Research Article

Evidence of Serial Processing in Visual Word Recognition

Psychological Science
1–10

DOI: 10.1177/0956797617751898
www.psychologicalscience.org/PS

Alex L. White, John Palmer, and Geoffrey M. Boynton
Department of Psychology, University of Washington

Abstract

To test the limits of parallel processing in vision, we investigated whether people can recognize two words at once. Participants viewed brief, masked pairs of words and were instructed in advance to judge both of the words (dual-task condition) or just one of the words (single-task condition). For judgments of semantic category, the dual-task deficit was so large that it supported all-or-none serial processing: Participants could recognize only one word and had to guess about the other. Moreover, participants were more likely to be correct about one word if they were incorrect about the other, which also supports a serial-processing model. In contrast, judgments of text color with identical stimuli were consistent with unlimited-capacity parallel processing. Thus, under these conditions, serial processing is necessary to judge the meaning of words but not their physical features. Understanding the implications of this result for natural reading will require further investigation.

Keywords

word recognition, divided attention, visual perception, visual attention, language, open data

Received 8/16/17; Revision accepted 12/1/17

Vision begins with parallel processing: The retina and early visual cortex encode many stimulus elements at once across the visual field. If independent and parallel processing continued all the way through the system, you would be able to perceive multiple objects simultaneously with no cost. At the other extreme, a serial bottleneck would allow only one stimulus to be recognized at a time, causing delays or errors when you must attend to multiple stimuli. Divided attention does often impair task performance (Braun, 1998; Carrasco, 2011), but the impairments are usually too small to be explained by serial processing.

Written words provide an important test of the limits of parallel processing, with clear applications to life in the modern world. Although crowding and poor peripheral acuity require a sequence of eye movements to scan a page, multiple words are still visible within each glance. Indeed, many words in a line of text are at least partially processed while the eyes fixate on the previous word, and some are skipped over by the eyes completely (Rayner, 2009). Within each glance, early visual mechanisms first encode the basic features of many letter shapes in parallel (Grainger, Dufau, & Ziegler, 2016). Eventually, representations of letter combinations lead to full semantic recognition. The present study was motivated by the following question: Does the visual system's parallel architecture allow for two words to be processed simultaneously, or is recognition constrained by a serial bottleneck?

The extent of parallel processing in natural reading has been fiercely debated (Murray, Fischer, & Tatler, 2013; Starr & Rayner, 2001). Some researchers argue for strictly serial processing along the line of text (Reichle, Liversedge, Pollatsek, & Rayner, 2009), and others argue for a graded allocation of attention and parallel processing of multiple words (Engbert, Nuthmann, Richter, & Kliegl, 2005; Kennedy, 2000). The parallel–serial debate has proved difficult to resolve with eye movement studies in natural reading.

To determine whether people can recognize two words simultaneously when they are forced to try, we

Corresponding Author:
Alex L. White, University of Washington, Department of Psychology, 119 Guthrie Hall, UW Box 351525, Seattle, WA 98195
E-mail: alexlw@uw.edu

abandoned natural reading conditions and focused on isolated word recognition during fixation. Our goal was to measure *capacity limits*—constraints on how much information the perceptual system can process per unit time. Previous studies have shown that word recognition is subject to capacity limits by using varieties of search paradigms (Harris, Pashler, & Coburn, 2004; Mullin & Egeth, 1989; Reichle, Vanyukov, Laurent, & Warren, 2008; Scharff, Palmer, & Moore, 2011). None of these visual search studies conclusively determined whether the capacity limit is due to a serial bottleneck (see the Discussion).

Rather than visual search, we adopted another classic paradigm: comparing accuracy in dual-task and single-task conditions. In two experiments, we presented participants with time-limited and masked pairs of parafoveal nouns. Participants fixated their gaze in the center and were instructed to detect targets on just one side (single-task condition) or to detect targets on both sides (dual-task condition). We then compared the relative deficit in the dual-task condition to the quantitative predictions of several models of parallel or serial processing.

In separate blocks of trials with identical stimuli and matched difficulty, participants detected either semantic targets or color targets. Semantic targets were nouns that belonged to a particular category, such as "professions." Color targets were nouns colored slightly reddish, and their semantic meaning was irrelevant. In our dual-task condition, participants always made the same type of judgment for both words (e.g., semantic-semantic or color-color). They never had to make a semantic judgment and a color judgment in the same trial.

Comparing semantic and color judgments allows us to test whether the capacity limit in divided attention depends on which stimulus aspect is task relevant. Previous studies have demonstrated that people have unlimited capacity for detecting changes in low-level features of nonlinguistic stimuli (Bonnel, Stein, & Bertucci, 1992; Scharff et al., 2011; White, Runeson, Palmer, Ernst, & Boynton, 2017). However, if word recognition is automatic (Augustinova & Ferrand, 2014; Stroop, 1935) and uses a common resource (Pastukhov, Fischer, & Braun, 2009), then we would predict similar dual-task deficits for both types of judgments.

In the first experiment, participants detected targets embedded in rapid-serial-visual-presentation (RSVP) streams (Fig. 1a). RSVP of single words has been studied as a way to present text that does not require saccades (Potter, 1984). On each trial, we presented five pairs of unrelated words in RSVP. The presentation rate was adjusted to keep each participant's single-task semantic performance below ceiling. Fast rates limit the time available to process each pair of words and reduce the likelihood of a serial shift of attention from one word to the other within one frame. To match color and semantic single-task difficulty levels, we adjusted the saturation of the red targets.

The second experiment tested whether the semantic dual-task deficit in Experiment 1 was due to the RSVP streams overloading memory or other cognitive mechanisms rather than to a limit on the immediate processing of each pair of words. Experiment 2 differed in that each trial contained only one masked pair of words (Fig. 1b).

Method

Experiment 1

Participants. Ten volunteers (4 female, ages 19–30 years) with normal or corrected-to-normal visual acuity and normal color vision participated in exchange for fixed monetary payment. All but 2 (including author A. L. White) were naive to the research aims. The sample size of 10 was chosen prior to data collection, on the basis of a pilot experiment, to produce standard deviations of the dual-task deficits near .01 (in A_g units, see below). This is small relative to the dual-task deficit of .08 A_g predicted by the fixed-capacity parallel-processing model described below (assuming equal division of attention).

Each participant gave informed consent in accordance with the Declaration of Helsinki and the University of Washington Institutional Review Board. All participants learned English as their first language, and all scored above the norm of 100 (M = 120, SEM = 3) on the composite Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999).

Equipment and stimuli. We presented stimuli on a linearized CRT monitor with a 120 Hz refresh rate while the right eye's gaze position was monitored by an EyeLink eye tracker (SR Research, Ottawa, Ontario, Canada). The stimuli consisted of a white background, a small dark fixation cross subtending 0.3° × 0.3° of visual angle, and dark letter strings in Courier font. The letter strings had 83% Weber contrast. The words were drawn from 12 distinct semantic categories (e.g., "professions"), each with 35 nouns (see the Supplemental Material available online). The median lexical frequency was 6.4 per million, according to the Clearpond database (Marian, Bartolotti, Chabal, & Shook, 2012). They ranged from four to six characters in length, subtending 2.7° to 4.1° in width and 1.1° in height. We also used six-character masks (#@#@#@ and @#@#@#). The words and masks were centered at 2.75° to the left and right of fixation.

All the words were dark gray except for the color targets, which were equiluminant but with higher saturation in the red hue. Using the measured luminance values of the monitor, we incremented the red gun and

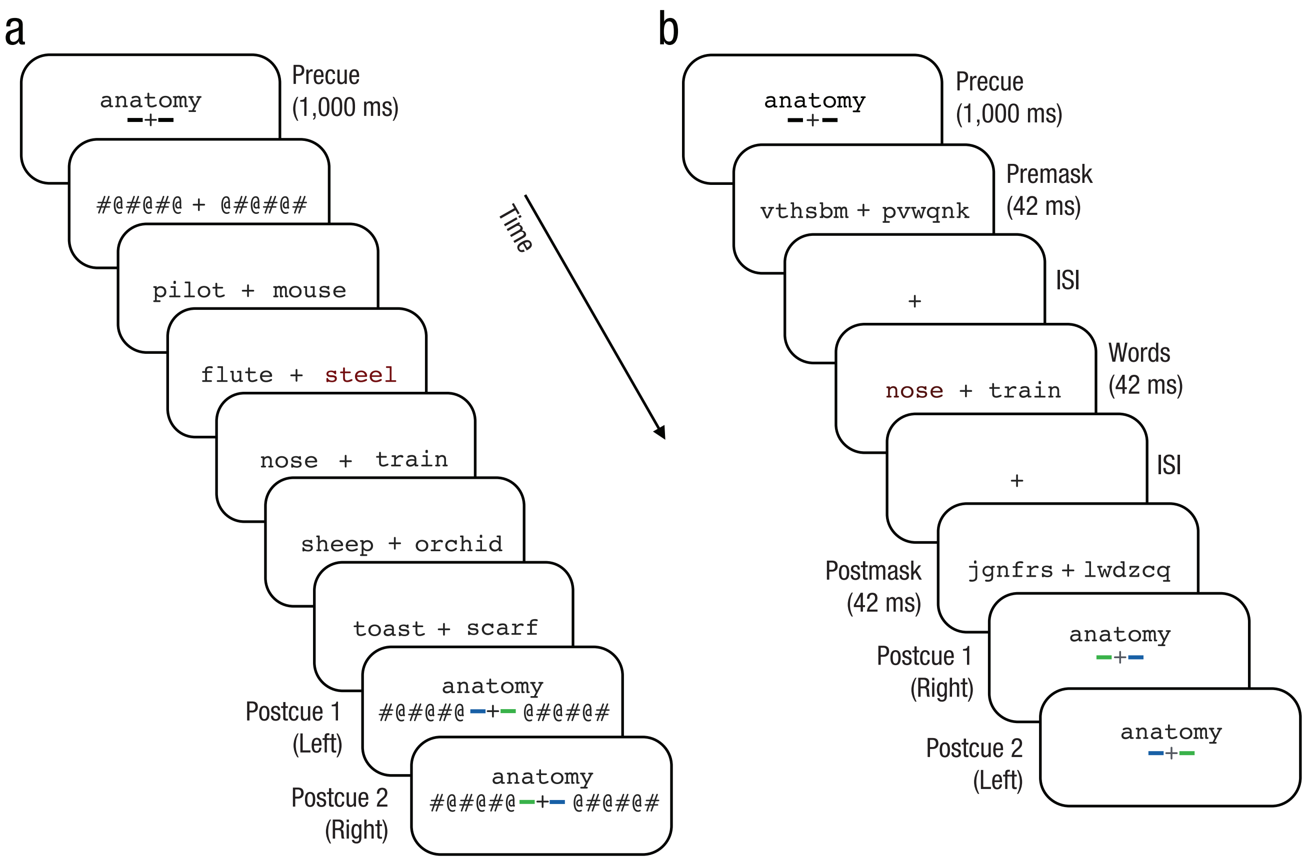

Fig. 1. Example dual-task semantic trials of Experiments 1 (a) and 2 (b). In Experiment 1, the rapid-serial-visual-presentation (RSVP) rate was adjusted for each participant (see the text). Not shown in panel (a) are the interstimulus intervals (ISIs) between each frame, during which only the fixation cross was presented. In Experiment 2, the duration of the two ISIs between masks and words was adjusted for each participant. The category name ("anatomy") was actually presented below fixation. In these examples, the postcue is the blue line that points to the side to be judged, but for half of the participants, it was the green line. Dual-task trials are shown here; in single-task trials, there was only one postcue, as the participant had to judge only one side.

decreased the green and blue guns to keep the total luminance constant. As described below, we set the magnitude of the red increment to each participant's detection threshold.

Trial sequence. The trial began with the participant fixating centrally for at least 1 s. Then the precue appeared: two 0.35° lines just to the left and right of the fixation cross. In the dual-task condition, both were black. In the single-task condition, one line was green and the other blue. Each participant was assigned to either green or blue and always attended to the side indicted by that color. A target-defining word (the name of the semantic category, or the word "color") also appeared 1° below fixation.

The 1-s precue was followed by a 600-ms interstimulus interval (ISI). Then the RSVP sequence began with premasks that covered the upcoming word locations, followed by a blank ISI containing only the fixation cross. Then five pairs of words were presented sequentially, separated by blank ISIs (the ISIs are not shown in Fig. 1). The premask, words, and ISIs all had the same duration, D_{RSVP}, which was adjusted to control semantic judgment difficulty (see below). After the last ISI, a postcue display appeared. This consisted of postmasks and a green and a blue line as in the precue. After 700 ms, a beep prompted the participant to respond to the side indicated by his or her assigned color. Responses before the beep were not recorded.

The postcue remained visible until the participant pressed one of four keys with the hand on the same side as the postcue, reporting with a 1 to 4 rating the level of confidence that a target was present: "sure absent," "guess absent," "guess present," or "sure present." Each key press was immediately followed by a high- or low-pitched feedback tone for correct or incorrect responses, respectively. On dual-task trials, as illustrated in Figure 1, a second postcue was then presented to prompt a response about the other side.

If the participant's gaze position moved too far from the fixation mark during the presentation of the stimuli, the trial was immediately terminated (see the Supplemental Material). This occurred on an average of 5.9% of trials (*SEM* = 1.0%).

Tasks. For semantic judgments, the targets were nouns of a particular semantic category. For color judgments, the targets were words colored slightly red, as defined above. On each trial, on each side, there was an independent 50% chance of semantic target presence and an independent 50% chance of color target presence. Therefore, for both target types in any given trial, there could be no targets, one target, or two targets (one on each side).

The time of each target type within the sequence was uniformly and independently distributed. When there were two targets of the same type, they appeared simultaneously. Only 10% of semantic targets happened to also be color targets. The particular words on each trial were chosen randomly from the entire set with the following constraints: (a) All words on each side were unique, (b) the same word was never presented on both sides simultaneously, (c) no more than one word from the target category could be present on one side, and (d) target words were not allowed to repeat across sequential trials.

In single-task trials, the precue instructed participants to attend to one side and ignore the other. The postcue always indicated the same side as the precue, requiring just one response. In dual-task trials, participants had to make independent judgments of target presence for both sides. The postcue first prompted the participant to respond to one side; then after the key press and feedback tone, the postcue switched, and the participant responded to the other side. The postcue order (left or right side first) was randomized across trials. Note again that in dual-task trials, both judgments were of the same type (semantic-semantic or color-color).

Procedure. In the first session, participants read the entire word list and practiced both judgment types. In the second session, they ran staircase blocks so we could estimate thresholds. A staircase for single-task semantic judgments established the threshold word duration, D_{RSVP}. Then, with the stimulus timing fixed, a staircase for single-task color judgments established the threshold color increment magnitude *I*. The average D_{RSVP} was 83 ms (*SEM* = 8.5 ms), corresponding to an average presentation rate of 6.7 Hz (*SEM* = 0.7 Hz). The average *I* (expressed as a proportion of the maximum possible equiluminant red saturation) was 0.09 (*SEM* = 0.01).

Main experimental trials were run in blocks of 16. The target type (semantic, color) was constant within each block, as was the semantic target's category and the precue condition (dual-task, single-task right, single-task left). Each participant completed 960 trials (60 blocks) of each judgment type.

Analysis. As a bias-free measure of accuracy, we computed the area under the receiver-operating-characteristic (ROC) curve, known as Green's area or A_g (Pollack & Hsieh, 1969). Like proportion correct, A_g is a proportion that ranges from .5 (guessing) to 1.0 (perfect). See the Supplemental Material for details.

Experiment 2

Participants. Ten volunteers participated (3 female, ages 19–31 years). Eight had also participated in Experiment 1. One additional participant chose to discontinue the study after one session. The mean TOWRE score was 119 (*SEM* = 2).

Trial sequence and procedure. The method of Experiment 2 (Fig. 1b) was identical to that of Experiment 1, except as follows. After the precue, a premask display appeared for 42 ms. It consisted of two strings of six randomly chosen consonants. Then only the fixation mark was presented for an ISI, the duration of which (D_{ISI}) was set to each subject's threshold for single-task semantic judgments. Then a pair of words appeared for 42 ms, followed by a second ISI of the same duration as the first. The following postmask was similar to the premask, 42 ms in duration, but with different random consonants. After the postmask, the trial sequence finished in the same way as in Experiment 1. Trials were run in blocks of 20.

There was an independent 50% chance that the word on each side was drawn from the target semantic category. Otherwise, it was drawn randomly from one of the other 11 categories. No word was allowed to appear on two consecutive trials or to appear on both sides in the same trial. There was also an independent 50% chance that each word was a color target.

The average D_{ISI} was 71 ms (*SEM* = 6.2 ms), and the average color target saturation increment *I* (proportion of the maximum possible) was .14 (*SEM* = .01). An average of 5.1% of trials were aborted because of breaks in fixation (*SEM* = 1.7%).

Results

Dual-task deficits

Semantic accuracy was significantly impaired in the dual-task conditions of both experiments (Table 1). The mean dual-task deficit (single-task A_g – dual-task A_g) was .12 (*SEM* = .01) in Experiment 1—comparison with 0: $t(9) = 17.17$, $p < .001$, 95% bootstrapped confidence interval (CI) = [.11, .13]. In Experiment 2, the mean semantic dual-task deficit was .14 (*SEM* = .01), $t(9) = 10.14$, $p < .001$, 95% CI = [.12, .17]. Color judgments, in contrast, suffered minimal deficits—Experiment 1: *M* = .02, *SEM* = .01, $t(9) = 1.55$, $p = .15$, 95% CI = [.0, .04];

Table 1. Mean Accuracy Level (A_g) in the Single- and Dual-Task Conditions

Measure	Experiment 1 Semantic judgments	Experiment 1 Color judgments	Experiment 2 Semantic judgments	Experiment 2 Color judgments
Single-task A_g	.83	.82	.84	.82
Dual-task A_g	.71	.80	.70	.78

Note: For all values, the standard error of the mean is .01.

Experiment 2: M = .04, SEM = .01, $t(9)$ = 4.63, p = .001, 95% CI = [.02, .06]. Dual-task deficits were significantly greater for semantic than color judgments, according to analysis of variance (ANOVA) interactions between judgment type and precue condition—Experiment 1: $F(1, 9)$ = 54.48, p < .001; Experiment 2: $F(1, 9)$ = 30.73, p < .001. The same analysis of d' (which, unlike A_g, is unbounded) supported the same conclusion—interactions: $F(1, 9)$ = 28.06, p < .001; $F(1, 9)$ = 40.21, p < .001.

Accuracy was consistently higher for targets on the right than left side of fixation, especially for semantic judgments, consistent with previous reports (e.g., Boles, 1983; Mishkin & Forgays, 1952). Details are available in the Supplemental Material.

Attention operating characteristics (AOCs)

The AOC allows the comparison of dual-task deficits to specific model predictions (Sperling & Melchner, 1978). Figure 2 plots accuracy for words on the left side of fixation against accuracy for words on the right. The single-task conditions are pinned to their respective axes. The accuracy levels in the dual-task condition form a single point (open circle) in that 2-D space. We compared that point to the predictions of three specific models of capacity limits (Bonnel & Prinzmetal, 1998; Scharff et al., 2011; Shaw, 1980; Sperling & Melchner, 1978), which are briefly described here (see the Supplemental Material for more detail).

The *unlimited-capacity parallel-processing model* assumes that two stimuli can be fully processed simultaneously just as well as one (i.e., there is no dual-task deficit). In the AOC plot, this model predicts that the dual-task data point falls at the intersection of the dashed lines.

The *fixed-capacity parallel-processing model* assumes that the perceptual system extracts a fixed amount of information from the whole display per unit time. Therefore, processing resources must be shared between both stimuli in the dual-task condition, which lowers sensitivity. As the proportion of resources given to the right stimulus increases from 0 to 1, this model traces out the black curve in the AOC plot. See the Supplemental Material for the calculation of this curve.

The *all-or-none serial-processing model* assumes that only one stimulus can be processed per trial, with equal sensitivity as in the single-task condition. The participant does not have time to even start processing the other stimulus and therefore must guess when asked about it. As the proportion v of trials in which the right side is processed increases from 0 to 1, this model traces out the diagonal black line in the AOC plot.

The serial-processing model can be generalized to account for less severe deficits by assuming that on some fraction b of dual-task trials, both sides are fully processed. The resulting accuracy is a mixture of trials in which only one stimulus is processed and trials in which both stimuli are processed. In conditions with dual-task deficits larger than predicted by either parallel-processing model, we solved for the serial-processing model parameters b and v that best fitted each participant's data.

Semantic judgments: Experiment 1. As can be seen in Figure 2a, dual-task accuracy fell below the fixed-capacity parallel-processing model's curve and near but above the all-or-none serial-processing model's line. The mean distance to the nearest point on the fixed-capacity parallel curve was .05 (SEM = .01), significantly greater than 0, $t(9)$ = 4.53, p = .001, 95% CI = [.03, .07]. The mean distance to the all-or-none serial line was .04 (SEM = .01), $t(9)$ = 4.87, p < .001, 95% CI = [.03, .06]. Fitting the more general serial-processing model, we found that the mean b parameter was .20 (SEM = .04, 95% CI = [.13, .28]), meaning that on an average of 80% of the trials, only one stimulus was processed. The mean v was .71 (SEM = .05, 95% CI = [.62, .82]), meaning that there was a significant bias to process the right stimulus when only one could be processed. This bias might reflect an attentional strategy to process the easier side, given that only one could be processed successfully.

Semantic judgments: Experiment 2. Mean dual-task accuracy fell on top of the all-or-none serial-processing model's prediction (Fig. 2c). The average distance from that line was .005 (SEM = .012), which was not significantly above 0, $t(9)$ = 0.41, p = .69, 95% CI = [−.02, .03]. The average distance to the closest point on the fixed-capacity parallel curve was .08 (SEM = .02), $t(9)$ = 4.52, p = .001, 95% CI = [.05, .12]. The mean value of b was .06 (SEM = .03, 95% CI = [.01, .14]), meaning that on 94% of trials, only one stimulus was processed. There was again a strong bias to processing the right stimulus (mean v = .79, SEM = .04, 95% CI = [.71, .86]).

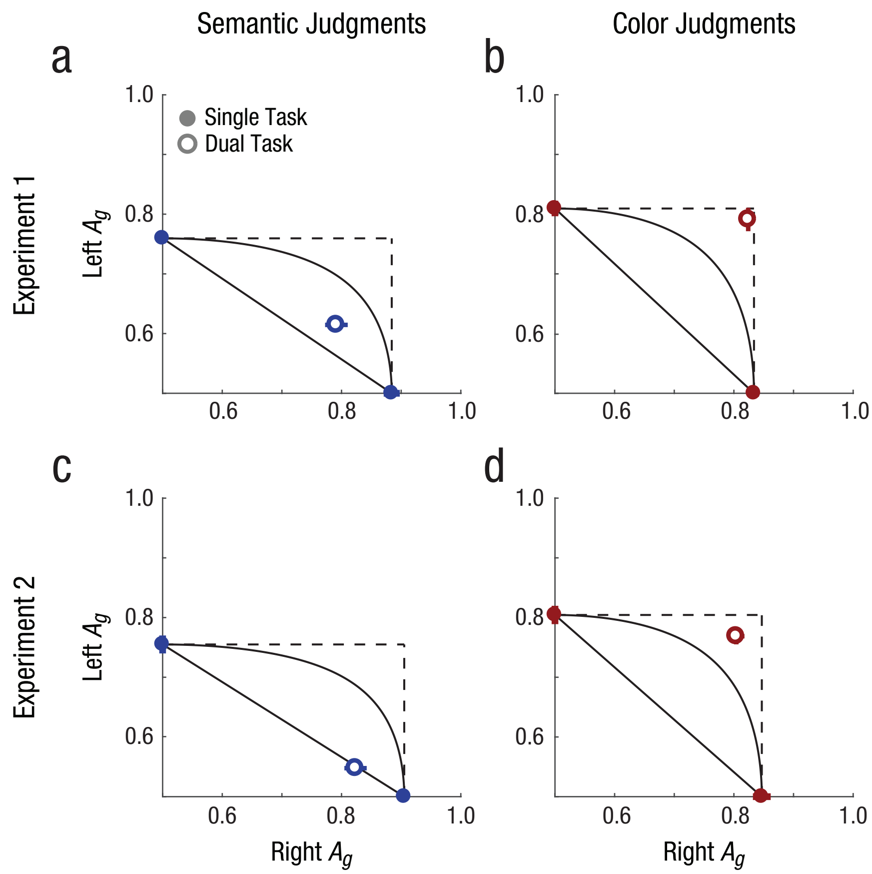

Fig. 2. Attention operating characteristics (AOCs) for semantic judgments (left column) and color judgments (right column) in Experiments 1 (top row) and 2 (bottom row). Solid points pinned to the axes are single-task accuracy levels (A_g); open points are dual-task accuracy levels. Error bars show ±1 *SEM*.

Color judgments: Experiment 1. The mean dual-task accuracy point fell well above the fixed-capacity parallel curve, near the intersection that marks the prediction of the unlimited-capacity parallel-processing model (Fig. 2b). The mean distance to the nearest point on the fixed-capacity curve was .09 (*SEM* = .01), which is significantly above 0, $t(9) = 7.07$, $p < .001$, 95% CI = [.06, .11].

Color judgments: Experiment 2. As in Experiment 1, dual-task accuracy fell near the unlimited-capacity parallel-processing model's prediction (Fig. 2d). The mean distance from the fixed-capacity parallel curve was .06 (*SEM* = .01), $t(9) = 6.25$, $p < .001$, 95% CI = [.04, .07].

Effects of accuracy on the other side

The all-or-none serial-processing model assumes that only one side can be processed per trial and no information is acquired about the other. If we also assume that the focus of attention switches across trials between the left and right sides, the model predicts a negative correlation between the accuracies of dual-task responses (Bonnel & Prinzmetal, 1998; Ernst, Palmer, & Boynton, 2012; Lee, Koch, & Braun, 1999; Sperling & Melchner, 1978). In other words, the participant is more likely to be correct about one side when he or she is incorrect about the other side.

There are several ways to test this prediction, including computing correlation coefficients between the accuracies of the left- and right-side responses. For semantic judgments, the correlation coefficients were negative, except when neither side contained a target (see the Supplemental Material). This complex pattern seemed to be related to changes in decision criterion or bias as a function of the other side's task. We therefore needed a bias-free measure of accuracy, such as area under the ROC curve (A_g). So, as another direct test of the serial-switching model's prediction (Braun & Julesz, 1998; Lee et al., 1999), we coded all dual-task responses by whether the response to the other side on the same trial was correct or incorrect and then computed A_g for both sets of trials (Fig. 3).

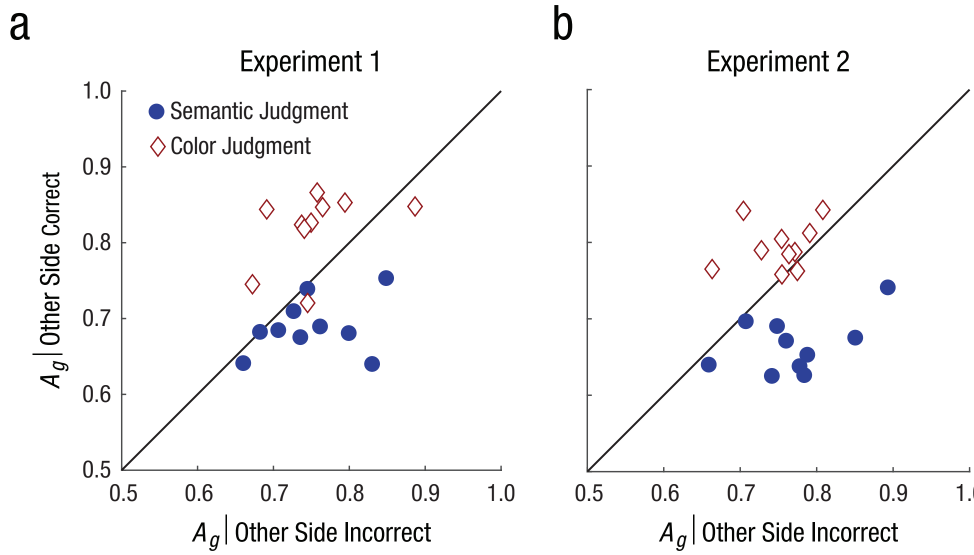

Fig. 3. Dual-task accuracy (A_g) when the response to the other side was incorrect plotted against accuracy when the response to the other side was correct, separately for (a) Experiment 1 and (b) Experiment 2. Each point is 1 participant's data for one judgment type. Points below the unity line indicate that accuracy was better when the other side's response was incorrect.

For semantic judgments in Experiment 1, accuracy was on average .06 A_g units (*SEM* = .02) lower when the other side's response was correct than incorrect, $t(9) = 3.31$, $p = .009$, 95% CI = [.03, .11]. In Experiment 2, this difference was .11 A_g units (*SEM* = .02), $t(9) = 5.87$, $p < .001$, 95% CI = [.07, .14]. This pattern supports the serial switching model.

For color judgments, the effect was exactly opposite: In Experiment 1, accuracy was on average .07 A_g units (*SEM* = .02) higher when the other side's response was correct than incorrect, $t(9) = 3.60$, $p = .006$, 95% CI = [.03, .10]. The same pattern occurred in Experiment 2: The mean difference was .04 A_g units (*SEM* = .02), $t(9) = 2.99$, $p = .015$, 95% CI = [.02, .08]. This positive effect of the other side's accuracy is consistent with fluctuations in overall effort or arousal that could cause positive correlations in sensitivity between the two sides (Bonnel & Prinzmetal, 1998).

In both experiments, the effect of the other side's accuracy significantly differed between semantic and color judgments—ANOVA interactions in Experiment 1: $F(1, 9) = 30.63$, $p < .001$; ANOVA interactions in Experiment 2: $F(1, 9) = 57.02$, $p < .001$.

Discussion

We measured dual-task deficits for color and semantic judgments of written words. For color judgments, there was minimal deficit, consistent with unlimited-capacity parallel processing. For semantic judgments, there was a large dual-task deficit that was inconsistent with unlimited-capacity or even fixed-capacity parallel processing. Instead, it supported an all-or-none serial-processing model: a bottleneck in the recognition process that allows only one word to be categorized per trial. The serial-processing model predicts that given limited processing time, participants process only the left stimulus on some trials and only the right on others. Fulfilling that prediction, semantic accuracy for each side was relatively impaired when the other side was judged correctly. (Note that our model fits suggested that participants could process both words on a minority of trials—20% and 6% in Experiments 1 and 2, respectively.) The opposite was true for color judgments. On the basis of these contrasting results, we argue that serial processing is necessary to judge the meaning of words but not their physical features.

The consistently large semantic dual-task deficit is unlikely to be explained by a memory limit, for several reasons. First, participants in Experiment 2 had to remember only two words for the brief interval between the stimuli and the postcue. Two words is within the limits of verbal working memory (Chen & Cowan, 2009). Second, dual-task color trials also required two reports with the same timing and suffered hardly any deficit. Finally, accuracy of the second responses on dual-task trials was not worse than the first responses, suggesting that both stimuli could be remembered for the whole trial. We therefore favor the hypothesis that the serial bottleneck lies in perceptual or linguistic analysis.

The color and semantic data fell at the opposite extremes despite the fact that the stimuli and single-task difficulty levels were the same. Therefore, the extent of parallel processing is not fixed by the stimulus but depends on the demands of the task. This finding also casts doubt on the view that semantic processing is automatic (Augustinova & Ferrand, 2014; Stroop, 1935). If both words were automatically processed in single-task semantic trials, accuracy would have been no better than in dual-task trials. Moreover, the lack of dual-task deficit for color judgments shows that participants can prevent semantic processing of the two attended words from overwhelming a central pool of limited resources. This supports the hypothesis that semantic processing requires attention and is under top-down control (e.g., Robidoux & Besner, 2015). Consistent with this view, a recent study demonstrated that when pairs of Chinese words are superimposed in different colors, readers are able to attend to one word at a time and avoid interference from the other (Liu & Reichle, 2017).

Although the color data were consistent with parallel processing, our design was not optimized to rule out a rapid switching of attention that sequentially processed both colors because the stimulus timing was set to control semantic difficulty only. The more predictable color target features could also account for the difference in results. Nonetheless, we showed that different tasks of equal difficulty performed on the same stimuli can have very different capacity limits.

It is likely that the large semantic dual-task deficit depended on our use of displays that limited the amount of time available to process each word pair. By the time one word was recognized, information about the other was lost or replaced by subsequent stimuli. In contrast, another study from our group using a semantic search task without postmasks found results consistent with the fixed-capacity parallel-processing model (Scharff et al., 2011). The lack of masking may have allowed multiple words to be processed sequentially. We propose that recognition is always serial, but multiple words can be processed in a single presentation if the stimuli or their memory traces last long enough.

Prior studies have demonstrated semantic capacity limits with evidence that response times increase with set size during search for particular words (Harris et al., 2004; Reichle et al., 2008). Reichle et al. (2008) further argued that visual features are detected in parallel but lexical processing is serial. However, search slopes alone are insufficient to distinguish between serial- and parallel-processing models (Carrasco & Yeshurun, 1998; Townsend, 1990).

The *redundant-target* paradigm has also been used to investigate capacity limits in word recognition. Assuming that search is self-terminating, unlimited-capacity parallel-processing models predict faster responses when the display consists of two targets compared with a single target. The first such study found no redundancy gains and concluded that capacity is limited, potentially (but not necessarily) because of a serial process (Mullin & Egeth, 1989). However, later studies reported contradictory results (Shepherdson & Miller, 2014, 2016).

More generally, the signatures of serial processing have rarely been observed in dual visual tasks. The studies that have observed them involved task conditions more complex than ours (Bonnel & Prinzmetal, 1998; Braun & Julesz, 1998; Lee et al., 1999; Sperling & Melchner, 1978). In all of these cases, the attentional set (i.e., which stimulus features the subject looks for and discriminates) was different for the two concurrent tasks. Our study is therefore unique: The left- and right-side tasks were identical.

Reading researchers have long debated whether words in a line of text are necessarily processed one at a time (Murray et al., 2013; Reichle et al., 2009). Much of the debate has focused on eye movement studies that demonstrate that readers begin processing word $n + 1$ before the eyes leave word n. There is even evidence that the *meaning* of word $n + 1$ is acquired before it is fixated (Hohenstein & Kliegl, 2014; Schotter, 2013), which is consistent with the assumptions of parallel-processing models. However, proponents of serial-processing models argue that they can also account for such semantic effects (Schotter, Reichle, & Rayner, 2014). Our study avoided that impasse and demonstrated a boundary condition in which parallel recognition of two words is not possible.

The conditions of our study differed from those of natural reading, so we cannot say with certainty how our findings apply to the processing of a whole page of text. By demonstrating one condition in which semantic processing of two words is strictly serial, we do not rule out the hypothesis that parallel processing is at least sometimes possible during natural reading. For instance, parallel processing might be possible when pairs of words are related to each other or when one word is fixated and the other is to the right. Reading is a remarkable skill, and future research must explore a wider range of tasks and stimulus configurations to map out the cognitive functions that support it.

Action Editor

Rebecca Treiman served as action editor for this article.

Author Contributions

A. L. White, J. Palmer, and G. M. Boynton designed the study; A. L. White collected and analyzed the data; A. L. White wrote the manuscript; and A. L. White, J. Palmer, and G. M. Boynton

edited it. All authors approved the final version of the manuscript for submission.

ORCID iD

Alex L White https://orcid.org/0000-0001-5530-742X

Acknowledgments

Thanks to Kadie Anne Brown, Micah Ketola, and Evan McAuley for help with data collection.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This work was supported by National Eye Institute Grants T32 EY007031 to the University of Washington, R01 EY12925 to G. M. Boynton and J. Palmer, and F32 EY026785 to A. L. White.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617751898

Open Practices

The data and analysis code have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/ewr45/ and https://osf.io/ev3w5/, respectively. The design and analysis plans for the experiments were not preregistered. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797617751898. This article has received the badge for Open Data. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

References

Augustinova, M., & Ferrand, L. (2014). Automaticity of word reading: Evidence from the semantic Stroop paradigm. *Current Directions in Psychological Science*, *23*, 343–348.

Boles, D. B. (1983). Hemispheric interaction in visual field asymmetry. *Cortex*, *19*, 99–113.

Bonnel, A.-M., & Prinzmetal, W. (1998). Dividing attention between the color and the shape of objects. *Perception & Psychophysics*, *60*, 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, 44A*, 601–626.

Braun, J. (1998). Divided attention: Narrowing the gap between brain and behavior. In R. Parasuraman (Ed.), *The attentive brain* (pp. 327–351). Cambridge, MA: MIT Press.

Braun, J., & Julesz, B. (1998). Withdrawing attention at little or no cost: Detection and discrimination tasks. *Perception & Psychophysics*, *60*, 1–23.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*, 1484–1525.

Carrasco, M., & Yeshurun, Y. (1998). The contribution of covert attention to the set-size and eccentricity effects in visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 673–692.

Chen, Z., & Cowan, N. (2009). Core verbal working-memory capacity: The limit in words retained without covert articulation. *The Quarterly Journal of Experimental Psychology*, *62*, 1420–1429.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777–813.

Ernst, Z. R., Palmer, J., & Boynton, G. M. (2012). Dividing attention between two transparent motion surfaces results in a failure of selective attention. *Journal of Vision*, *12*(12), Article 6. doi:10.1167/12.12.6

Grainger, J., Dufau, S., & Ziegler, J. C. (2016). A vision of reading. *Trends in Cognitive Sciences*, *20*, 171–179.

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

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 166–190.

Kennedy, A. (2000). Parafoveal processing in word recognition. *The Quarterly Journal of Experimental Psychology*, *53*, 429–455.

Lee, D. K., Koch, C., & Braun, J. (1999). Attentional capacity is undifferentiated: Concurrent discrimination of form, color, and motion. *Perception & Psychophysics*, *61*, 1241–1255.

Liu, Y., & Reichle, E. D. (2017). Eye-movement evidence for object-based attention in Chinese reading. *Psychological Science*, *29*, 278–287.

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). Clearpond: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLOS One*, *7*(8), Article e43230. doi:10.1371/journal.pone.0043230

Mishkin, M., & Forgays, D. G. (1952). Word recognition as a function of retinal locus. *Journal of Experimental Psychology*, *43*, 43–48.

Mullin, P. A., & Egeth, H. E. (1989). Capacity limitations in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 111–123.

Murray, W. S., Fischer, M. H., & Tatler, B. W. (2013). Serial and parallel processes in eye movement control: Current controversies and future directions. *The Quarterly Journal of Experimental Psychology*, *66*, 417–428.

Pastukhov, A., Fischer, L., & Braun, J. (2009). Visual attention is a single, integrated resource. *Vision Research*, *49*, 1166–1173.

Pollack, I., & Hsieh, R. (1969). Sampling variability of the area under the ROC-curve and of d′e. *Psychological Bulletin*, *71*, 161–173.

Potter, M. C. (1984). Rapid serial visual presentation (RSVP): A method for studying language processing. In D. Kieras

& M. Just (Eds.), *New methods in reading comprehension research* (pp. 91–118). Hillsdale, NJ: Erlbaum.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.

Reichle, E. D., Liversedge, S. P., Pollatsek, A., & Rayner, K. (2009). Encoding multiple words simultaneously in reading is implausible. *Trends in Cognitive Sciences*, *13*, 115–119.

Reichle, E. D., Vanyukov, P. M., Laurent, P. A., & Warren, T. (2008). Serial or parallel? Using depth-of-processing to examine attention allocation during reading. *Vision Research*, *48*, 1831–1836.

Robidoux, S., & Besner, D. (2015). Conflict resolved: On the role of spatial attention in reading and color naming tasks. *Psychonomic Bulletin & Review*, *22*, 1709–1716. doi:10.3758/s13423-015-0830-7

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

Schotter, E. R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language*, *69*, 619–633.

Schotter, E. R., Reichle, E. D., & Rayner, K. (2014). Rethinking parafoveal processing in reading: Serial-attention models can explain semantic preview benefit and *N*+2 preview effects. *Visual Cognition*, *22*, 309–333.

Shaw, M. L. (1980). Identifying attentional and decision-making components in information processing. In R. Nickerson (Ed.), *Attention and performance VIII* (pp. 277–296). New York, NY: Routledge.

Shepherdson, P. V., & Miller, J. (2014). Redundancy gain in semantic categorisation. *Acta Psychologica*, *148*, 96–106.

Shepherdson, P. V., & Miller, J. (2016). Non-semantic contributions to "semantic" redundancy gain. *The Quarterly Journal of Experimental Psychology*, *69*, 1564–1582.

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

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: Some current controversies. *Trends in Cognitive Sciences*, *5*, 156–163.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Torgesen, J., Wagner, R., & Rashotte, C. (1999). *Test of Word Reading Efficiency: Manual*. Austin, TX: Pro-Ed.

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*, 46–54.

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), Article 19. doi:10.1167/17.6.19.