

Is there a serial bottleneck in visual object recognition?

Dina V. Popovkina

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 Psychological and Brain Sciences,
University of Iowa, Iowa City, IA, USA



Geoffrey M. Boynton

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



Divided attention has little effect for simple tasks, such as luminance detection, but it has large effects for complex tasks, such as semantic categorization of masked words. Here, we asked whether the semantic categorization of visual objects shows divided attention effects as large as those observed for words, or as small as those observed for simple feature judgments. Using a dual-task paradigm with nameable object stimuli, performance was compared with the predictions of serial and parallel models. At the extreme, parallel processes with unlimited capacity predict no effect of divided attention; alternatively, an all-or-none serial process makes two predictions: a large divided attention effect (lower accuracy for dual-task trials, compared to single-task trials) and a negative response correlation in dual-task trials (a given response is more likely to be incorrect when the response about the other stimulus is correct). These predictions were tested in two experiments examining object judgments. In both experiments, there was a large divided attention effect and a small negative correlation in responses. The magnitude of these effects was larger than for simple features, but smaller than for words. These effects were consistent with serial models, and rule out some but not all parallel models. More broadly, the results help establish one of the first examples of likely serial processing in perception.

Introduction

Visual tasks can produce a variety of divided attention effects for making multiple judgments, ranging from little or no effect, to large effects. One way to measure divided attention effects is to use dual tasks,

in which participants perform the same task in two locations. Some simple judgments, such as detecting luminance increments, can be performed in two locations as well as in one, as if there are independent parallel processes for each location (Bonnell, Stein, & Bertucci, 1992). Additional evidence of independent parallel processing is found in experiments on summary statistics (Attarha, Moore, & Vecera, 2014; Sun, Chubb, Wright, & Sperling, 2016). Other judgments, such as semantic categorization of masked words, have large divided attention effects, as if they can be carried out in only one location at a time (White, Palmer, & Boynton, 2018). Our question is whether processing of multiple visual objects is subject to large divided attention effects, like words, or little or no divided attention effects, like simple features. Specifically, we consider semantic categorization of nameable visual objects.

The anatomic organization of the early stages of the visual system is capable of parallel processing. For example, early cortical areas have retinotopic organization, and stimuli presented in the left and right visual hemifields are processed by primary visual cortices in opposite hemispheres. This separation can support parallel processing of visual input. Indeed, there is behavioral and physiological evidence consistent with parallel processing of multiple stimuli for simple judgments, such as contrast detection (Scharff, Palmer, & Moore, 2011; Chen & Seidemann, 2012; White, Runeson, Palmer, Ernst, & Boynton, 2017). For judgments of simple features, there may be little or no effect of divided attention if the information about each stimulus can be represented by a distinct neuronal subpopulation in early retinotopic areas.

In contrast, semantic categorization of simultaneously presented masked words results in a severe

Citation: Popovkina, D. V., Palmer, J., Moore, C. M., & Boynton, G. M. (2021). Is there a serial bottleneck in visual object recognition? *Journal of Vision*, 21(3):15, 1–21, <https://doi.org/10.1167/jov.21.3.15>.



processing limit, as if participants can recognize only one word at a time (White, Palmer, & Boynton, 2018; White, Palmer, & Boynton, 2020). This large effect of divided attention is consistent with a serial processing bottleneck. Indeed, processing of words appears to be mediated by the visual word form area, which is less retinotopically organized (Rauschecker, Bowen, Parvizi, & Wandell, 2012; Le, Witthoft, Ben-Shachar, & Wandell, 2017). Because this higher-level visual area might be unable to represent two words at once, this could explain the serial processing for reading words (White, Palmer, Boynton, & Yeatman, 2019). Might judgments of multiple visual objects be similarly constrained?

Previous studies of object perception

In an early study, Biederman and colleagues (Biederman, Blickle, Teitelbaum, & Klatsky, 1988) examined whether object perception is limited by divided attention. The stimuli were line drawings of objects associated with basic-level category names (e.g. “traffic light” or “file cabinet”). The task was to search for a specific object among a display of one to six objects. Response time to find the target increased with the addition of more distractor objects to the display, suggesting that perceptual processing of objects is limited in some way. However, such response time set-size effects do not clearly distinguish serial and parallel processing, because for this measurement parallel processes can mimic serial processing (Townsend, 1971; Townsend, 1990; Palmer, Verghese, & Pavel, 2000).

Potter and Fox (2009) used the simultaneous-sequential paradigm (Shiffrin & Gardner, 1972) to measure how object categorization is limited by divided attention. Stimuli were pictures of objects in scenes with associated verbal descriptions, presented as rapid serial visual presentation (RSVP) sequences of eight displays with one to four stimuli simultaneously displayed. The task was to search for the presence of a picture matching a target verbal description (e.g. “balloons” or “cut up fruit”). The target picture could either appear together with one or more distractors in the same presentation interval (simultaneous), or in the other presentation interval (sequential). The key comparison was between simultaneous and sequential presentations. Performance was worse for targets presented simultaneously with one or more distractors, compared with performance for targets and distractors presented sequentially. This sequential advantage is consistent with limited-capacity processing in perception. Although Potter and Fox showed that object processing is limited in capacity, they did not distinguish whether it was serial or parallel, because

both serial and limited-capacity parallel processes produce similar results in these experiments.

Scharff and colleagues (Scharff, Palmer, & Moore, 2011) also examined whether object processing has limited capacity using the simultaneous-sequential paradigm. The stimuli were pictures of animals that were members of categories (e.g. fox or deer). The task was to determine which of two categories of animal was present in a multi-object display. For example, the target might be a fox or a deer presented among distractors that were a squirrel and a moose. Scharff and colleagues used simpler displays than the relatively long RSVP sequences used by Potter and Fox (2009). The simultaneous condition presented four objects on the display at a time, whereas the sequential condition presented two objects on the display at a time over two intervals. The results showed a sequential advantage in this task, consistent with limited-capacity perceptual processing. However, this study was also unable to distinguish between a serial process and a limited-capacity parallel process, because both could account for the sequential advantage.

Alternative hypotheses

Our focus is the hypothesis that the perceptual processing of objects, like words, is severely limited under divided attention. For example, if there are large divided attention effects for the semantic categorization of nameable objects that are similar to those observed for words, there might be a serial bottleneck in object processing. In that case, one possibility is that words and objects share serial processes beyond retinotopic cortex. Such a bottleneck might restrict the extraction of the meaning of the word or object (Broadbent, 1958). Another possibility is that a serial bottleneck might constrain a certain type of higher-level process that is similar for objects and words, but need not be the same. For example, the bottleneck may constrain the formation of an object representation (Kahneman, Treisman, & Gibbs, 1992).

An alternative possibility is the hypothesis that object judgments depend on only parallel processes, either similar to simple feature judgments or intermediate between results for features and words. This predicts that object judgments are not subject to large divided attention effects. This idea is consistent with the argument that some visual processing, such as the discounting of distractors, can occur in parallel for multiple stimuli, and limitations in processing constrain only judgments of multiple simultaneous targets (Duncan, 1980). According to this hypothesis, the decreased performance during divided attention tasks is due to limitations beyond perception, such as in sensorimotor processes, which link percepts to actions

(Allport, 1987). This view stands in contrast to a serial bottleneck in object perception.

As described above, although there is consistent evidence that object tasks do show effects of divided attention, unlike some simple feature tasks that show no such effects, this evidence is ambiguous with respect to whether object tasks show divided attention effects as large as those caused by a serial bottleneck. Specifically, the studies above present evidence only for limited capacity, which is consistent with either serial or parallel processes. Our experimental approach follows the studies by White and colleagues that have uncovered evidence for a serial bottleneck for word categorization. These studies take advantage of a dual-task paradigm that can distinguish between specific serial and parallel models.

Benchmark models of perceptual dual tasks

Here, we consider three models of serial and parallel processes as benchmarks for judging experimental evidence in favor of or against the existence of a serial processing bottleneck. These models were implemented as in White et al. (2020).

The independent parallel model allows for two stimuli to be processed independently (with unlimited capacity). Because processing is not affected by the number of objects to be processed, this model predicts no divided attention effect - judging two stimuli will be as accurate as judging one. This prediction has been satisfied for detecting simple features such as luminance increments (Graham, Kramer, & Haber, 1985; Bonnel, Stein, & Bertucci, 1992).

The fixed-capacity parallel model is a special case of a limited-capacity parallel model. It assumes that parallel processing is limited such that the total amount of information obtained from a display is constant (Taylor, Lindsay, & Forbes, 1967). Hence the name: “fixed capacity.” One way to implement such a constraint is to use the metaphor of statistical sampling as done in Shaw’s (1980) sample size model. This theory starts with a signal detection theory framework in which the quality of a percept, and therefore the probability of its detection, is a function of a single random variable. In particular, the quality of a percept is assumed to correspond to the variability of estimates based upon a set of samples of the underlying random variable. When one object is relevant, all of the available samples can be directed to this one object. When there are two objects, the samples must be shared among the objects. Consequently, each object is sampled less often, and the quality of the percept per object is lower. For the case of equally sampling two objects instead of one object, the standard deviation of the mean of the samples increases by

the square root of two (in d' units). This prediction has been satisfied for discriminating some simple features (Miller & Bonnel, 1994) and for some simple visual memory tasks (Smith, Lilburn, Corbett, Sewell, & Kyllingsbaek, 2016).

The all-or-none serial model represents a processing bottleneck, which allows for only one stimulus to be processed at a time. For this “all-or-none” model, we also assume that there is no time to process a second stimulus: no switching of a single serial process between two stimuli. The model thus predicts the largest effect of divided attention. In addition, because only one stimulus out of two is processed, there is a negative correlation in the accuracy of the two responses: correct responses for one stimulus co-occur with incorrect (or chance) responses for the other. These predictions have been satisfied for letter-digit tasks with conflicting S-R mapping (Sperling & Melchner, 1978), for certain multiple object dual tasks (Bonnel & Prinzmetal, 1998), and for masked words (White et al., 2018; White et al., 2020). In summary, the all-or-none serial model predicts both a large magnitude effect of divided attention and a negative correlation between dual-task responses.

Effects of intermediate magnitudes can be predicted with generalizations of each model. For example, a fixed-capacity parallel process can predict a larger dual-task deficit when target detection uses discretized states than when target detection uses continuous information (Swagman, Province, & Rouder, 2015; and see Appendix). Similarly, a serial process can produce a smaller dual-task deficit if there is enough time within one trial to complete processing one stimulus and switch to processing a second stimulus (White et al., 2020). To interpret our findings, we consider these generalized versions of these models alongside the three benchmark models.

Overview of experiments

In this article, we ask whether semantic categorization of visual objects shows large divided attention effects, consistent with that predicted by a serial bottleneck. The observed accuracy was compared with predictions of the three benchmark models to interpret the effects of divided attention. Critically, we use brief stimulus presentations and masking to minimize the opportunity for the switching of any serial processes. Without time constraints, a serial process could completely process a stimulus in one location, and start processing a stimulus in another location, reducing the observed effect of divided attention. This brief timing is implemented in two ways: in Experiment 1, multiple stimuli were shown using RSVP; in Experiment 2, single stimuli were shown with pre- and post-masks.

Experiment 1: RSVP

In the first experiment, stimuli were presented using brief durations and RSVP to limit the time available to process stimuli, and thus help distinguish serial and parallel processing predictions (Forster, 1970; Potter & Haggmann, 2015; Robinson, Grootswagers, & Carlson, 2019). The task was semantic object categorization, similar to the task used with words by White and colleagues (White et al., 2018).

Methods

Participants

For Experiment 1, 12 paid participants (6 men and 6 women) were recruited from the University of Washington and greater Seattle community; author D.V.P. was one of the participants for Experiment 1. Participants had normal or corrected-to-normal visual acuity. All participants gave written and informed consent in accordance with the Declaration of Helsinki and the human subjects Institutional Review Board at the University of Washington.

Apparatus and eyetracking

Stimuli were presented on a linearized CRT monitor (Sony GDM-FW900) with a resolution of 1024×640 pixels and a 120 Hz refresh rate. The monitor was viewed from a 60 cm distance and had a peak luminance of 90 cd/m^2 . Presentation of stimuli was controlled using MATLAB (MathWorks, Natick, MA) and the Psychophysics Toolbox (Brainard, 1997). An Eyelink 1000 (SR Research, Ontario, Canada) and the Eyelink Toolbox (Cornelissen, Peters, & Palmer, 2002) were used to monitor and enforce fixation during the experiment. A trial was terminated if the participant blinked or moved their eyes outside of a 2 degree window while stimuli were present on the screen. On average over all experiments, $0.7\% \pm 0.1\%$ of trials were terminated due to blinks or apparent eye movements.

Stimuli

The stimuli were photographs of nameable objects removed from the background context. Stimuli were hand-selected from an internet image search and from the Massive Memory Object Categories image set (Konkle, Brady, Alvarez, & Oliva, 2010). Each image was adjusted to maximize contrast and remove color, and was resized to a 100 pixel \times 100 pixel square (4.2 degrees \times 4.2 degrees).

Stimuli were from eight categories: plants, food, clothing, animals, furniture, household devices,



Figure 1. Example stimuli used in both experiments. All images from the category “animals” are shown.

transport, and musical instruments. Two judges confirmed that all examples were easy to identify and clearly belonged to the assigned category, and not the other categories. Each category had 50 exemplar objects; Figure 1 shows the 50 objects in the category “animals.” With eight categories, the stimulus set had a total of 400 objects.

Procedure

Figure 2 shows a schematic of the RSVP task, which was similar to Experiment 1 in White et al. (2018). On each trial, the participants saw a category word, followed by briefly presented visual objects, and a

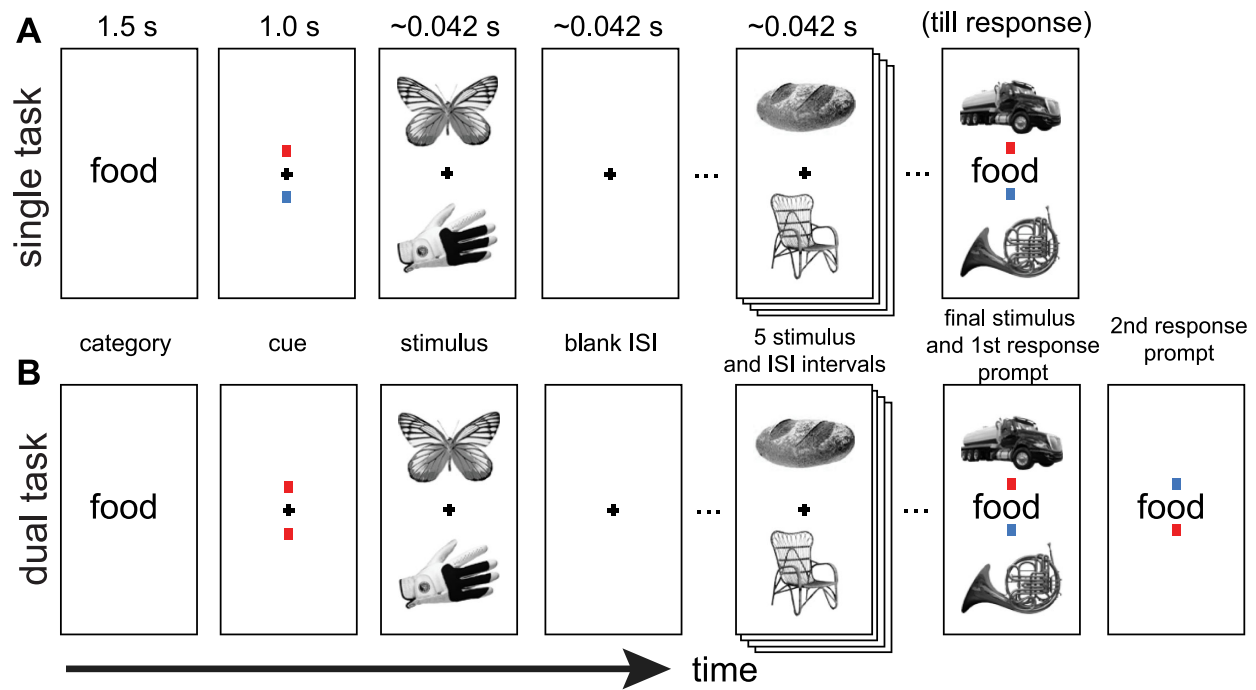


Figure 2. Rapid serial visual presentation (RSVP) procedure in [Experiment 1](#). Trial sequences for the single task (**A**, top location cued) and dual task (**B**, both locations cued) are shown. Ellipses indicate more intervals of the same duration, for a total of seven intervals containing objects and six intervening blank intervals. In this example, the observer cue color is red. Mean stimulus and blank ISI durations shown; these were adjusted separately for each observer to produce approximately 80% accuracy in the single task.

response prompt. Participants reported with a button press whether an object from the target category had appeared in the cued location. For example, for the trial in [Figure 2](#), participants were looking for food objects and a target object (bread) was presented in the top location. The relevant location(s) were cued both before presentation and during the response prompt. Red and blue colored lines were used as cues, with one color assigned as a relevant cue for each participant and the other color serving as the irrelevant cue. The assignment was balanced such that for half of the participants, the relevant cue was red, and for the other half, the relevant cue was blue (in [Figure 2](#), the relevant cue is red).

Stimuli were presented above and below a 0.5 degree fixation cross (top-side and bottom-side, respectively) and were centered at 4 degrees away from fixation. In [Experiment 1](#), stimuli were presented as a RSVP sequence. [Figure 2](#) shows a schematic of an example trial sequence, with each box representing a time interval. The RSVP sequence contained seven object presentations separated by equal duration intervals with a blank screen (only 3 object presentations are shown in the figure). The first and last object presentations never contained an object from the target category (serving as pre- and post-masks). In the second to sixth intervals, one object from the target category could appear amid a stream of objects from other categories. Over the course of the entire sequence, there was a 50% chance

of a target object appearing within the stream at a given stimulus location. This probability was independent for the two locations: that is, the presence of a target object in one location gives no information about the presence of a target object in the other location. The only dependency between locations was that in trials with a target present in both locations, the targets appeared in the same interval to make switching ineffective. All other stimuli, including masks, were randomly chosen from nontarget categories. The post-mask stayed on the screen for 700 ms, at which time a brief tone accompanied the response prompt. The post-mask was replaced by a blank as soon as there was a response.

Conditions

Stimuli were presented in three different conditions, which were blocked:

In the single-task condition, there was a single task to perform on each trial. Objects were presented in two locations, but only one location was relevant. A label at the beginning of a block indicated whether the relevant location was on the top or bottom side of the display; the relevant location stayed the same for the duration of the block. Participants judged the object in the cued location only.

In the dual-task condition, there were two tasks to perform on each trial. Again, objects were presented

in two locations, but both locations were relevant and participants judged the objects separately for each location. If a target object was present in both locations, the target objects were shown at the same time in the sequence to make switching strategies ineffective. The order of testing the two locations was randomized.

In the control single-stimulus condition, there was a single task to perform on each trial. Participants saw an object in only one location and judged the object in that location. The relevant location stayed the same for the duration of the block. The only difference from the single-task condition was the absence of the irrelevant stimulus. This condition was included to check for crowding and similar interference effects.

Timing

Before the main experiment, the RSVP timing was adjusted for each participant to achieve approximately 80% accuracy in the single-task condition by manipulating the duration of the stimulus and blank intervals. The stimulus and interval durations were always identical and adjusted together. The mean stimulus and interval duration across 12 participants was 42 ms (individual timings ranged from 33 ms to 58 ms), and the resulting mean accuracy was $82\% \pm 1\%$ in the single-task condition. For the main experiment, the same customized timing was used in all conditions.

It is possible for factors other than timing to affect accuracy: for example, stimuli could be inherently difficult to discriminate, or the RSVP paradigm could be limited by a memory requirement. In a control experiment, we verified that timing was a primary factor limiting accuracy by increasing all interval durations to 150 ms (for 2 early participants, the duration was 100 ms and 125 ms, respectively). Over 128 trials in the single-task condition, average accuracy was $93.8\% \pm 1.4\%$ ($n = 12$), only about 6% below perfect. Thus, other phenomena, such as discrimination difficulty and memory processes, limited accuracy only slightly.

Responses

Participants made unspeeded responses using one of four buttons. They reported “yes” / “no” answers to the core question: “on the prompted side, did any object belong to the target category?” Participants also gave a confidence rating (“likely” or “guess”) associated with their report. Specifically, the four buttons represented the following responses: “likely no,” “guess no,” “guess yes,” and “likely yes.” Button layout was horizontal, orthogonal to the vertical stimulus layout. Responses about the top location were arranged along the top row of a keypad; responses about the bottom location were arranged along the bottom row of a second keypad. After the response, feedback was given in the form of a high- or low-frequency tone for correct and incorrect

responses, respectively. Feedback for the responses in the dual-task condition was provided only after both responses were given.

Design

The experiment was carried out in sessions of six blocks of 16 trials: two blocks of dual task; two block of single task, one cued to the top location and one cued to the bottom location; two blocks of single stimulus, one cued to the top location and one cued to the bottom location. Trials within each block had the same target category and the order of blocks was randomized for each session. Each session took about 15 minutes to complete. A complete experiment included at least 38 sessions for a total of at least 1188 trials per task condition.

Analysis

Accuracy was measured as the percentage of area under the receiver operating characteristic (ROC). This metric has properties similar to two traditional accuracy measures: like percent correct, it is bounded by 50% (chance accuracy) and 100% (perfect accuracy); and like d' , it is an unbiased measure of accuracy. The ROC curves were constructed using the confidence ratings reported by the participants. All accuracy results are reported as mean \pm standard error of the mean. For significance testing, all alpha levels were set to 0.05 and all t -tests were two-tailed.

Number of participants

To determine the appropriate sample size (number of participants), we examined data from four previous dual-task experiments using RSVP and masked word stimuli (White et al., 2018; White et al., 2020). In each, participants ($n = 10$) performed judgments of words with similar methods as the current study. A power analysis was conducted to determine the sample size needed to distinguish the predictions of the fixed-capacity, parallel model and the all-or-none serial model. This was done for the dual-task deficit and a conditional accuracy measure of response correlation. Our calculations assumed alpha and beta errors of 0.05 (power of 95%). The estimated minimum sample size was five for the dual-task deficit and eight for the conditional accuracy measure. To be conservative, we used a minimum sample size of 10. In practice, we collected data from a few additional participants, for a total of 12 participants in [Experiment 1](#) and 11 participants in [Experiment 2](#).

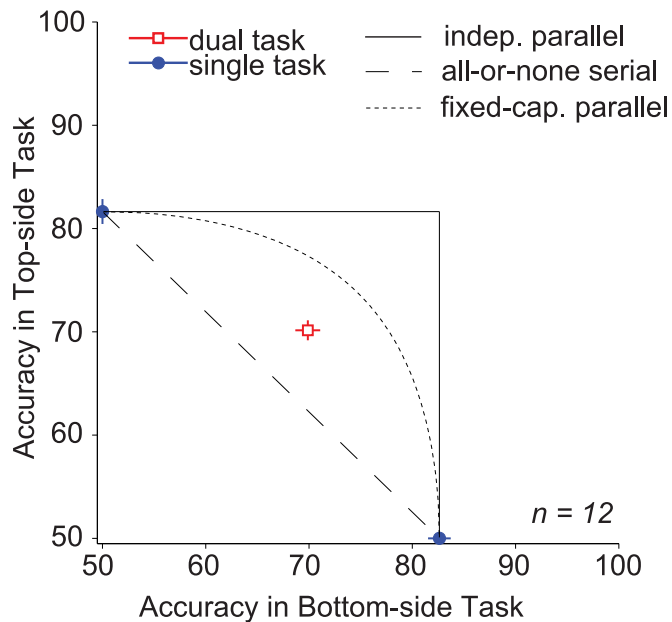


Figure 3. Attention operating characteristic for Experiment 1. Observed behavioral accuracy, measured as percent area under the ROC curve, in single (blue) and dual (red) tasks. Error bars: standard error of the mean. Solid line: prediction of the independent parallel model. Dashed line: prediction of the all-or-none serial model. Dotted curve: prediction of the fixed-capacity parallel model.

Main results

Dual-task deficit

Accuracy in the semantic categorization task was worse for categorizing two objects (dual-task accuracy: $70.0\% \pm 0.8\%$) compared with categorizing one object (single-task accuracy: $82.1\% \pm 1.1\%$). The difference is the dual-task deficit: $12.1\% \pm 1.0\%$ (significantly different from zero, $t(11) = 12.3$, $p < 0.001$).

Figure 3 shows average accuracy in the form of an attention operating characteristic (Sperling & Melchner, 1978). Accuracy (measured as area under ROC, see Methods section) for the task in the top location (y -axis) is plotted against accuracy for the task in the bottom location (x -axis). The blue circles on the axes indicate the single-task accuracy for the respective locations; the red square indicates accuracy for each of the locations in the dual-task condition. The overlaid lines correspond to the predictions of the three benchmark models: the independent parallel model (solid line); the all-or-none serial model (dashed line); and the fixed-capacity parallel model (dotted line). The observed results are inconsistent with the independent parallel model and the fixed-capacity model because the dual-task deficit is larger than the deficit predicted by either of these models. The results are also inconsistent with an all-or-none serial model because the dual-task deficit is smaller than the deficit it predicts. In summary,

there was a large dual-task deficit in the semantic categorization task; but the magnitude of the observed deficit was smaller than predicted by an all-or-none serial model.

Response correlation

In the presence of a serial bottleneck, the observer in a dual-task trial can perform the task for the stimulus in one location, and not in the other. To extract this hallmark of a bottleneck, we use a trial-by-trial analysis of response correlation. Specifically, an all-or-none serial process predicts a negative correlation between the accuracy of responses. Accuracy should be higher for a response in one location if the response in the other location was wrong, rather than correct. One way to quantify such a response correlation is to use a conditional accuracy measure (see White et al., 2018; White et al., 2020). Conditional accuracy can be calculated only on dual-task trials. Responses are separated into two sets of trials: one set where the response about the other stimulus in the same trial was correct, and another set where the response about the other stimulus in the same trial was wrong. Then, accuracy is calculated separately for each set.

Figure 4 shows accuracy conditioned on whether the response about the stimulus in the other location was correct (ordinate) or wrong (abscissa). The dashed line shows the conditional accuracy predicted by the all-or-none serial model: higher accuracy when the response about the other stimulus was wrong than when the response about other stimulus was correct (this prediction was generated using simulated dual-task trials from an all-or-none serial model; for details; see White et al., 2018). Neither of the parallel models predicts any difference in conditional accuracy (solid line). In dual-task trials, the observed conditional accuracy was higher when the response on the other side was wrong (70.7%) than when the response on the other side was correct (68.1%), a difference of $-2.5\% \pm 1.1\%$ (significantly different from zero, $t(11) = -2.24$, $p = 0.046$). This negative correlation had a smaller magnitude than predicted by the all-or-none serial model, but it was reliably different from zero, the prediction of the parallel models. In summary, there was a negative conditional accuracy difference that is often considered to be a signature of serial processing.

Secondary results

Effect of crowding from the second stimulus

In the single-stimulus condition, participants performed the single task with stimuli presented only in the relevant location. Accuracy in the single-stimulus

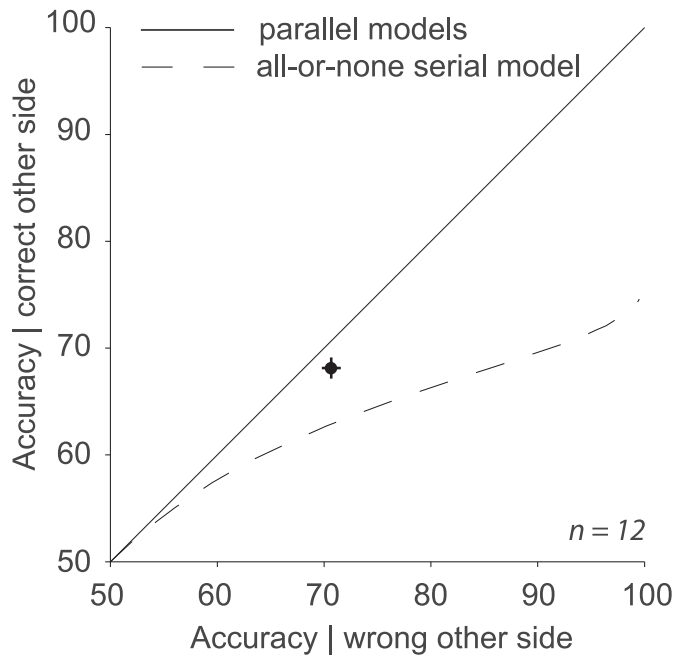


Figure 4. Conditional accuracy in [Experiment 1](#). Observed behavioral accuracy, measured as percent area under the ROC curve, in dual-task trials conditioned on whether the response about the other side was wrong (abscissa) or correct (ordinate). Error bars: standard error of the mean. Solid line: prediction of both parallel models. Dashed line: prediction of the all-or-none serial model.

condition ($83.0\% \pm 0.9\%$) was slightly better, but similar to accuracy in the single-task condition ($82.1\% \pm 1.1\%$). The difference was $0.9\% \pm 0.5\%$ (not significantly different from zero, $t(11) = 1.98$, $p = 0.073$). Thus, there was no evidence of crowding in this experiment.

Effect of response order in the dual task

In the dual task, one of the two locations was randomly chosen as the first response, and the other as the second response. Accuracy for the first response ($70.0\% \pm 0.8\%$) was similar to accuracy for the second response ($70.0\% \pm 1.0\%$). The difference was $0.04\% \pm 0.8\%$ (not significantly different from zero, $t(11) = 0.049$, $p = 0.96$). Thus, neither memory nor response interference appeared to differentially affect the second response.

Effect of stimulus order in the RSVP sequence

Across all trials and participants, the accuracy in each interval that could contain a target was: 73.3%, 74.2%, 73.4%, 70.2%, and 70.9% (listed in chronological order). There was a small advantage for detecting a target object in the first possible stimulus interval ($73.3\% \pm 1.0\%$) compared with the last possible stimulus interval ($70.9\% \pm 1.4\%$). This difference was small ($2.4\% \pm$

1.3%) and not significant ($t(11) = 1.83$, $p = 0.094$). Such small “primacy” effects are often reported for RSVP procedures ([Coltheart, 1999](#)).

Two-target effects

In some cases, target detection can be affected by the presence of another target in the display ([Duncan, 1980](#)). The difference between trials where the other stimulus was a distractor and trials where the other stimulus was a target was $3.4\% \pm 1.7\%$ in the single-task condition (significantly different from zero, $t(11) = 1.95$, $p = 0.08$), and $3.8\% \pm 0.8\%$ in the dual-task condition (significantly different from zero, $t(11) = 4.64$, $p < 0.001$). These differences are consistent with a performance deficit in the presence of another target in the display. However, both serial and parallel models can give rise to such effects; see [Appendix](#) and General Discussion.

Discussion

In summary, in the first experiment, participants cannot categorize two objects as well as they can categorize one. The large dual-task deficit and the negative correlation between responses were generally consistent with, but smaller than, the predictions of an all-or-none serial model. Our findings also reject the fixed-capacity parallel model. There was a two-target effect, but the results were not mediated by other stimulus- and task-related factors, such as crowding or response order effects (see [Appendix](#) for similar analyses of response bias and stimulus location). Before considering the implications of these results more deeply, we present a second version of the experiment to test the generality and reliability of these results.

Experiment 2: Masking

In the first experiment, we used brief stimulus durations and RSVP to differentiate predictions of serial and parallel models in a semantic object categorization task. In this second experiment, we asked whether removing the RSVP component of the task can produce the same results. Specifically, brief masked presentation of a single object was used, similar to [Experiment 2](#) of [White et al. \(2018\)](#) with words. This change helps address potential confounds arising from the temporal uncertainty of target appearance in the RSVP stream, or some effect of interference or overload in short-term memory ([Akyurek & Hommel, 2005](#)), whereas the remaining masks continue to make the task challenging.

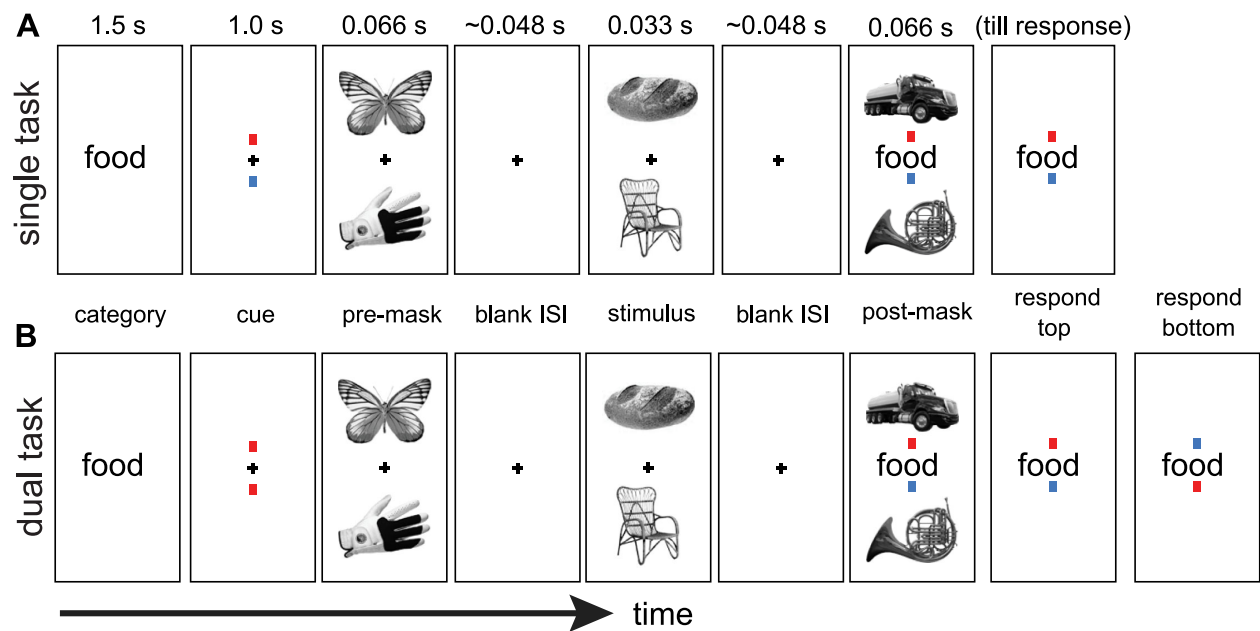


Figure 5. Masking procedure in Experiment 2. Trial sequences for the single task (A, top location cued) and dual task (B, both locations cued) are shown. Unlike in Experiment 1, the target can occur in only one time interval. In this example, the observer cue color is red. Mean ISI durations shown; these were adjusted separately for each observer to produce approximately 80% accuracy in the single task.

Methods

The methods were the same as in Experiment 1, with the exception of differences described below.

Participants

For Experiment 2, 11 paid participants (5 men and 6 women) were recruited from the University of Washington and greater Seattle community. Seven of these participants also completed Experiment 1. In Experiment 2, the participants completed a minimum of 1129 trials per condition.

Procedure

In Experiment 2, a single stimulus display was presented with pre- and post-masks, rather than an RSVP sequence. Figure 5 shows a schematic of an example trial sequence, with each box representing a time interval. In this experiment, the display sequence contained three object presentation intervals separated by intervals with a blank screen. The first and last object presentation intervals never contained an object from the target category (serving as pre- and post-masks). In the second interval, one object from the target category could appear. There was a 50% chance of a target object appearing, independently at each stimulus location: that is, the presence a target object in one location gives no information about the presence of a target object in

the other location. The stimuli shown in mask intervals were randomly chosen from nontarget categories.

Timing

The object presentations were fixed in duration: pre-mask = 66 ms, stimulus interval = 33 ms, and post-mask = 66 ms. Timing of both intervening blank intervals was adjusted for each participant to achieve approximately 80% accuracy in the single-task condition. The mean interval duration across 11 participants was 48 ms (range: 25 ms – 91 ms), and the resulting mean accuracy was $80\% \pm 1\%$ in the single-task condition. In a control experiment, we verified that timing was a primary factor limiting accuracy by setting the blank interval duration to 400 ms in a short session of 128 trials. With this longer blank interval duration, average accuracy in the single task was $95.2\% \pm 1.3\%$ ($n = 8$ participants). Thus, like in Experiment 1, other phenomena, such as discrimination difficulty, did not limit accuracy with longer intervals.

Main results

Dual-task deficit

Accuracy in the semantic categorization task was worse for categorizing two objects (dual-task accuracy: $68.2\% \pm 1.1\%$) compared with categorizing one

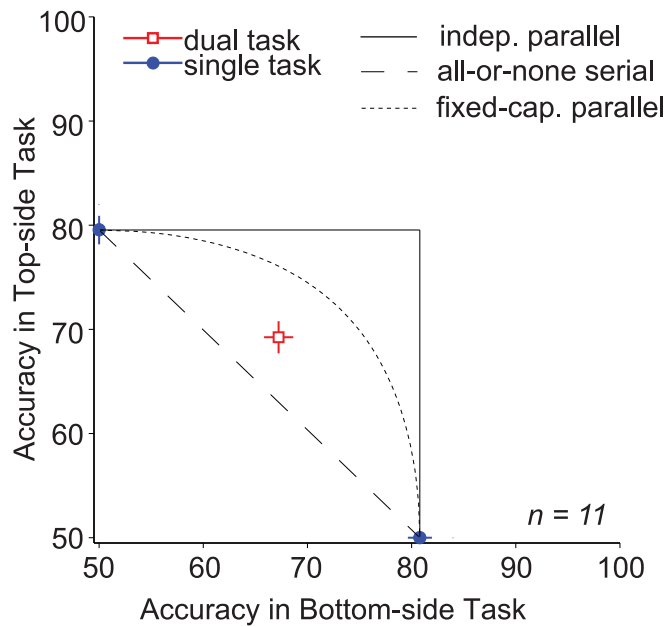


Figure 6. Attention operating characteristic for Experiment 2. Observed behavioral accuracy, measured as percent area under the ROC curve, in single (blue) and dual (red) tasks. Error bars: standard error of the mean. Solid line: prediction of the independent parallel model. Dashed line: prediction of the all-or-none serial model. Dotted curve: prediction of the fixed-capacity parallel model.

object (single-task accuracy: $80.2\% \pm 1.3\%$). The dual-task deficit was $11.9\% \pm 1.2\%$ ($t(10) = 9.85$, $p < 0.001$). Figure 6 shows average accuracy in Experiment 2 in the form of an attention operating characteristic. As in Figure 3, accuracy for the task in the top location is plotted against accuracy for the task in the bottom location. The blue circles on the axes indicate the single-task accuracy for the respective locations; the red square indicates accuracy for each of the locations in the dual-task condition. The overlaid lines correspond to predictions of three theoretical models: the independent parallel model (solid line); the all-or-none serial model (dashed line); and the fixed-capacity parallel model (dotted line). The results are inconsistent with the independent parallel model and the fixed-capacity parallel model because the dual-task deficit is larger than the deficit predicted by either of these models. The results are also inconsistent with an all-or-none serial model because the dual-task deficit is smaller than the deficit it predicts. In summary, there was a large dual-task deficit, but it was smaller than that predicted by the all-or-none serial model.

Response correlation

Figure 7 shows accuracy conditioned on whether the response about the stimulus in the other location was correct (ordinate) or wrong (abscissa). The dashed line

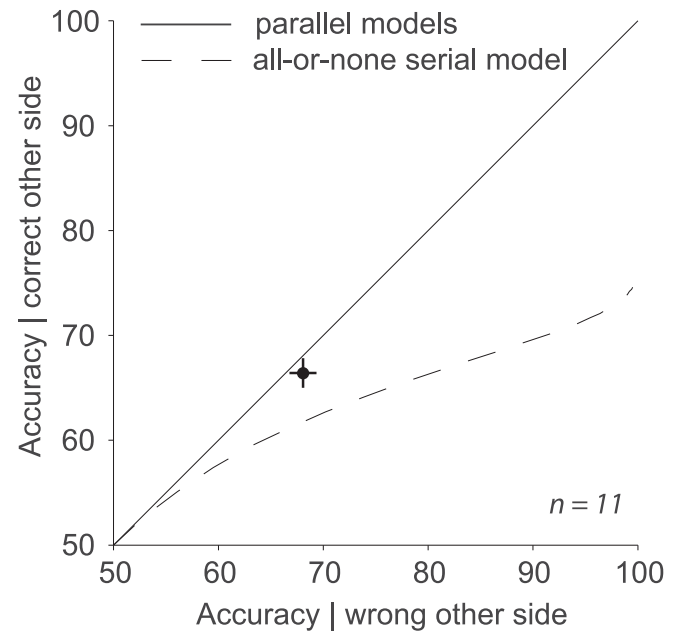


Figure 7. Conditional accuracy in Experiment 2. Observed behavioral accuracy, measured as percent area under the ROC curve, in dual-task trials conditioned on whether the response about the other side was wrong (abscissa) or correct (ordinate). Error bars: standard error of the mean. Solid line: prediction of both parallel models. Dashed line: prediction of the all-or-none serial model.

shows the prediction of the all-or-none serial model: higher accuracy when the response about the stimulus was wrong (dashed line). Neither of the parallel models predicts any difference in conditional accuracy (solid line). Accuracy in the dual-task condition was higher when the response on the other side was wrong (68.1%) than when the response on the other side was correct (66.4%). This difference of $-1.6\% \pm 1.3\%$ was not reliable ($t(10) = -1.26$, $p = 0.25$). In summary, the difference in conditional accuracy was consistent in sign with the prediction of the all-or-none serial model, but not reliably different than zero.

Secondary results

Effect of crowding from the second stimulus

Accuracy in the single-stimulus condition ($83.0\% \pm 1.4\%$) was similar to accuracy in the single-task condition ($80.2\% \pm 1.3\%$). The difference was $2.8\% \pm 0.6\%$ (significantly different from zero, $t(10) = 4.47$, $p = 0.0012$). Thus, displaying two stimuli had a small effect in this experiment.

Effect of response order in the dual task

In the dual task, accuracy for the first response ($68.9\% \pm 1.2\%$) was similar to accuracy for the

second response ($67.7\% \pm 1.1\%$). The difference was $1.3\% \pm 0.5\%$ (significantly different from zero, $t(10) = 2.58$, $p = 0.027$). Thus, memory or interference had a small effect on the second response; however, it was much smaller than the dual-task deficit.

Two-target effects

The difference between trials where the other stimulus was a distractor and trials where the other stimulus was a target was $3.4\% \pm 2.3\%$ in the single-task condition (not significantly different from zero, $t(10) = 1.46$, $p = 0.17$) and $3.6\% \pm 1.3\%$ in the dual-task condition (significantly different from zero, $t(10) = 2.67$, $p = 0.02$). These differences are consistent with a performance deficit in the presence of another target in the display. However, both serial and parallel models can give rise to such effects; see [Appendix](#) and General Discussion.

Discussion

In [Experiment 2](#), we found that participants categorized two visual objects worse than they categorized one. The large dual-task deficit was consistent with, but less than predicted by, an all-or-none serial model. Although the response correlation was negative (as predicted by serial models) it was not significantly different from the parallel processing prediction of zero correlation. There was a two-target effect, but the divided attention effects were not mediated by any other stimulus- and task-related factors observed in this study, such as crowding or response order (see [Appendix](#) for similar analyses of response bias and stimulus location).

The magnitude of the dual-task deficit was similar in [Experiments 1](#) and [2](#): both were larger than the fixed-capacity parallel model prediction, and smaller than predicted by an all-or-none serial model. This result is distinct from the observations for both simple feature judgments and word judgments (see Summary of Results below). The sign of the response correlation was the same in [Experiments 1](#), and [2](#), consistent with a serial bottleneck; but for [Experiment 2](#), it was not statistically different from the prediction of parallel models. In sum, three of four lines of evidence reject the fixed-capacity parallel model for visual object processing.

General discussion

Summary of results

In this study, we asked whether the semantic categorization of nameable objects shows large

effects of divided attention, like those observed for categorization of words. We performed the experiment using two presentation paradigms, RSVP and masking, and found similar divided attention effects in both. Specifically, there was a large dual-task deficit, and a negative correlation in responses. Both findings reject the fixed-capacity parallel model, and are smaller than the prediction of an all-or-none serial model.

[Figure 8](#) summarizes these two metrics for seven studies of objects, words, and simple features ([White et al., 2018](#); [White et al., 2020](#); all studies used the same metrics). In [Figure 8A](#), the x - and y -axes show single- and dual-task accuracy, respectively. The crossed squares represent results from experiments involving judgments of a simple feature (color); the open diamonds represent results from experiments involving semantic categorization of words; the closed circles represent results from the experiments in the current study; and the lines represent the predictions of the benchmark models. Results from our [Experiments 1](#) and [2](#) fall nearest the results from word judgments, indicating that words and objects show a similar large dual-task deficit under divided attention. In [Figure 8B](#), the x -axis shows the magnitude of the dual-task deficit and the y -axis shows the response correlation for the same studies. Results from our [Experiments 1](#) and [2](#) fall in between the results from simple feature judgments and the results from word judgments. Thus, while object judgments show a large dual-task deficit, overall, the divided attention effects are smaller than found with words.

Working hypothesis

Our working hypothesis is that semantic categorization of nameable objects is constrained by a serial bottleneck in perceptual processing. Because the results fell short of the benchmark all-or-none serial model prediction, we considered a more general model that can predict any magnitude of dual-task deficit or negative correlation. The serial model with partial switching represents a bottleneck where processing can occur for only one stimulus at a time. Moreover, on some proportion of trials there is enough time to switch to processing a second stimulus. For example, if 100% of trials allow processing of the second stimulus, the prediction becomes identical to the independent parallel model (no dual-task deficit). Conversely, if 0% of trials allow processing of the second stimulus, the prediction becomes identical to the all-or-none model with no switching (a large dual-task deficit). Intermediate proportions of trials with processing of the second stimulus produce intermediate dual-task deficits. Changing the specific proportion of trials in which only one stimulus is processed allows the model to predict any magnitude of the dual-task deficit; and

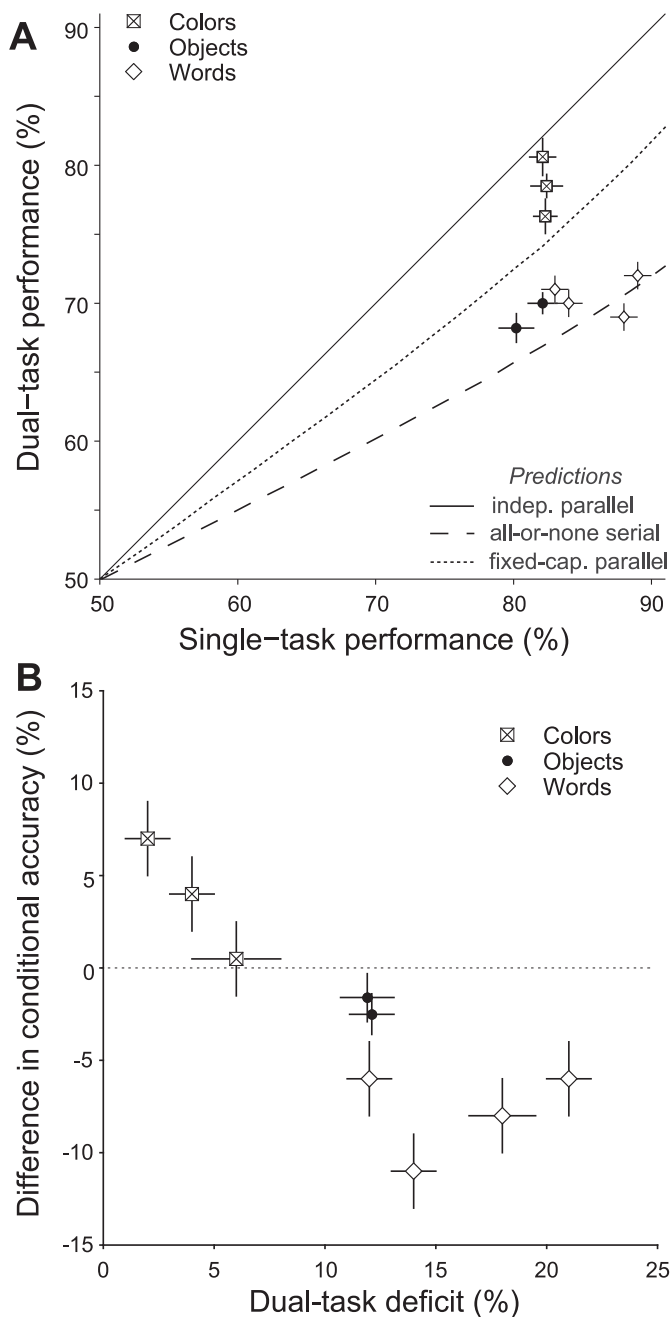


Figure 8. Summary of divided attention effects in object, feature, and word judgments. **(A)** Relationship between single- and dual-task performance. **(B)** Relationship between dual-task deficit and conditional accuracy. Solid circles: object judgments (present study). Crossed squares: color judgments (White et al., 2018; White et al., 2020). Open diamonds: word judgments (White et al., 2018; White et al., 2020). Dotted line: no difference in conditional accuracy predicted by parallel models. Error bars: standard error of the mean.

we can interpret our results in this context by estimating this proportion from the model using the observed performance. The dual-task deficit magnitude in both experiments can be described by switching on about 20% of trials, and processing of only one stimulus on

the remaining 80% of trials. The observed negative correlation was consistent with the predictions of a partial switching model where only one stimulus was processed on about 60% of trials (Experiment 1) or 50% of trials (Experiment 2). More generally, the results are consistent with a model where in at least half of the trials only one stimulus can be processed. By this interpretation, object processing does not rely on only parallel processes, like simple features, and instead is constrained by a serial bottleneck. Objects differ from words in that more than one object can be processed on some trials despite the brief masked displays.

This working hypothesis can be further extended to account for the observed two-target effect. Sometimes this effect is taken as a sign of parallel processing and “late selection” (Duncan, 1980). In fact, a small modification of the partial switching model can account for the two-target effect: a reduction in the proportion of trials in which the second stimulus is processed after a target has been processed, compared with after a distractor. In the extreme, a second stimulus might be processed only on trials in which a distractor is processed first. An alternative account is to adapt the two-stage models proposed by Duncan (1980) or Corbett and Smith (2017) by appending them to a first stage consisting of the partial switching model. When the first stage processes two targets, information about the targets is subject to target-specific interference at the second stage (e.g. limited memory encoding), producing the two-target effect. Without two targets, there would be no effect of the second stage. Either of these accounts can yield the modest 3% two-target effects found in the current experiments.

An alternative hypothesis

An alternative to our working hypothesis is that semantic categorization of objects is mediated by some kind of limited-capacity parallel process. While the divided attention effects for objects are larger than the predictions of the two parallel models we used as benchmarks, a discrete fixed-capacity parallel model can capture the magnitude of the dual-task deficit. This model is similar to the fixed-capacity parallel model, but information from the stimulus informs two discrete states: “detect target” or “detect no target” (Luce, 1963; Swagman et al., 2015; see Appendix for details). The discrete model predicts a dual-task deficit magnitude similar to what we observed in this study. However, like other parallel models, the discrete model predicts no response correlation. To predict a negative correlation between the two responses, one can add parameter variability to the attention parameter that assigns the relative number of samples to one task or the other task (see the section at the end of the Appendix on response correlation). If this parameter varies from trial

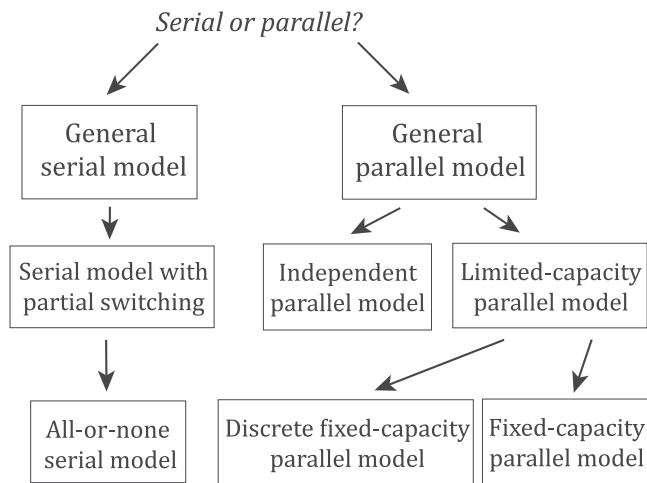


Figure 9. Summary of the relevant models. Arrows lead from more general to more specific models.

to trial, then some trials have higher performance for one response and other trials have higher performance for the other response. This modification can predict the observed small negative correlation between the two responses. But unlike the other changes to the model, this change is ad hoc.

This alternative hypothesis can also be extended to account for the two-target effect. One can adapt the two-stage models proposed by Duncan (1980) or Corbett and Smith (2017). These models add a second stage (e.g. limited memory encoding) to the simple parallel perception models described here. For example, after a target is processed, there is a memory encoding process that delays processing of additional targets, but not distractors. Alternatively, the two-target effect can be accounted for by attenuating the gain of a stimulus in a target context, relative to a distractor context (analogous to crosstalk, but opposite in sign). In summary, there are many ways to modify a parallel model to yield the two-target effects found in the current studies. Our larger point is that by themselves, two-target effects are not diagnostic for distinguishing between parallel and serial models.

Figure 9 summarizes the specific models considered and tested here in the context of other general models of processing. In the present study, we provide evidence rejecting all of the common specific models suggested for object processing: the independent parallel model, the fixed-capacity parallel model, and the all-or-none serial model. Our working hypothesis consists of a serial model with partial switching. This model allows switching on some proportion of trials. It provides the most parsimonious explanation of the results. However, we cannot yet rule out an alternative: a discrete fixed-capacity parallel model, with an ad hoc addition such that on some trials attention is allocated unequally between the two stimulus locations. Although the

current data cannot be used to definitively distinguish these possibilities, future studies could be targeted to accomplish this: for example, a redundant-target experiment could provide further evidence for or against serial processing (Mullin, Egeth, & Mordkoff, 1988; Mullin & Egeth, 1989; Shepherdson & Miller, 2014).

Stepping back from the details of the models, this study informs a larger set of questions about perception within a single fixation (or single brief display). Although there are many examples of likely parallel processing (e.g. contrast detection; Bonnel et al., 1992; Scharff et al., 2011), there are few examples of likely serial processing. Early work suggesting serial processing in visual search has been rejected both because of mimicry between serial and parallel processes (e.g. Townsend, 1990; Palmer et al., 2000) and because the detailed predictions of serial models have not been satisfied (e.g. Ward & McClelland, 1989). In contrast, White et al. (2018) proposed that the perception of words is a good example of serial processing. Evidence from dual-task paradigms in that study builds on the important existing evidence from the redundant target experiments of Mullin and Egeth (1989). The research in this article extends this example to nameable objects. We intend future studies to further evaluate the case of words and objects as an example of perception limited by a serial process.

Relationship to other behavioral studies

The results of our study are compatible with previous literature suggesting that processes governing object perception have limited capacity (Potter & Fox, 2009; Scharff et al., 2011). Our results are similar to but smaller than the effects of divided attention reported for the semantic categorization of words. White et al. (2018 and 2020) found a large dual-task deficit and a negative correlation consistent with the prediction of the all-or-none serial model. The methods used in the present study are especially compatible to allow a direct comparisons of effect magnitudes to those found by White and colleagues, as summarized in Figure 8.

We see our results as broadly compatible with the literature on the automaticity of perceptual processes (Shiffrin & Schneider, 1977). Our experiments use familiar objects, but the specific images are not particularly familiar. While one might find parallel processing of a particular familiar feature, it is less likely that there would be parallel processing of the diverse images of a familiar object. Therefore, we see no conflict between finding evidence of serial processing for recognition of familiar objects and theories of automaticity.

A different point of contact to this literature is Cousineau and Shiffrin (2004). In this study, the authors

created a task with difficult discriminations involving the relative position of multiple features (4 spokes around a central circle). Their goal was to find a task that required serial processing in visual search. Using a response time paradigm, the results are among the most convincing for serial processing in visual search. Thus, that paper is one of the closest response time analogues to the current study.

In contrast to these studies, recent work with natural scene stimuli has proposed that processing in complex visual recognition tasks might be parallel. For example, Thorpe and colleagues (Thorpe, Fize, & Marlot, 1996) investigated the speed of visual recognition using objects in natural scenes. A natural scene stimulus was presented briefly for 20 ms; and the task was to report the presence of an animal in that scene by releasing a button. Participants were very accurate while maintaining fast reaction times, which the authors suggest reveals an underlying rapid and efficient process for recognition of objects in complex scenes. We point out that although objects might be highly discriminable, this observation does not necessarily imply a reduced effect of divided attention.

Relevant to divided attention, Rousselet and colleagues used the same task to ask whether this processing occurs in parallel when two or four scenes are presented simultaneously (Rousselet, Thorpe, & Fabre-Thorpe, 2004). The authors observed set-size effects in accuracy and response time when multiple scenes were presented simultaneously. They argued that these divided attention effects were consistent with an unlimited-capacity parallel model with late selection, but the authors did not consider, nor rule out, alternative serial models. In summary, Rousselet and colleagues make the case that parallel processes can plausibly underlie object recognition in their experiment, but do not distinguish parallel and serial model predictions for their task. Indeed, that remains a largely unsolved problem for such visual search tasks that use response time.

The neural basis of divided attention effects

Visual word form area (VWFA) has been proposed as the locus of the serial bottleneck for words (White et al., 2019). Although simple features can be processed using information in earlier visual cortex (from V1 to V4), word processing depends on VWFA. White and colleagues present evidence that signals in the anterior part of VWFA can represent only one word at a time. This is a candidate neural correlate of a serial bottleneck, but does not rule out additional bottlenecks elsewhere in processing, for example, in semantic processing within the language areas.

Analogous to the VWFA, the candidate locus for a serial bottleneck for objects is likely to be in the ventral

pathway beyond retinotopic cortex. One candidate is lateral occipital complex (LOC). In this area, signals represent complex object characteristics such as category, and thus underlie more complex judgments than simple features (Kourtzi & Kanwisher, 2000; Eger, Ashburner, Haynes, Dolan, & Rees, 2008). Retinotopy is weaker and receptive field sizes are larger in this area than in early visual cortex, opening up the possibility that two objects cannot be represented at the same time; that is, neural responses carry information about only one of the two objects (Larsson & Heeger, 2006; Cichy, Chen, & Haynes, 2011). Another candidate is anterior or anteromedial temporal cortex. These regions are downstream of LOC, and are known to be involved in object-based and semantic processing (Moss, Rodd, Stamatakis, Bright, & Tyler, 2005; Patterson, Nestor, & Rogers, 2007).

Comparison of processing for words and objects

Is processing different for words and objects? Several aspects of word processing may be unique; for example, word judgments require lexical access, whereas many object judgments do not (although there may be some overlap for object naming tasks; Biggs & Marmurek, 1990). Word processing also appears to have a visual hemifield effect: English words shown in the right hemifield are recognized more effectively than in the left (Chiarello, 1988; Bub & Lewine, 1988; Brysbaert, Vitu, & Schroyens, 1996; Simola, Holmqvist, & Lindgren, 2009). Nameable objects generally do not show such asymmetry (Biederman & Cooper, 1991; but see McAuliffe & Knowlton, 2001).

Word and object recognition processes also appear to be nonoverlapping at the conceptual processing level. For example, Endress and Potter (2012) asked participants to recognize a scene or its verbal description while performing a simultaneous secondary task: scenes were presented together with words, or non-word stimuli. Scene understanding was impaired only when the secondary task involved non-word stimuli, suggesting that the processing of words is distinct from the processing of scenes.

Conclusion

In this study, we examined whether object perception showed large divided attention effects like those found for words, or little or no divided attention effects like those found for simple features. We used a semantic categorization task with nameable object stimuli and two presentation paradigms: RSVP and brief masking. There was a large divided attention effect and a negative correlation in responses for the semantic categorization

of two nameable objects. These results are consistent with a serial model in which only one stimulus is processed on most, but not all trials. The results are also consistent with an alternative discrete fixed-capacity parallel process with differential allocation of attention. In conclusion, the effects of divided attention for objects are smaller than observed for words, but might still reflect a serial bottleneck in object processing.

Keywords: divided attention, serial processing, object perception

Acknowledgments

The authors thank Justin Harshman for assistance in collecting these data, and Miranda Johnson and Alex White for helpful discussions of the results.

Supported in part by Grants from the National Eye Institute (F32 EY030320 to D.V.P. and EY12925 to G.M.B.).

The data reported in this article are available in the Open Science Framework repository (<https://osf.io/p3yan/>).

Commercial relationships: none.

Corresponding author: Dina V. Popovkina.

Email: dina4@uw.edu.

Address: Department of Psychology, University of Washington, 3920 15th Avenue NE, Seattle, WA 98195, USA.

References

- Akyürek, E. G., & Hommel, B. (2005). Short-term memory and the attentional blink: Capacity versus content. *Memory & Cognition*, *33*(4), 654–663.
- Allport, A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. *Perspectives on Perception and Action*, *15*, 395–419.
- Attarha, M., Moore, C. M., & Vecera, S. P. (2014). Summary statistics of size: Fixed processing capacity for multiple ensembles but unlimited processing capacity for single ensembles. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1440.
- Biederman, I., Blickle, T. W., Teitelbaum, R. C., & Klatsky, G. J. (1988). Object search in nonscene displays. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 456.
- Biederman, I., & Cooper, E. E. (1991). Evidence for complete translational and reflectional invariance in visual object priming. *Perception*, *20*(5), 585–593.
- Biggs, T. C., & Marmurek, H. H. (1990). Picture and word naming: Is facilitation due to processing overlap? *The American Journal of Psychology*, *103*(1 Spring 1990), 81–100.
- Bonnel, A. M., & Prinzmetal, W. (1998). Dividing attention between the color and the shape of objects. *Perception & Psychophysics*, *60*(1), 113–124.
- Bonnel, A. M., Stein, J. F., & Bertucci, P. (1992). Does attention modulate the perception of luminance changes? *The Quarterly Journal of Experimental Psychology Section A*, *44*(4), 601–626.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.
- Broadbent, D. E. (1958). *Perception and communication*. Elmsford, NY: Pergamon Press US.
- Brysbaert, M., Vitu, F., & Schroyens, W. (1996). The right visual field advantage and the optimal viewing position effect: On the relation between foveal and parafoveal word recognition. *Neuropsychology*, *10*(3), 385.
- Bub, D. N., & Lewine, J. (1988). Different modes of word recognition in the left and right visual fields. *Brain and Language*, *33*(1), 161–188.
- Chen, Y., & Seidemann, E. (2012). Attentional modulations related to spatial gating but not to allocation of limited resources in primate V1. *Neuron*, *74*(3), 557–566.
- Chiarello, C. (1988). Lateralization of lexical processes in the normal brain: A review of visual half-field research. In *Contemporary Reviews in Neuropsychology* (pp. 36–76). New York, NY: Springer.
- Cichy, R. M., Chen, Y., & Haynes, J. D. (2011). Encoding the identity and location of objects in human LOC. *Neuroimage*, *54*(3), 2297–2307.
- Coltheart, V. (Ed.). (1999). *Fleeting memories: Cognition of brief visual stimuli*. New York, NY: MIT Press.
- Corbett, J., & Smith, P. L. (2017). The magical number one-on-square-root-two: The double-target detection deficit in brief visual displays. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(7), 1376.
- Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 613–617.
- Cousineau, D., & Shiffrin, R. (2004). Termination of a visual search with large display size effects. *Spatial Vision*, *17*(4), 327–352.

- Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, *87*(3), 272.
- Eger, E., Ashburner, J., Haynes, J. D., Dolan, R. J., & Rees, G. (2008). fMRI activity patterns in human LOC carry information about object exemplars within category. *Journal of Cognitive Neuroscience*, *20*(2), 356–370.
- Endress, A. D., & Potter, M. C. (2012). Early conceptual and linguistic processes operate in independent channels. *Psychological Science*, *23*(3), 235–245.
- Forster, K. I. (1970). Visual perception of rapidly presented word sequences of varying complexity. *Perception & Psychophysics*, *8*(4), 215–221.
- Graham, N., Kramer, P., & Haber, N. (1985). In Posner, M. I., & Marison, O. S. M. (Eds). Attending to the spatial frequency and spatial position of near-threshold visual patterns. *Mechanisms of Attention and Performance XI*, 269–284.
- Graham, N., Robson, J. G., & Nachmias, J. (1978). Grating summation in fovea and periphery. *Vision Research*, *18*(7), 815–825.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics (Vol. 1)*. Hoboken, NJ: Wiley.
- Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219.
- Kellen, D., Erdfelder, E., Malmberg, K. J., Dubé, C., & Criss, A. H. (2016). The ignored alternative: An application of Luce’s low-threshold model to recognition memory. *Journal of Mathematical Psychology*, *75*, 86–95.
- Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General*, *139*(3), 558.
- Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience*, *20*(9), 3310–3318.
- Krantz, D. H. (1969). Threshold theories of signal detection. *Psychological Review*, *76*(3), 308.
- Larsson, J., & Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *Journal of Neuroscience*, *26*(51), 13128–13142.
- Le, R., Witthoft, N., Ben-Shachar, M., & Wandell, B. (2017). The field of view available to the ventral occipito-temporal reading circuitry. *Journal of Vision*, *17*(4), 6.
- Luce, R. D. (1963). A threshold theory for simple detection experiments. *Psychological Review*, *70*(1), 61.
- McAuliffe, S. P., & Knowlton, B. J. (2001). Hemispheric differences in object identification. *Brain and Cognition*, *45*(1), 119–128.
- Miller, J., & Bonnel, A. M. (1994). Switching or sharing in dual-task line-length discrimination? *Perception & Psychophysics*, *56*(4), 431–446.
- Moss, H. E., Rodd, J. M., Stamatakis, E. A., Bright, P., & Tyler, L. K. (2005). Anteromedial temporal cortex supports fine-grained differentiation among objects. *Cerebral Cortex*, *15*(5), 616–627.
- Mullin, P. A., & Egeth, H. E. (1989). Capacity limitations in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(1), 111.
- Mullin, P. A., Egeth, H. E., & Mordkoff, J. T. (1988). Redundant-target detection and processing capacity: The problem of positional preferences. *Perception & Psychophysics*, *43*(6), 607–610.
- Navon, D., & Miller, J. (1987). Role of outcome conflict in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(3), 435.
- Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*(10-12), 1227–1268.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), 976–987.
- Potter, M. C., & Fox, L. F. (2009). Detecting and remembering simultaneous pictures in a rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 28.
- Potter, M. C., & Haggmann, C. E. (2015). Banana or fruit? Detection and recognition across categorical levels in RSVP. *Psychonomic Bulletin & Review*, *22*(2), 578–585.
- Rauschecker, A. M., Bowen, R. F., Parvizi, J., & Wandell, B. A. (2012). Position sensitivity in the visual word form area. *Proceedings of the National Academy of Sciences*, *109*(24), E1568–E1577.
- Robinson, A. K., Grootswagers, T., & Carlson, T. A. (2019). The influence of image masking on object representations during rapid serial visual presentation. *NeuroImage*, *197*, 224–231.
- Rousselet, G. A., Thorpe, S. J., & Fabre-Thorpe, 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. (2011). Evidence of fixed capacity in visual object categorization. *Psychonomic Bulletin & Review*, *18*(4), 713–721.
- Shaw, M. L. (1980). Identifying attentional and decision-making components. *Attention and Performance VIII*, *8*, 277.
- Shepherdson, P., & Miller, J. (2014). Redundancy gain in semantic categorisation. *Acta Psychologica*, *148*, 96–106.
- Shiffrin, R. M., & Gardner, G. T. (1972). Visual processing capacity and attentional control. *Journal of Experimental Psychology*, *93*(1), 72.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127.
- Simola, J., Holmqvist, K., & Lindgren, M. (2009). Right visual field advantage in parafoveal processing: Evidence from eye-fixation-related potentials. *Brain and Language*, *111*(2), 101–113.
- Smith, P. L., Lilburn, S. D., Corbett, E. A., Sewell, D. K., & Kyllingsbæk, S. (2016). The attention-weighted sample-size model of visual short-term memory: Attention capture predicts resource allocation and memory load. *Cognitive Psychology*, *89*, 71–105.
- Sperling, G., & Melchner, M. J. (1978). The attention operating characteristic: Examples from visual search. *Science*, *202*(4365), 315–318.
- Sun, P., Chubb, C., Wright, C. E., & Sperling, G. (2016). Human attention filters for single colors. *Proceedings of the National Academy of Sciences*, *113*(43), E6712–E6720.
- Swagman, A. R., Province, J. M., & Rouder, J. N. (2015). Performance on perceptual word identification is mediated by discrete states. *Psychonomic Bulletin & Review*, *22*(1), 265–273.
- Taylor, M. M., Lindsay, P. H., & Forbes, S. M. (1967). Quantification of shared capacity processing in auditory and visual discrimination. *Acta Psychologica*, *27*(1967), 223–229.
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522.
- Townsend, J. T. (1971). A note on the identifiability of parallel and serial processes. *Perception & Psychophysics*, *10*(3), 161–163.
- Townsend, J. T. (1990). Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science*, *1*(1), 46–54.
- Ward, R., & McClelland, J. L. (1989). Conjunctive search for one and two identical targets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(4), 664.
- White, A. L., Runeson, E., Palmer, J., Ernst, Z. R., & Boynton, G. M. (2017). Evidence for unlimited capacity processing of simple features in visual cortex. *Journal of Vision*, *17*(6), 19.
- White, A. L., Palmer, J., & Boynton, G. M. (2018). Evidence of serial processing in visual word recognition. *Psychological Science*, *29*(7), 1062–1071.
- White, A. L., Palmer, J., Boynton, G. M., & Yeatman, J. D. (2019). Parallel spatial channels converge at a bottleneck in anterior word-selective cortex. *Proceedings of the National Academy of Sciences*, *116*(20), 10087–10096.
- White, A. L., Palmer, J., & Boynton, G. M. (2020). Visual word recognition: Evidence for a serial bottleneck in lexical access. *Attention, Perception, & Psychophysics*, *82*(4), 2000–2017.
- Yantis, S., & Johnston, J. C. (1990). On the locus of visual selection: Evidence from focused attention tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 135.

Appendix

Performance metrics

Performance was calculated as percent area under the Receiver Operating Characteristic (ROC) curves. Like percent correct, this metric ranges from 50% (chance) to 100% (perfect); and like d' , it is an unbiased estimate of performance. These curves were constructed by plotting proportion of hits against the proportion of false alarms. The aggregate ROC curves for each experiment are shown in [Figure A1](#). The x-axis shows the false alarm rate, and the y-axis shows the hit rate. The filled blue circles represent single-task data and the open red squares represent dual-task data. The positive diagonal represents chance performance (50% of the area lies under this diagonal).

As usual for rating data, the points defining each curve are derived from differentially grouping the participant responses (“likely yes”, “guess yes”, “guess no”, and “likely no”) to produce different criterion levels for treating a response as a “yes”. For example, treating the first three responses as a “yes” produces the point towards the top of the plot; treating the first two responses as a “yes” produces the middle point; and treating only the first response as a “yes” produces the point toward the left of the plot. Lines are connected from the origin (0,0) through the three points and then to (1,1). The percent area under these connected lines defines our measure of accuracy. The effect of divided

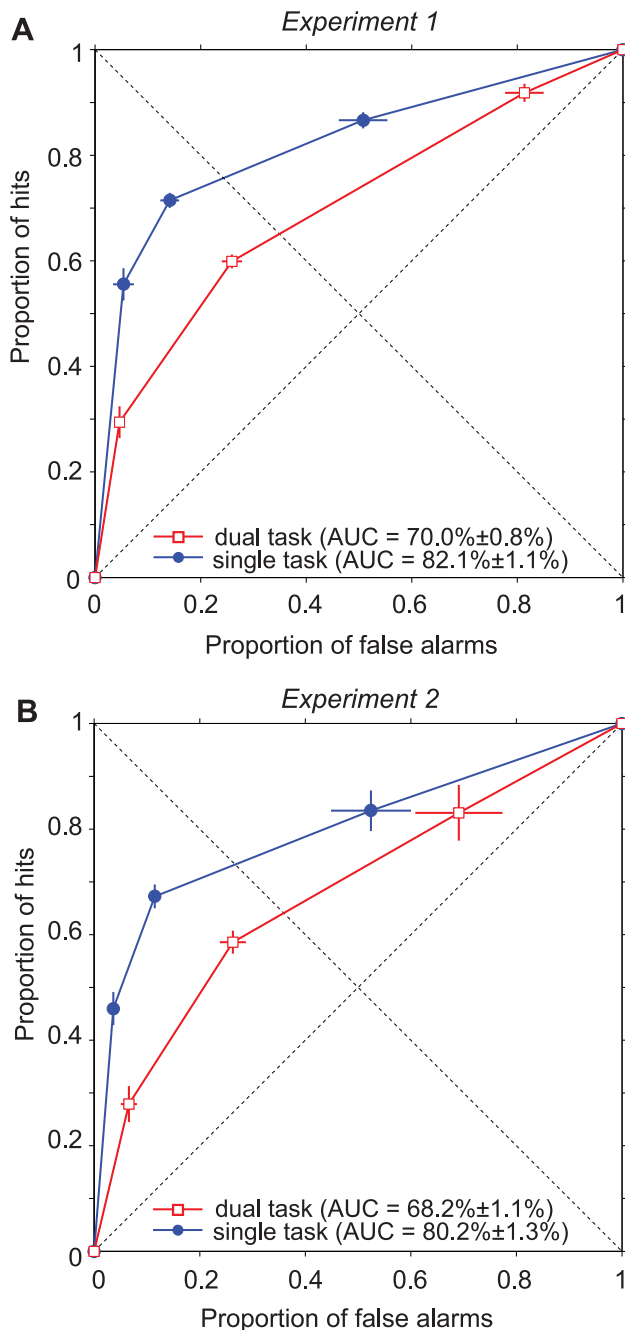


Figure A1. Average ROC curves for Experiments 1 and 2 (A and B, respectively). Single (blue) and dual (red) tasks are shown. Error bars: standard error of the mean. AUC: Area under the curve.

attention is a shift of the entire ROC curve in the dual task, relative to the single task; and the area under the dual-task curve is smaller than the area under the single-task curve.

These aggregate curves are representative of individual data: participants were largely consistent in their confidence ratings, as suggested by the small error bars. Most of the difference between individual

participants was in their use of the more extreme rating “likely no”. Across participants, there was also a consistent bias towards responding “no” more often than “yes” (unbiased performance produces points along the negative diagonal; see below for a detailed analysis of response bias in each experiment).

Additional secondary results for Experiment 1

Effect of response bias. The percent of “yes” responses was similar in the single- and dual-task conditions ($42.8\% \pm 1.4\%$ and $42.9\% \pm 1.5\%$, respectively). The difference was $0.1\% \pm 1.5\%$ (not significantly different from zero, $t(11) = 0.07$, $p = 0.94$). Thus, dividing attention did not affect response bias.

Effect of stimulus location. Accuracy was similar for both locations, both in the single-task condition (top location: $81.6\% \pm 1.3\%$ vs bottom location: $82.6\% \pm 1.1\%$) and in the dual-task condition (top location: $70.1\% \pm 1.0\%$ vs. bottom location: $69.9\% \pm 1.3\%$). The aggregate difference was $0.4\% \pm 1.2\%$ (not significantly different from zero, $t(11) = -0.29$, $p = 0.78$).

Additional secondary results for Experiment 2

Effect of response bias. The percent of “yes” responses was similar in the single- and dual-task conditions ($39.3\% \pm 1.4\%$ and $42.4\% \pm 2.1\%$, respectively). The difference was $3.1\% \pm 1.8\%$ (not significantly different from zero, $t(10) = 1.74$, $p = 0.11$). Thus, the task conditions had little or no effect on response bias.

Effect of stimulus location. Accuracy was similar for both locations, both in the single-task condition (top location: $79.6\% \pm 1.4\%$ vs bottom location: $80.7\% \pm 1.2\%$) and in the dual-task condition (top location: $69.2\% \pm 1.6\%$ vs. bottom location: $67.2\% \pm 1.5\%$). The aggregate difference was $0.4\% \pm 1.2\%$ (not significantly different from zero, $t(10) = 0.32$, $p = 0.76$).

Two-target effect or a congruency effect?

Among the results described for each experiment was an analysis of possible two-target effects, which compared accuracy in one location when there was a target present in the other location, to accuracy when there was no target present in the other location. For both experiments and for both single and dual tasks, this comparison yielded between 3 and 4% worse performance in the context of a target. Such a comparison is a common way to measure two-target effects (e.g. Duncan, 1980). However, this analysis does not clearly distinguish between two-target effects (Duncan, 1980) and congruency effects (Navon &

	Irrelevant target context (<i>T</i>)	Irrelevant distractor context (<i>D</i>)	<i>T-D</i>
Relevant target (hits)	57.0 ± 1.3%	62.8 ± 1.6%	−5.8 ± 1.3%
Relevant distractor (correct rejections)	72.5 ± 1.8%	75.7 ± 2.4%	−3.1 ± 1.7%

Table 1. **Experiment 1**: Percent Hit and Correct Rejection by Target or Distractor Context

	Irrelevant target context (<i>T</i>)	Irrelevant distractor context (<i>D</i>)	<i>T-D</i>
Relevant target (hits)	57.6 ± 2.2%	59.5 ± 2.5%	−2.0 ± 1.8%
Relevant distractor (correct rejections)	68.7 ± 2.8%	78.7 ± 2.9%	−10.1 ± 3.0%

Table 2. **Experiment 2**: Percent hit and correct rejection by target or distractor context.

Miller, 1987). Moreover, it does not show that the effect is specific to there being two targets, rather than a general interference effect of a target on any other stimulus.

Both the two-target effect and the congruency effect are context effects of one location on the other. They differ in whether two targets hurt performance or improve performance. To distinguish these possibilities, one must break down the results by both the different contexts and by the relevant stimulus. Because targets and distractors must be considered separately, we must distinguish hits and correct rejections rather than use a measure that combines them, such as the area under the ROC. The results of such an analysis for **Experiment 1** are presented in **Table 1**. For simplicity, we consider in this analysis only dual-task trials. The rows specify the relevant stimulus: target or distractor. The columns specify the irrelevant stimulus context present on the other side of the display for the other task: target (*T*) or distractor (*D*). The final column is the difference in performance for a stimulus with a target context minus performance for a stimulus with a distractor context (*T-D*).

For **Experiment 1**, the comparison of different contexts shows a deficit for the target context compared to the distractor context. To distinguish whether this context effect is due to a two-target effect or to a congruency effect, one must examine the upper right cell: *T-D* for a relevant target. For a two-target effect, the irrelevant target context reduces performance and this cell should be negative. For a congruency effect, the irrelevant target context matches the relevant stimulus to improve performance and this cell should be positive. The observed difference is a negative value of about 6%. Furthermore, this difference is reliable ($t(11) = -4.51$, $p < 0.001$). This is consistent with a two-target effect rather than a congruency effect.

The difference in the lower right cell is harder to interpret. A two-target effect that is specific to targets predicts no difference, but a more general target

interference effect that also affects distractors predicts a negative value. A congruency effect also predicts a negative value in the lower right cell. Consequently, based on this cell one cannot distinguish a congruency effect from a more general target interference effect. For this experiment, the result for this cell was modestly negative, which does not distinguish the possibilities.

For **Experiment 2**, the comparison of the different contexts is shown in **Table 2**. As before, it shows a deficit for the target context compared to the distractor context. For the critical upper right cell, there is a negative value of 2%. Thus, this experiment is also consistent with a two-target effect rather than a congruency effect. But in this case, the effect is not reliable ($t(10) = -1.11$, $p = 0.29$). This experiment also shows a relatively large difference in the lower right cell, which is not consistent with the strict version of the two-target effect which predicts zero for this cell. Thus, the data suggest that a “pure” two-target effect cannot account for the observed effects. The possibilities that can account for the effects include a more general target interference effect, or the combination of a two-target effect and a congruency effect. More extensive experiments are needed to sort out these possibilities.

Serial model with partial switching

This is a generalized version of the all-or-none serial model. It allows a mixture of trials where only one stimulus is processed (all-or-none trials) and other trials where both stimuli are processed (switching trials). This model was implemented as described in **White et al. (2020)**.

When all trials are those in which only one stimulus is processed, this model predicts the same magnitude of dual-task deficit and difference in conditional accuracy as the all-or-none serial model. When all trials are those in which both stimuli are processed, this model predicts

the same magnitude of effects as the independent parallel model. For intermediate proportions of trials in which only one stimulus is processed, this model predicts intermediate effect magnitudes.

Discrete fixed-capacity parallel model

Models with discrete rather than continuous representations have a long history. Early work is reviewed in Green and Swets (1966) and a more recent discussion is in Kellen, Erdfelder, Malmberg, Dube, and Criss (2016). Our interest was sparked by Swagman et al. (2015) that make the case for discrete states in word identification. Perhaps object identification is similar. Our extension of the discrete models combines ideas from Luce's (1963) two-state low-threshold model, and Shaw's (1980) sample size model.

The two-state, low-threshold model. To begin, consider the idea of discrete states. Instead of representing the relevant stimulus information by a single continuous value as in signal detection theory, the relevant information is represented by two discrete states. For this model, the presence or absence of the relevant stimulus (a target) is represented by a corresponding *detect state D* or a *no-detect state not-D*. The special feature of Luce's (1963) model, relative to the well-known high-threshold model, is that errors can occur when in either of the states. Given a target, the probability of being in state *D* is q_t and the probability of being in state *not-D* is $1 - q_t$. Given a distractor, the probability of being in state *D* is q_d and the probability of being in state *not-D* is $1 - q_d$. This model results in an ROC with a single point at the location defined by the probability of hit q_t and the probability of a false alarm q_d . If $q_d = 0$, then this model simplifies to the high-threshold model.

To allow for response bias, one needs to add a guessing mechanism. A guessing parameter serves the same function as the criterion parameter in signal detection models. Because of the two states in this discrete model, there are two kinds of guesses. If in the detect state, one might guess “no” and if in the not-detect state one might guess “yes”. Thus, one can adjust response bias by modifying the responses when in the detect state or by modifying the responses when in the no-detect state. We assume an optimal guessing strategy. The strategy can be understood by starting with the extreme case in which one always responds “no” regardless of the state. What kind of guess should one make given this start point to optimize performance? As long as performance is above chance, the better strategy is to add “yes” responses when in the detect state. Once one is always responding “yes” when in the detect state and always responding “no” when in the no-detect state, one has reached the point on the ROC defined by (q_d, q_t) . To further increase “yes”

responses, the only choice is to guess “yes” when in the not-detect state. In summary, the optimal guessing strategy is to guess “no” from the detect state when being conservative about responding “yes”, and to guess “yes” from the no-detect state when being liberal. This guessing strategy can be implemented with a guessing parameter g that goes from 0 to 1 and switches from conservative to liberal guessing at the value $g = .5$. Similar to Kellen et al. (2016), the predicted hits and false alarms are:

$$\begin{aligned} p_{hit} &= q_t + q_t(2g - 1), & \text{if } g < 0.5, \\ &= q_t + (1 - q_t)(2g - 1), & \text{if } g \geq 0.5, \text{ and} \\ p_{FA} &= q_d + q_d(2g - 1), & \text{if } g < 0.5, \\ &= q_d + (1 - q_d)(2g - 1), & \text{if } g \geq 0.5. \end{aligned} \quad (1)$$

A special case with a single sensitivity parameter.

Our next step is to identify the sensitivity parameter of the model that varies with the stimulus. Luce (1963) intentionally did not specify the role of his two parameters to keep the model general. But for our purpose, we want a simple model in which there is one parameter that describes sensitivity as a function of the stimulus and then compare it to signal detection theory with its single sensitivity parameter (d'). One possibility is to let q_t vary with the stimulus and fix q_d . But as discussed by Krantz (1969) that is too restrictive and prevents one from describing a full range of performance by varying q_t with the stimulus and keeping a fixed q_d . Instead, we restrict the model to be symmetric such that $q_d = 1 - q_t$. This restriction is intended to be analogous to signal detection theory when simplified by assuming equal-variance distributions.

The next step is to add further assumptions about how the model depends on the stimulus. For the high-threshold model, the initial state without the target is always the not-detect state. As a function of the stimulus, there is some probability of entering the detect state. For our symmetrical case, we assume the probability of the initial states is equal. The probability of entering the detect state is denoted by the sensitivity parameter d which can vary from 0 to 1. Specifically, given a target, d is the probability of entering the detect state *D*. Otherwise one remains in an initial state. Similarly, given a distractor, d is the probability of entering the no-detect state and otherwise one remains in an initial state. The probability d is assumed to be independent of the initial state, so using the definition of independent joint probabilities:

$$\begin{aligned} p(D|\text{target}) &= 0.5 + d - (0.5d) \\ &= 0.5 + 0.5d. \end{aligned} \quad (2)$$

These values can be converted into the parameters used by Luce of q_t and q_d (hit and false alarm

probabilities):

$$\begin{aligned} q_t &= p(D|\text{target}) = 0.5 + 0.5d, \text{ and} \\ q_d &= 1 - q_t. \end{aligned} \quad (3)$$

To summarize, we pursue the special case of a symmetric, low-threshold model with a single sensitivity parameter that gives the probability of the stimulus moving one into the appropriate state for that stimulus.

Adding the effects of attention. The next step is to modify the sample size model (Shaw, 1980) to these discrete representations. Recall that the sample size model has some number samples, m , that are distributed over the relevant stimuli to represent the effect of attention. Each sample is assumed to provide independent information about the stimulus. In the continuous case, the mean of multiple samples improves the estimate of the stimulus value. For the discrete case, we assume that each sample provides an independent chance of entering the detect state given a target. This is similar to the idea of probability summation (e.g. Graham, Robson, & Nachmias, 1978). Let $d_{1\text{sample}}$ be the probability of entering the detect state given a target and with only one sample. Then for r independent samples the probability of entering the detect state given the target becomes $1 - [1 - d_{1\text{sample}}]^r$. In words, this is one minus the joint probability of not detecting from r samples. From this equation, one can describe how the fraction of samples affects the probability of entering the detect state given the target. Let a be the fraction of samples allocated to Stimulus 1. Then $d(a)$ is given by

$$d(a) = 1 - [1 - d_{1\text{sample}}]^{am}. \quad (4)$$

The next step is to derive how $d_{1\text{sample}}$ is related to the sensitivity parameter d defined in the previous section. The d parameter is the sensitivity for a single task in which all of the samples are devoted to a single stimulus ($a=1$). Assuming $a=1$ and substituting d for $d(a)$ turns Equation 4 into $d = 1 - [1 - d_{1\text{sample}}]^m$. Solving for $d_{1\text{sample}}$ yields:

$$d_{1\text{sample}} = 1 - (1 - d)^{(1/m)}. \quad (5)$$

These equations can be combined to solve for $d(a)$ as a function of d and a . Equation 5 can be written as $[1 - d_{1\text{sample}}]^m = (1 - d)$. Next, Equation 4 can be rewritten using the power law of exponents as $d(a) = 1 - \{[1 - d_{1\text{sample}}]^m\}^a$. Now substitute the rewritten Equation 5

into Equation 4 to obtain the result:

$$d(a) = 1 - [1 - d]^a. \quad (6)$$

This result does not depend on the total number of samples m . It can be substituted into Equation 3 to obtain:

$$\begin{aligned} q_t &= 0.5 + 0.5[1 - (1 - d)]^a, \text{ and} \\ q_d &= 1 - q_t. \end{aligned} \quad (7)$$

In summary, the new model is described by Equation 1 from low-threshold theory and Equation 7 which specifies the effects of sensitivity and attention. The model has three parameters that are similar to parameters used in signal detection models:

- a sensitivity parameter, d ,
- a guessing parameter, g , and
- an attentional parameter a .

To give a numerical example of the predictions of this model, assume $d = 0.6$ and $g = 0.5$. A value of $a = 1$ (single-task condition), yields 80% correct and $a = 0.5$ (dual-task condition), yields 68.4% correct. This is a dual-task deficit of 11.6%. For the corresponding conditions, the dual-task deficit of the continuous fixed-capacity parallel model is 7.6%, and the dual-task deficit of the all-or-none serial model is 15%. Thus, for this performance level, the prediction of the discrete model is about halfway between these two benchmark models.

Response correlations. Like other parallel models, this discrete fixed-capacity parallel model predicts no correlation between responses. But, such a correlation can be generated from the model with the following addition. The parameter a describes how attention (samples) is allocated to the two stimuli. Specifically, the proportion of samples allocated to the first stimulus is a , and the proportion of samples allocated to the second stimulus is $1-a$. Because of this dependence, adding parameter variability to the attention parameter introduces a small negative correlation between responses. When the number of samples increases for one stimulus, it must decrease for the other. Such parameter variability can account for the small negative correlations found here for objects. But it cannot account for the larger negative correlations found for words by White et al. (2018, 2020).