with one target, designated T, and $n-1$ distracters, designated $D_1 \ldots, D_{n-1}$. The corresponding density distributions are denoted $f_T(x), f_{D1}(x), \ldots, f_{Dn-1}(x)$. The corresponding cumulative distributions are denoted $F_T(x), F_{D1}(x), \ldots, F_{Dn-1}(x)$. The expected value of the evidence variable is $\mu = 0$ for distracters and $\mu = w$ for targets, where w is a parameter that represents discrimination difficulty. For both targets and distracters, the random variables have unit variance. Except where otherwise noted, models use a maximum decision rule, selecting the stimulus with the highest likelihood of being the target. Beyond these assumptions, the differences between the three models are straightforward.

Standard parallel model

The standard parallel model assumes no effect of divided attention on the formation of random variable

representations. The observer selects the stimulus yielding the maximum evidence value. The probability of choosing the right location is

$$p(\text{correct}) = p[T > \max(D_1, \ldots, D_{n-1})]. \quad (1)$$

To make predictions using the distributional assumptions stated above, we use the function

$$p(\text{correct}) = \int_{-\infty}^{\infty} f(x - w) F_1(x)^{n-1} \, dx. \quad (2)$$

Parallel sampling model

The parallel sampling model assumes that the variance of each random variable is proportional to the number of simultaneously presented relevant stimuli, so that the precision of each representation decreases as attention is divided over more stimuli. Thus the variance of each random variable is doubled in the simultaneous and repeated conditions relative to the sequential condition.

Standard serial model

The standard serial model has an additional parameter m, designating the number of objects the observer can scan during a brief display. The observer forms a random variable representation for each of the m scanned stimuli but has no knowledge about the $n - m$ unscanned stimuli. When all stimuli are scanned, the maximum evidence rule is applied as above. However, if not all stimuli are scanned, we use a guessing rule in which the maximum evidence value obtained from the scanned stimuli must exceed an optimal criterion c (determined by simulations to be the midpoint of zero and w) or else one of the unscanned stimuli is selected at random. Given these definitions, the probability of selecting the correct location is

$$\begin{aligned} p(\text{correct}) = {}& (m/n)p[T > \max(c, D_1, \ldots D_{m-1})] \\ & + (1 - m/n)(1/n - m)p \\ & \times [c > \max(D_1, \ldots D_m)]. \end{aligned} \quad (3)$$

Given the distributional assumptions above, proportion correct can be predicted using the function

$$\begin{aligned} p(\text{correct}) = {}& \frac{m}{n} \int_c^{\infty} f(x - w) F(x)^{m-1} dx \\ & + \left(1 - \frac{m}{n}\right)\left(\frac{1}{n - m}\right) \int_{-\infty}^{c} f(x)^m dx. \end{aligned} \quad (4)$$

To generate the serial model predictions, we assume that m changes between conditions, reflecting the number of stimuli that the observer can scan during each brief display. The value of m was doubled for the

sequential and repeated conditions with the constraint that m cannot exceed n.

Sequential effect predictions

The sequential effect refers to the predicted advantage of the sequential condition over the simultaneous condition. Sequential effect predictions are plotted in Panel A of Figure 11. At each difficulty level, predicted performance in the sequential condition is plotted against predicted performance in the simultaneous condition. Each curve corresponds to a specific model across all values of w from zero to six with the most difficult discriminations at the lower-left point of each curve and the easiest discriminations at the upper right most point of each curve. The vertical distance from the diagonal is the predicted magnitude of the sequential effect.

The curves of each standard serial model show large sequential advantages are predicted when m is one or two with a more modest advantage predicted when $m = 3$. When m is less than three, performance in the simultaneous condition reaches an asymptote well below 100%, reflected by the fact that the model curves do not reach to the upper-right corner of the plot.

Under the standard parallel model, the simultaneous-sequential manipulation has no effect on the formation of representations. Consequently, the predictions for simultaneous and sequential are equivalent at all difficulty levels. This is reflected by the fact that the curve representing the standard parallel model follows the diagonal.

The curve reflecting the parallel sampling model shows that the magnitude of the sequential effect depends on difficulty: It is small at high and low values of w and is largest at intermediate values of w.

Repeated effect predictions

The repeated effect refers to the predicted advantage of the repeated condition over the simultaneous condition. Repeated effect predictions are plotted in Panel B of Figure 11. At each difficulty level, predicted performance in the repeated condition is plotted against predicted performance in the simultaneous condition. A dotted reference line follows the diagonal of the plot and it does not reflect the predictions of any model.

Under the serial models, twice as many stimuli are scanned in the repeated as in the simultaneous, with a maximum of four. Thus the sequential and repeated effects are equivalent under these serial models. Each of the serial models shown predicts equality of the sequential and repeated effects.

Here, the parallel sampling model is not distinguished from the standard parallel model because the simultaneous-repeated manipulation does not change the number of simultaneously presented relevant stimuli. However, to predict the magnitude of the repeated effect under parallel models, we need to consider how observers combine information across the two brief displays. Two parallel models are plotted, representing two ways of integrating information across displays in the repeated condition. In the integration model, samples from corresponding stimuli are combined across the two displays, halving the variance of each random variable representation. In the independent decisions model, each brief display of each stimulus is considered separately, and the maximum decision rule is applied to the set of eight random variables with two ways to pick the correct target (first and second interval). As the plot shows, the integration model predicts a larger repeated effect than the independent decisions model. The parallel integration model predicts equality of the sequential and repeated effects, but the parallel independent decisions model predicts a larger sequential effect than repeated effect, yielding the predicted pattern simultaneous < repeated < sequential.

Generalized equality predictions

Next we consider some ways these models can be generalized without changing the equality predictions that are at the core of this article. Each equality prediction rests on a single critical criterion. As long as this criterion is satisfied, various features can be added to the models without breaking the equality prediction.

The criterion for sequential-repeated equality is that the sequential effect must be equal in magnitude to the repeated effect. That is, the relative benefit of seeing all stimuli twice must precisely match the benefit of seeing two stimuli at a time instead of four stimuli. Of the models described above, the standard serial models ($m = 1$, $m = 2$, and $m = 3$) and parallel sampling model with integration predict equality between repeated and sequential. The standard parallel model (either with integration or independent decisions) predicts better performance in repeated than sequential, and the parallel sampling model with independent decisions predicts better performance in sequential than repeated.

The criterion for simultaneous-sequential equality is that the extent of stimulus processing must be similar in the sequential or simultaneous conditions. The standard parallel model makes this prediction, but all other models discussed are limited-capacity and predict better performance in the sequential conditions.

To generalize these models, consider some features that could be added that preserve the equality predictions. Equalities are maintained as long as new features impact all conditions equally. For one, we can

Models that predict simultaneous = sequential < repeated
  • Standard parallel model
Models that predict simultaneous < sequential = repeated
  • Standard serial models ($m = 1$, $m = 2$, $m = 3$, or a mixture of these)
  • Parallel sampling model with integration
Model that make other predictions
  • Parallel sampling with independent decisions predicts simultaneous < repeated < sequential
Ways to generalize models while preserving both predicted equalities
  • Lapse rate added to all conditions
  • Temporal correlation in sampling
  • Mixtures of models or parameters that predict the same equalities
Ways generalization can fail
  • Mixtures of models or parameters that predict different equalities

Table 2. Specific model predictions.

assume that observers lapse on proportion of trials, performing at chance levels for these. As long as the lapse rate is consistent across conditions, these lapses would attenuate the predicted magnitude of the effects while preserving the equality predictions.

Another plausible cause of effect size attenuation is temporal correlation in perceptual sampling. The parallel models described above assume independent successive samples, a mechanism that describes how observers perform better with long viewing times. However, if successive samples are correlated, the benefit of extended viewing time is reduced. For the models described here, this temporally correlated sampling would attenuate the predicted effect sizes while preserving the equalities.

Mixture models can also preserve equality predictions. For example, the standard serial model described above has a single value for parameter m, reflecting the number of stimuli that can be scanned during a trial. Mixture models allow the value of m to vary between trials (e.g., $m = 1$ for some trials and $m = 2$ for others), changing the predicted magnitude of effects. As long as the distribution of m is equivalent across conditions, the equality predicted performance for sequential and repeated is preserved.

Similarly, the models above assume a single value for w, the difficulty of discriminating targets and distracters. If w is allowed to vary from trial-to-trial (reflecting the fact that some target-distracter pairs are more discriminable than others) then predicted effect sizes change but equalities are preserved.

Thus far, our discussion has focused on the sequential-repeated equality prediction. In fact, all of the generalizations also apply to the simultaneous-sequential equality prediction. Indeed, it is probably even more robust than the sequential-repeated prediction.

For the sequential-repeated equality, the key is whether the effect of increasing the number of stimuli is balanced out by presenting them a second time. They will cancel if both are limited by the same constraints of fixed capacity. For example, in the parallel sampling model there would be fewer samples with more stimuli but more samples with a second look. Thus, details of both the effect of additional stimuli and the effect of a second presentation matter for this equality prediction.

Some mixture models do not preserve the equalities. First, mixing models with qualitatively different predictions produce intermediate results. For example, if a standard parallel model operates on some trials and a standard serial model operates on others, an intermediate result of simultaneous < sequential < repeated is predicted. Second, in the case of mixed parameter models, mixtures of extreme parameter values may produce ambiguous results. For example, any model with low values of w predict near-chance performance in all conditions, while parallel models with large values of w predict perfect performance in all conditions. A mixture of trials with low w (near-chance performance) in some trials and high w (near-perfect performance) in others could produce an ambiguous result in which simultaneous = sequential = repeated.

Another example class of models that might not preserve the equalities is the self-terminating models described by Townsend (1974). He considered the possibility that observers might terminate stimulus processing rather than continue until the trial's end. These models are more relevant to conditions in which the observers face instructions to respond quickly, but their implications are worth considering here.

Table 2 gives a summary of the model predictions and generalizations reviewed in this appendix.