Effects of selective and divided attention arise from either serial or parallel processing depending on the task

Miranda Johnson

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2021

Reading Committee:
John Palmer, Chair
Geoffrey M. Boynton
Cathleen M. Moore

Program Authorized to Offer:
Psychology

University of Washington

Abstract

Effects of selective and divided attention arise from either serial or parallel processing depending on the task

Miranda Johnson

Chair of the Supervisory Committee:
John Palmer
Department of Psychology

Visual attention involves selecting relevant information for further processing. In selective attention, a single stimulus is selected while distractors are filtered out. In divided attention, multiple stimuli are selected. Both types of visual attention have been described by two competing theories: selective perception or selective decision. Selective perception can be described as a process in which attention enhances perceptual encoding. Selective decision can be described as a process in which attention influences decision making. Within these two broad theories, there are specific models. This dissertation focuses on distinguishing serial models of attention, in which only a single stimulus is processed at a time, and parallel models of attention, in which multiple stimuli are processed at the same time. The first two experiments of this dissertation asked whether selective attention is accounted for by a serial or a parallel process. In the first experiment, selective attention to Gabor patches was measured using the partially valid cueing paradigm where a cue indicated the location where a stimulus was most likely to appear. Results of this experiment indicate that selective attention for simple stimuli, such as Gabor patches, is consistent with selective decision and a model that assumes parallel processing for perception. In the second experiment, selective attention to masked words was measured using partially valid cueing. For this more complex stimulus, selective attention was consistent with an all-or-none serial model for perception. The final experiment asked whether divided attention to moving stimuli is accounted for by the all-or-none serial model or fixed-capacity parallel model . Divided attention was measured using the multiple object tracking paradigm. Results of this experiment were not consistent with either model, but elaborations of either model can account for the results. Together, the results of this dissertation indicate that the mechanisms underlying selective and divided attention depend on the stimulus being attended.

TABLE OF CONTENTS

CHAPTER ONE

Dissertation Introduction

Visual attention is instrumental in guiding our perception of the world. Attended information can be accessed, while unattended information is filtered out. This process of attending to information can be further divided into *selective attention* and *divided attention*. Selective attention is the selection of a relevant stimulus for processing while not selecting irrelevant distractors. Divided attention is the selection of multiple stimuli for processing rather than a single stimulus. This dissertation focuses on two paradigms that have been developed to measure effects in both selective and divided attention. Specifically, I test whether selective and divided attention to different types of stimuli are consistent with serial or parallel processes. For serial processes, only a single stimulus is selected at a time. An example of a serial process is the serial switching model, where attentional selection is switched between stimuli over time. For parallel processes, multiple stimuli are processed at the same time in parallel. An example of a parallel process is the weighted parallel model, where two stimuli are processed at once, but one of the stimuli is given priority over the other.

Overview of Selective Attention and Partially Valid Cueing

One way that selective attention is measured is by the partially valid cueing paradigm (Posner, Snyder, & Davidson, 1980). In partially valid cueing, participants are tasked with making judgements about a stimulus, and are cued to the spatial location where the stimulus is most likely to occur. Performance is better when the stimulus appears at a cued location versus when it appears at an uncued location. There are two competing hypotheses for the locus in processing that causes this cueing effect. The first hypothesis, *selective perception*, posits that perceptual encoding is enhanced for cued information. Under this hypothesis, selective processing is often described as a limited resource (Norman & Bobrow, 1975), or it is described using a spotlight metaphor (Posner et al., 1980). The second hypothesis, *selective decision*, posits that perceptual encoding is the same at the cued and uncued locations and the relevant perceptual processing is unlimited in capacity. Instead, the cue is used as a prior in decision making. As in Bayesian decision making, such a prior can influence performance.

Two experiments in this dissertation seek to distinguish these hypotheses. The first experiment uses simple stimuli and distinguishes the two hypotheses by comparing cuing effects for simultaneous and sequential displays. The simultaneous-sequential paradigm manipulates the amount of perceptual information encoded during a given time interval, while holding the decision component of the task constant (Eriksen & Spencer, 1969; Shiffrin & Gardner, 1972). The second experiment uses a more complex stimulus—words—and distinguishes these two hypotheses by comparing performance in the invalid condition to the predictions of the two models.

Endogenous cueing effects for detection

The first of my two cueing experiments, which has been completed and published, used a detection-like task where participants report the orientation of a low-contrast Gabor (Johnson, Palmer, Moore & Boynton, 2020). In the simultaneous condition, a spatial cue indicated with 80% validity which of two spatial locations the target was most likely to occur, and the temporal interval during which the stimulus would occur was known. In the sequential condition, a temporal cue indicated with 80% validity which of two temporal intervals the target was most likely to occur, and the spatial location where the stimulus would occur was known. The selective perception and selective decision hypotheses can be distinguished by comparing performance in the sequential condition.

Selective perception posits that perceptual encoding involves a limited resource that can be selectively allocated in space and switched to different spatial locations over time. For brief stimulus displays like those used in our simultaneous condition, there is not enough time to switch attention, therefore participants have only one opportunity to allocate resources. Because the cued location is much more likely to contain the target, it is best to allocate resources to that location, leading to a cueing effect in the simultaneous condition. In contrast, in the sequential condition, the spatial location of the target is known with 100% validity, and the temporal interval is cued. We used long delays between stimulus intervals to allow plenty of time for attentional switching. Because of this, participants can allocate resources to the location where the target will appear and encode the target regardless of whether it appears in the cued or uncued interval. Thus, selective perception predicts a cueing effect for the simultaneous condition, but not the sequential condition.

Selective decision posits that perceptual encoding is unlimited in capacity, meaning that the quality of encoding is not enhanced by the cue. Instead, cueing effects occur because the cue is used as a prior to improve decision making. In both the simultaneous and sequential conditions, cued information is weighted more heavily than uncued information simply because the target is more likely to occur at the cued location. Therefore, selective decision predicts a cueing effect for both the simultaneous and sequential conditions.

Endogenous cueing effects for word recognition

The second of the cueing experiments used a word categorization task to distinguish serial models and parallel models of selective attention. An 80% valid cue indicated which of two spatial locations the target was most likely to occur. Participants reported which of two categories a target word belonged to.

Prior work has found that participants are more accurate and faster to respond when the target word appears at a cued location than when it appears at an uncued location (McCann, Folk & Johnston, 1992). However, it is not yet known whether cueing effects for tasks involving complex stimuli, such as words, are best accounted for by selective perception or selective decision. We distinguish two specific versions of each hypothesis: the all-or-none serial model and the weighted parallel model. To distinguish these models, I use masks to limit attentional switching. There has been evidence from divided attention studies that word processing is serial (White, Palmer & Boynton, 2018; White, Boynton & Yeatman, 2019), meaning that attentional resources can only be allocated to one word at a time in an all-or-none fashion.

Overview of Divided Attention and Multiple Object Tracking

In divided attention, information from multiple stimuli is selected for further processing. For some tasks, visual information can be selected from multiple objects in parallel with unlimited capacity, while for other tasks there is a limit to how much information can be selected at once. For example, it has been found that participants can divide attention across simple features (e.g. the contrast of discs in noise; Scharff, Palmer, & Moore, 2011), but not more complex stimuli (e.g. discrimination of masked words; Whites, Palmer, & Boynton, 2018; 2020). This limit can be all-or-none in nature, such that only one object can be selected at once, or it can be parallel in nature, such that multiple objects can be selected, but each object gets less of the attentional resource than a single object would. The third experiment in my dissertation focuses on limits in dividing attention when observers track moving objects.

The ability to attend to multiple moving objects is often measured using the multiple object tracking (MOT) paradigm, where participants fixate on a central point in the display while tracking a subset of moving targets among distractors. Typically, performance decreases with the number of objects tracked (Pylyshyn & Storm, 1988).

There are several general theories for why it is difficult to track multiple moving objects, some of which fall within the general categories of serial switching theory and parallel resource theory. We focus on specific models from each of the two broader theories. (1)The *all-or-none serial model,* which is an extreme version of serial switching, posits that attention is a limited resource that can be allocated to only one object at a time, and we must guess on the location of additional objects (Holcombe & Chen, 2013). (2) The *fixed-capacity parallel model*, which is a specific version of parallel resource theory, posits that attention is a limited, continuous resource that is shared across objects in parallel (Alvarez & Franconeri, 2007; Horowitz & Cohen, 2010). The more targets tracked, the less of that resource each individual target gets. To test these specific models, we must limit perceptual crowding using sparse displays. Perceptual crowding occurs when objects close in visual space are confused. Another general theory, *crowding theory* posits that we can attend to an unlimited number of objects, but tracking is hindered by perceptual crowding (Franconeri, Jonathan, & Scimeca, 2010). Under crowding theory, we can theoretically track an unlimited number of objects in the absence of crowding. By using sparse displays, we can measure divided attention effects that are not likely due to perceptual crowding.

Dual-task deficits in MOT

The two hypotheses described above were distinguished using a dual-task design. On each trial, participants were presented with a display containing three objects in the upper visual field and three in the lower visual field. The discs were widely-spaced to prevent perceptual crowding. In the single-task condition, participants tracked one target object that appeared in the upper or lower visual field. In the dual-task condition, participants simultaneously tracked two target objects, one in the upper visual field and one in the lower visual field.

I used an analysis of the attentional operating characteristic (AOC; Sperling & Melchner, 1978) to make distinct predictions for these three hypotheses. Each of the three hypotheses predicts a different-magnitude deficit for tracking one vs. two targets. The AOC analysis has previously been used to identify an all-or-none serial limitation for recognition of stationary word stimuli (White, Palmer, & Boynton, 2018).

The all-or-none serial model predicts a large divided attention effect that is consistent with participants being able to track only one target, and having to guess on a second target. Additionally, if only one target is tracked, we expect a negative correlation in the accuracy between the top and bottom targets because a correct response for the top side is associated with an incorrect response for the bottom side, and vice versa. Resource theory also predicts a divided attention effect, but one of a smaller magnitude. Importantly, while the all-or-none serial model predicts a negative correlation in the accuracy between the top and bottom targets, the fixed-capacity parallel model predicts no correlation in dual-task accuracy.

References

Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track ?: Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, *7*(4), 1–10. https://doi.org/10.1167/7.13.14.Introduction

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

Franconeri, S. L., Jonathan, S. V., & Scimeca, J. M. (2010). Tracking multiple objects is limited only by object spacing, not by speed, time, or capacity. *Psychological Science*, *21*(7), 920–925. https://doi.org/10.1177/0956797610373935

Horowitz, T. S., & Cohen, M. A. (2010). Direction information in multiple object tracking. *Attention, Perception, & Psychophysics*, *72*(7), 1765–1775. https://doi.org/10.3758/APP

Howe, P. D. L., & Holcombe, A. O. (2012). Motion information is sometimes used as an aid to the visual tracking of objects. *Journal of Vision*, *12*(13), 10–10. https://doi.org/10.1167/12.13.10

Johnson, M. L., Palmer, J., Moore, C. M., & Boynton, G. M. (2020). Endogenous cueing effects for detection can be accounted for by a decision model of selective attention. *Psychonomic Bulletin and Review*, *27*(315–321).

McCann, R. S., Folk, C. L., & Johnston, J. C. (1992). The role of spatial attention in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1015–1029.

Norman, D. A., & Bobrow, D. G. (1975). On Data-limited and Resource-limited Process. *Cognitive Psychology*, *7*, 44–64.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*(2), 160–174. https://doi.org/10.1037/0096-3445.109.2.160

Pylyshyn, Z. W., & Storm, R. O. N. W. (1988). Tracking multiple independent targets : Evidence for a parallel tracking mechanism *. *Spatial Vision*, *3*(3), 179–197.

Scharff, A., Palmer, J., & Moore, C. M. (2013). Divided attention limits perception of 3-D object shapes. *Journal of Vision*, *13*(2), 1–24. https://doi.org/10.1167/13.2.18.doi

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

Sperling, G., & Melchner, M. J. (1978). The Attention Operating Characteristic : Examples from Visual Search. *Science*, *202*(4365), 315–318.

White, A. L., Palmer, J., & Boynton, G. M. (2018). Evidence of Serial Processing in Visual Word Recognition. *Psychological Science*, *29*(7), 1062–1071. https://doi.org/10.1177/0956797617751898

White, A. L., Palmer, J., & Boynton, G. M. (2020). Visual word recognition : Evidence for a serial bottleneck in lexical access. *Attention, Perception, and Psychophysics*, *82*, 2000–2017.

CHAPTER TWO:

Endogenous cueing effects for detection can be accounted for by a decision model of selective attention

Miranda L. Johnson[1]

John Palmer[1]

Cathleen M. Moore[2]

Geoffrey M. Boynton[1]

[1]University of Washington, Department of Psychology

[2]University of Iowa, Department of Psychology

Funding
This research was partly supported by grants to G. M. Boynton from the National Institute of Health EY12925.

Abstract

Spatial cues help participants detect a visual target when it appears at the cued location. One hypothesis for this cueing effect, called *selective perception*, is that cueing a location enhances perceptual encoding at that location. Another hypothesis, called *selective decision*, is that the cue has no effect on perception, but instead provides prior information that facilitates decision-making. We distinguished these hypotheses by comparing a simultaneous display with two spatial locations to sequential displays with two temporal intervals. The simultaneous condition had a partially-valid *spatial* cue, and the sequential condition had a partially-valid *temporal* cue. Selective perception predicts no cueing effect for sequential displays given there is enough time to switch attention. In contrast, selective decision predicts cueing effects for sequential displays regardless of time. We used endogenous cueing of a detection-like coarse orientation discrimination task with clear displays (no external noise or post-masks). Results showed cueing effects for the sequential condition, supporting a decision account of selective attention for endogenous cueing of detection-like tasks.

Whether reading a book or having a conversation in a crowded room, selective attention allows one to select relevant information while limiting distractions. One way to measure selective attention is *partially-valid cueing* (Posner, Snyder & Davidson, 1980), in which a cue indicates where a target stimulus is most likely to appear. For some proportion of trials, the cue is valid and the target appears at the cued location. For a smaller proportion of trials, the cue is invalid and the target appears at an uncued location. Targets are more likely to be detected at the cued location.

What causes such *spatial cueing effects*? One hypothesis, *selective perception,* posits that perceptual encoding is enhanced for information at the cued location. Under this hypothesis, the selective processing is often described using a spotlight metaphor (Posner, et al., 1980), or a limited resource (Norman & Bobrow, 1975). It is assumed that, given time, one can switch selective processing to different spatial locations. This idea has been generalized to allow allocation of renewable resources over space and time (Denison, Carrasco, & Heeger, 2017).

A second hypothesis, *selective decision*, posits that perceptual encoding has unlimited capacity, and therefore encoding is the same at cued and uncued locations. Rather than enhancing perceptual encoding, the cue is used in decision-making to increasingly weight the cued information. This hypothesis is a generalization of Bayesian decision-making in that the cue provides a prior that influences how information for the different locations is used in decision. An early example of this hypothesis is the weighting model of Kinchla, Chen and Evert (1995). For an optimal Bayesian model and a detailed development of this hypothesis, see Shimozaki, Eckstein and Abbey (2003). For this hypothesis, attentional switching is irrelevant because perception has unlimited capacity and thus cannot be improved by selective processing of the stimulus.

In the literature, some studies support selective perception (Dosher & Lu, 2000; Posner et al., 1980), and others support selective decision (Kinchla, et al., 1995; Shimozaki, et al., 2003). One recent hypothesis is that selective decision mediates endogenous cueing with clear displays and selective perception contributes to both exogenous and endogenous cuing of displays with external noise or post-masks (Dosher & Lu, 2000; Smith, 2000). In this article, we focus on endogenous cueing and clear displays, which appear most likely to be mediated by selective decision.

Our aim is to distinguish these hypotheses by comparing simultaneous and sequential displays (Eriksen & Spencer, 1969; Shiffrin & Gardner, 1972). This paradigm manipulates the amount of perceptual information that must be encoded within a given time interval, while keeping the decision component of the task constant. In their groundbreaking study, Shiffrin and Gardner (1972) tested whether the detection of letters is limited or unlimited in capacity. Participants detected a target letter among distractor letters that were shown either simultaneously or sequentially. A limited-capacity model predicts better performance for sequential compared to simultaneous displays. An unlimited-capacity model predicts equal performance for simultaneous and sequential displays. They found equal performance in the simultaneous and sequential conditions, which is consistent with letters being processed with unlimited capacity.

The simultaneous-sequential paradigm has since been used to test capacity limits for a variety of stimuli. The results support unlimited-capacity processing for simple stimuli, such as simple features, alphanumeric digits, and simple surface completion, but limited-capacity processing for more complex stimuli, such as words and objects (Attarha, Moore, Scharff, &

Palmer, 2014; Duncan, 1980; Huang & Pashler, 2005; Pashler & Badgio, 1987; Scharff, Palmer, & Moore, 2011, 2013).

Key to distinguishing limited and unlimited capacity is the assumption that participants can shift attention with sequential displays. To pursue this, Duncan, Ward, and Shapiro (1994) measured *attentional dwell time*: the time needed to completely shift attention from one object to another. The identification of the stimulus in the second display depended on the SOA which indicates that attending to the first stimulus interfered with identificiation of the second stimulus. Across several studies, estimates of the mean dwell times vary from 150 to 600 ms (Moore, Egeth, Berglan, & Luck, 1996; Petersen, Kyllingsbæk, & Bundesen, 2012). Thus, an SOA of about 1000 ms is sufficient to switch attention on almost all trials.

In the current study, we used the simultaneous-sequential paradigm to distinguish the predictions of selective perception and selective decision. Comparisons of spatial and temporal cueing have not often been made (Rohenkohl, Gould, Pessoa, & Nobre, 2014). We focused on conditions with clear displays (no external noise or post-masks), and endogenous cues. Stimuli were presented in one of two temporal intervals. In the simultaneous condition, a partially-valid cue indicated the most likely spatial location of a stimulus. In the sequential condition, a partially-valid cue indicated the most likely temporal interval. Both hypotheses predict a cueing effect in the simultaneous condition. Selective decision also predicts a cueing effect in the sequential condition because the cue is used as a prior for decision-making. In contrast, selective perception predicts no cueing effect in the sequential condition given there is time to switch attention from one interval to the other.

Experiment

Method

Overview. A single brief Gabor patch was presented that was either tilted left or right, and the task was to indicate whether the tilt was to the left or the right. Such a coarse orientation discrimination yields the same performance as detection (Thomas & Gille, 1979). Hence, we consider it a detection-like task. This stimulus was presented to either the left or right of fixation and in one of two possible intervals (Figure 2). In the simultaneous condition, the stimulus was presented at one of two locations in a known interval. In the sequential condition, the stimulus was presented in one of two intervals at a known location.

Participants. There were thirteen paid participants included one author (MJ). All participants had normal or corrected-to-normal acuity. All gave written and informed consent in accord with the human subjects Institutional Review Board at the University of Washington, in adherence with the Declaration of Helsinki.

To determine the number of participants, we used pilot data from a previous partially-valid cueing experiment. Participants (N = 12) each completed a coarse orientation discrimination experiment with similar methods as the simultaneous condition of the present experiment. We observed a cueing effect of 7 ± 1%. Seeking this size of cueing effect, a power analysis suggested a minimum of seven participants. To be conservative, we decided to use a minimum of 12. Due to accidents of scheduling, we tested a total of 13.

Apparatus. Displays were presented on a linearized CRT monitor (Sony GDM-FW900) with resolution 1024 by 640 pixels refreshing at 120 Hz. The monitor was viewed from 60 cm and had a mean luminance of 56 cd/m^2. Stimuli were created with MATLAB (MathWorks) and Psychophysics Toolbox (Brainard, 1997). Gaze position was monitored for all trials using an

EyeLink 1000 (SR Research). Trials containing blinks or broken fixation were excluded from analysis. Such excluded trials were infrequent and occurred on only 3 ± 0.5% of trials across participants.

Stimuli. Participants judged the orientation of a Gabor patch that had a spatial frequency for the grating component of 4 cycles per degree, a standard deviation of the Gaussian envelope of 0.5 degrees and was truncated to be 3 degrees in diameter. The Gabor was in one of two orientations: 130 degrees (left-tilting) or 40 degrees (right-tilting). It was presented 10 degrees to the left or right of fixation. The contrast of the Gabor was adjusted for each participant such that their performance with a valid cue was between 70-80% correct for both simultaneous and sequential trials. The mean contrast used was 8%.

Procedure. The simultaneous and sequential conditions are shown schematically in Figure 1. In the simultaneous condition, the cue indicated both the most likely target location and, with certainty, the temporal interval. For *early-interval blocks*, the trials began with fixation for 1500 ms, followed by the cue for 550 ms. The cue consisted of one red and one blue square, each 0.75 degrees in width and height. Both squares were one degree above fixation, and one degree to the left and right of fixation. Participants were assigned a cue color of red or blue. They were told that the cue indicated which side of fixation the target was most likely to appear, and the position of the cue above fixation indicates an early-interval trial. The random assignment of color was to ensure the cue was fully endogenous. The probability of the target appearing at the cued location was .8, and for the uncued location was .2.

Following the cue was a 500 ms delay with only the fixation cross. To reduce spatial uncertainty, there was a second 1050 ms delay with fiducial markers that indicated the locations

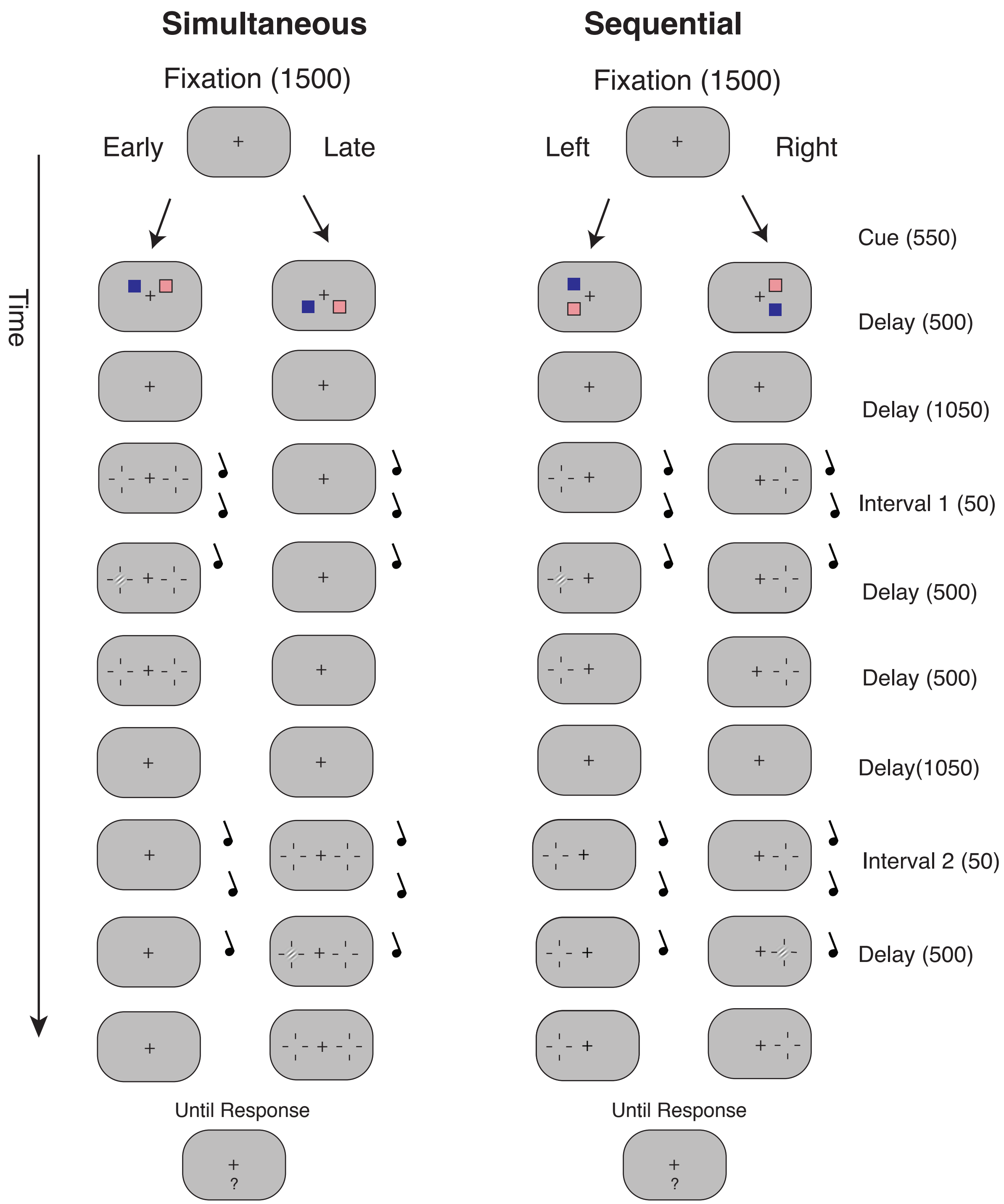

Figure 1. Schematic of the trial sequence for the simultaneous and sequential conditions (not to scale). Display durations are shown in milliseconds. In the simultaneous condition, the cue was spatial, indicating which side of fixation the target was most likely to appear. In the sequential condition, the cue was temporal, indicating which stimulus interval was most likely to contain the target. In the example sequence, the cue color was blue and valid trials are shown with a right-leaning Gabor target. The red cue is shown here with an outline to make it more distinct for reprinting.

where the target could appear. Additionally, to reduce temporal uncertainty, there was a sequence of three tones. During the time before the stimulus, two 500 Hz tones were played for 250 ms each with a 250 ms break following each tone.

After the delay, the stimulus was presented for 50 ms and a single 750 Hz tone was played for 250 ms. Following the stimulus was a 500 ms delay with fiducial markers. This continued display of the fiducial markers was to avoid an offset of the marker near in time to the presentation of the Gabor. An additional 500 ms delay then occurred without the fiducial markers. There was then a second 1050 ms delay during which two 500 Hz tones were played as before. Following the 1050 ms delay was a 50 ms blank interval, during which a single 750 Hz tone was played for 250 ms. There was a final 500 ms delay with only the fixation cross, after which participants were prompted to respond, and given unlimited time to do so. Reaction time was not recorded.

The trial sequence for *late-interval blocks* was similar to that of early-interval blocks. The differences were that the first interval was blank with no stimulus or fiducial markers, and instead the second interval contained the stimulus and fiducial markers. Additionally, the cue appeared below fixation to indicate a late-interval trial.

In the sequential condition, the cue indicated when the target was most likely to appear and with certainty indicated the location. The trial began with a fixation cross for 1500 ms, followed by a cue for 550 ms that consisted of a red and a blue square either one degree to the left or one degree to the right of fixation to indicate where the target would appear. The squares were positioned vertically such that one square appeared one degree above fixation, and the other square appeared one degree below fixation. The vertical location of the cue indicated which interval the target was most likely to appear. Specifically, a cue above fixation indicated that the target had an 80% chance to appear in the first interval, and a cue below fixation indicated that the target had an 80% chance to appear in the second interval.

Following the cue was a 500 ms delay with only the fixation cross. There was then a 1050 ms delay with a fiducial marker, and two 500 Hz tones were played as before. The first interval was then shown for 50 ms, and a single 750 Hz tone was played for 250 ms. The first interval was followed by a 500 ms delay with the fiducial markers. An additional 500 ms delay then occurred without the fiducial markers. There was then a second 1050 ms delay during which the fiducial marker was displayed and two 500 Hz tones were played as before. The second interval was then shown for 50 ms, accompanied by a 750 Hz tone for 250 ms. Following the second interval was a 500 ms delay with the fiducial marker. Participants were then prompted to respond, and given unlimited time to do so. The total SOA between the onset of the intervals was 2100 ms.

Prior to the experiment, participants completed 2-3 training sessions during which they learned to use the cues and perform the task. Participants then completed 25 experimental sessions. Each session consisted of 8 randomly-ordered blocks of 20 trials, making 160 trials per session and 4000 trials per participant.

Predictions

According to selective perception, perceptual encoding involves a limited resource that can be selectively allocated in space. This resource can be switched between locations, and is renewable over time. For brief simultaneous displays, there is no time to switch attention, so participants have only one chance to allocate resources. The cued location is more likely to contain the target, therefore it helps to allocate resources to that location. In contrast, in the sequential condition the location of the target is known, and the cue indicates the temporal interval that is likely to contain the target. Given a long enough SOA, participants can switch attention and allocate resources to both the cued and uncued intervals. Under these conditions, the selective perception hypothesis predicts no advantage for valid over invalid cues.

According to selective decision, perceptual encoding is unlimited in capacity: the quality of perceptual encoding cannot be enhanced by selective attention. Instead, cueing effects arise because the cue is used to improve decision-making. In both the simultaneous and sequential conditions, participants weight information at the cued location more heavily than information at the uncued location. Therefore, the selective decision hypothesis predicts a cueing effect for both conditions.

Results

In Figure 2a, percent correct is shown for both the valid and invalid cueing, and for both simultaneous and sequential displays. A repeated-measures ANOVA was used with stimulus condition (simultaneous/sequential) and cue condition (valid/invalid) as factors. Performance was better when the cue was valid compared to invalid, $F(1, 12) = 20.23$, $p < .001$. There was no effect of stimulus condition, $F(1, 12) = 0.31$, $p = .59$, and no interaction, $F(1, 12) = 0.44$, $p = .52$.

For the simultaneous condition, average performance was 76.7 ± 1% for the valid condition, and 68.5 ± 2% for the invalid condition. As shown in Figure 2b, the mean difference between the valid and invalid conditions was 8 ± 2% and was reliable, $t(12) = 4.27$, 95% CI [3.9, 12.3]. For the sequential condition, average performance was 76.6 ± 1% for the valid condition, and 69.4 ± 1% for the invalid condition. The mean difference was 7 ± 2% and was reliable, $t(12) = 4.13$, 95% CI [3.4, 11.1].

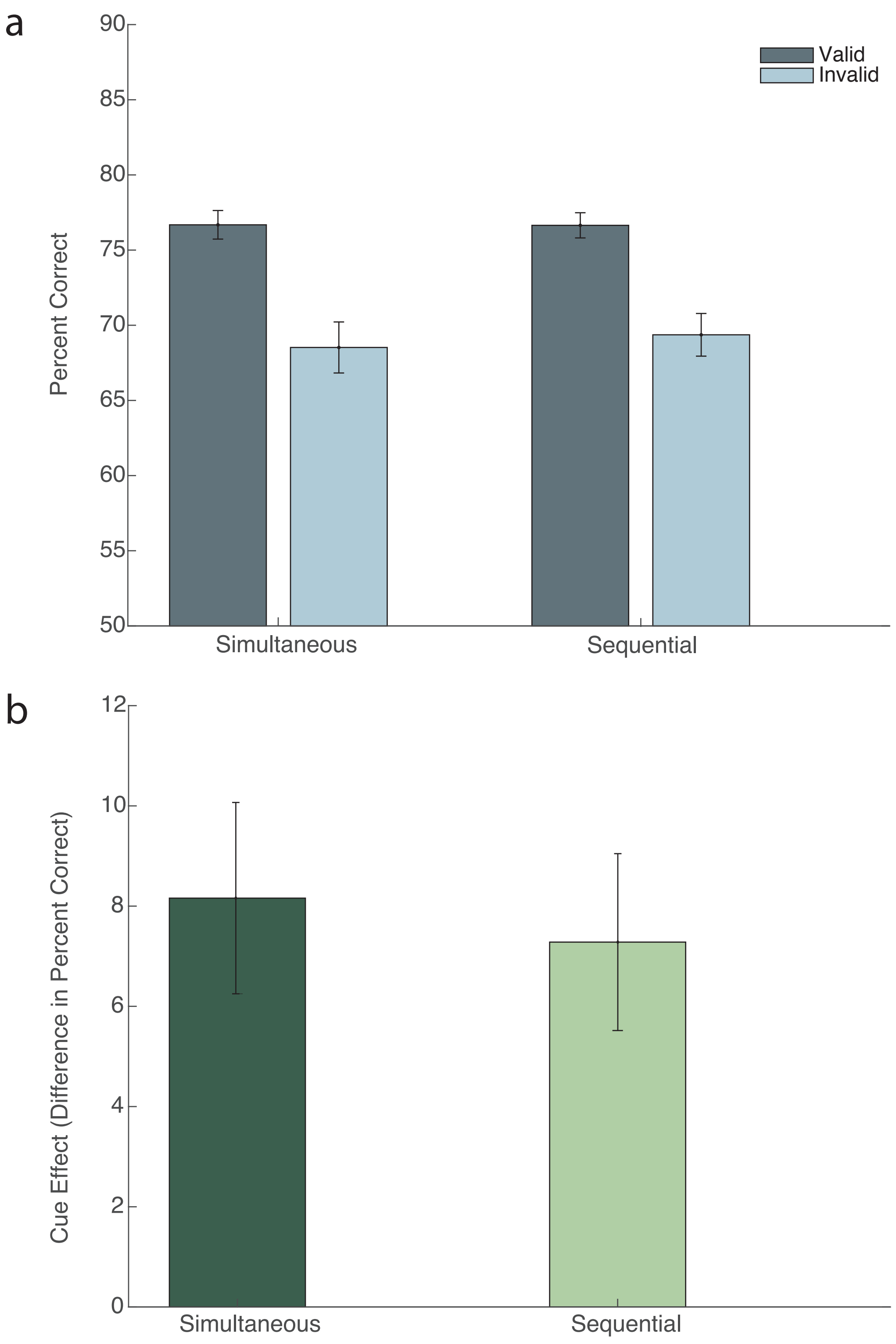

Figure 2. a) Percent correct for valid and invalid cues, for both the simultaneous and sequential condition. Error bars represent standard errors. b) The cueing effect for the simultaneous and sequential conditions, which is calculated as the difference in performance when the cue was valid versus when it was invalid.

General Discussion

Summary of Results

Selective perception and selective decision can both account for typical partially-valid cueing effects. To distinguish them, we used the simultaneous-sequential paradigm in which stimuli in the high- and low-probability locations were either presented simultaneously within one interval, or sequentially across two intervals. The SOA for the sequential condition was long enough to make it likely that participants could switch attention on all trials. We found similar cueing effects for simultaneous and sequential conditions. This result is inconsistent with selective perception, which posits the cueing effect is due to limited capacity in encoding. Instead, it is consistent with selective decision, which posits the cueing effect is due to using cue information in decision-making.

Relation to Previous Research on Temporal Cueing

Consider next the literature on temporal cueing. We focus on studies with SOAs of 500 ms or more because both selective perception and selective decision predict cueing effects for short SOAs. There is not enough time to switch attention. For general reviews see Correa, Lupianez, Madrid and Tudela (2006); Rolke and Ulrich (2012); and, Shimozaki, Schoonveld and Eckstein (2012).

In an early temporal cueing study, Coull and Nobre (1998, see also Griffin, Miniussi, & Nobre, 2001) used a character discrimination task, and their most relevant conditions had SOAs of 0.3 and 1.5 s. The magnitude of the cueing effects declined sharply with SOA. In another study, Denison, Heeger and Carrasco (2017) measured temporal cueing using fine orientation discrimination among Gabor patches and a post-cue. They found near-zero cueing effects for a

SOA of 800 ms, and cueing effects of around 0.3 d' units for short SOAs. Both of these results are consistent with selective perception because the cue effect declined with SOA.

Other studies have found results consistent with selective decision. Kinchla et al. (1995) had participants detect a target among four sequentially presented letters with an SOA of 1500 ms. Participants were better at detecting the target when it appeared in the cued interval than when it appeared in the uncued interval, which is consistent with selective decision. Correa, Lupianez and Tudela (2005) measured temporal cueing in an RSVP sequence using letter detection with SOAs of 414 and 1,057 ms. For the long SOA, they found a cueing effect of about 0.4 d' units. These effects were similar to what was found for the short SOA, which is consistent with selective decision.

Why might there be such divergent results? An interesting explanation is that different tasks have different capacity limits on the processing of multiple stimuli. For example, simple detection and word categorization have been shown to have different capacity limits (Pashler & Badgio, 1987; Scharff et al., 2011). These findings, taken together with the results of the current study, are consistent with selective perception for more complex tasks, and selective decision for simple detection.

Another explanation is that temporal uncertainty might have varied with SOA. Decreases in temporal uncertainty increase accuracy (Lasley & Cohn, 1981) and decrease response time (Niemi & Näätänen, 1981). Furthermore, some have shown that decreases in temporal uncertainty decrease cueing effects (Gould, Wolfgang, & Smith, 2007). In our experiments, a sequence of warning tones and fiducial markers minimized temporal and spatial uncertainty. Our pilot studies indicated that minimizing uncertainty was necessary to prevent changes in performance as a function of SOA.

The present study has larger implications. A common view is that selective perception is sufficient to account for all effects of partially-valid cueing (e.g. Denison, Heeger, & Carrasco, 2017; Nobre & Ede, 2017). Based on our results, and the results we have cited in the literature, we argue that selective perception alone can be rejected. Instead, theories are needed that include a role for both selective decision and selective perception.

Conclusion

The simultaneous–sequential paradigm was adapted to distinguish two hypotheses for selective attention: selective perception versus selective decision. For a detection-like coarse orientation discrimination task with endogenous cues and clear displays, the results were consistent with selective decision and not selective perception. Other studies using different tasks have found contrary results. While the differences among these studies need to be sorted out, we argue that one can reject selective perception as the universal account of partially-valid cueing.

Authorship

Palmer and Moore developed the study concepts. All authors contributed to the design. Data collection and analysis was done by Johnson and she drafted the manuscript. All authors contributed to its revisions and approved the final version.

Author's Note

Miranda L. Johnson, John Palmer, and Geoffrey M. Boynton, Department of Psychology, University of Washington; Cathleen M. Moore, Department of Psychology, University of Iowa. Correspondence concerning this article should be addressed to Miranda L. Johnson, Department of Psychology, University of Washington, Seattle, WA 98195. Email: mlpetty@uw.edu

Open Practices

This experiment was not preregistered. Datasets generated and analysed during the current study are available in the Open Science Framework repository: osf.io/nhpqu/. Other materials are available upon request.

References

Attarha, M., Moore, C. M., Scharff, A., & Palmer, J. (2014). Evidence of Unlimited-Capacity Surface Completion. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(2), 556–565.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 443–446.

Correa, Á., Lupiáñez, J., Madrid, E., & Tudela, P. (2006). Temporal attention enhances early visual processing : A review and new evidence from event-related potentials. *Brain Research*, *1076*, 116–128.

Correa, Á., Lupiáñez, J., & Tudela, P. Í. O. (2005). Attentional preparation based on temporal expectancy modulates processing at the perceptual level. *Psychonomic Bulletin and Review*, *12*(2), 328–334.

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *Journal of Neuroscience*, *18*(18), 7426–7435.

Denison, R. N., Carrasco, M., & Heeger, D. J. (2017). A dynamic normalization model of temporal attention. Poster session presented at the meeting of Society for Neuroscience, Washington, D.C.

Denison, R. N., Heeger, D. J., & Carrasco, M. (2017). Attention flexibly trades off across points in time. *Psychonomic Bulletin and Review*, *24*(4), 1142–1151. https://doi.org/10.3758/s13423-016-1216-1

Dosher, B. A., & Lu, Z.-L. (2000). Noise exlusion in spatial attention. *Psychological Science*, *11*(2), 139–146.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli.

Psychological Review, *87*(3), 272–300.

Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, *369*(6478), 313–315.

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

Gould, I. C., Wolfgang, B. J., & Smith, P. L. (2007). Spatial uncertainty explains exogenous and endogenous attentional cuing effects in visual signal detection. *Journal of Vision*, *7*(13), 1–17.

Griffin, I. C., Miniussi, C., & Nobre, A. C. (2001). Orienting attention in time. *Frontiers in BioScience*, *6*(12), 660–671.

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

Kinchla, R. a, Chen, Z., & Evert, D. (1995). Precue effects in visual search: data or resource limited? *Perception & Psychophysics*, *57*(4), 441–450.

Lasley, D. J., & Cohn, T. (1981). Detection of a luminance increment: effect of temporal uncertainty. *Journal of the Optical Society of America*, *71*(7), 845.

Moore, C. M., Egeth, H., Berglan, L. R., & Luck, S. J. (1996). Are attentional dwell times inconsistent with serial visual search ? *Psychonomic Bulletin and Review*, *3*(3), 360–365.

Niemi, P., & Näätänen, R. (1981). Foreperiod and Simple Reaction Time, *89*(1), 133–162.

Nobre, A. C., & Ede, F. Van. (2017). Anticipated moments : temporal structure in attention. *Nature Reviews Neuroscience*, *19*(1), 34–48.

Norman, D. A., & Bobrow, D. G. (1975). On Data-limited and Resource-limited Process.

Cognitive Psychology, *7*, 44–64.

Pashler, H., & Badgio, P. C. (1987). Attentional issues in the identification of alphanumeric characters. *Attention and Performance 12: The Psychology of Reading*, 63–81.

Petersen, A., Kyllingsbæk, S., & Bundesen, C. (2012). Measuring and modeling attentional dwell time, 1029–1046.

Posner, Michael I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*(2), 160–174.

Rohenkohl, G., Gould, I. C., Pessoa, J., & Nobre, A. C. (2014). Combining spatial and temporal expectations to improve visual perception. *Journal of Vision*, *14*(4), 1–13.

Rolke, B., & Ulrich, R. (2012). On the locus of temporal preperation: enhancement of premotor processes? *Attention and Time*, 227–241.

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*(3), 813–833.

Scharff, A., Palmer, J., & Moore, C. M. (2013). Divided attention limits perception of 3-D object shapes. *Journal of Vision*, *13*(2), 1–24.

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

Shimozaki, S. S., Eckstein, M. P., & Abbey, C. K. (2003). Comparison of two weighted integration models for the cueing task: linear and likelihood. *Journal of Vision*, *3*(3), 209–229.

Shimozaki, S. S., Schoonveld, W. A., & Eckstein, M. P. (2012). A unified Bayesian observer analysis for set size and cueing effects on perceptual decisions and saccades. *Journal of*

Vision, *12*(6), 1–26.

Smith, P. L. (2000). Attention and Luminance Detection : Effects of Cues , Masks , and Pedestals. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(4), 1401–1420.

Thomas, J. P., & Gille, J. (1979). Bandwidths of orientation channels in human vision. *Journal of the Optical Society of America*, *69*(5), 652–660.

CHAPTER THREE:

Selective attention to masked words is consistent with an all-or-none serial process

Miranda Johnson[1]

John Palmer[1]

Cathleen M. Moore[2]

Geoffrey M. Boynton[1]

University of Washington[1]

University of Iowa[2]

Key words: attention, selective attention, cueing, visual word recognition

Abstract

The partially valid cueing paradigm, in which a cue indicates where a target will most likely appear, is a common way to investigate selective attention. For reading words, results using this paradigm are mixed, with some results consistent with a parallel processing model and others consistent with a serial processing model. One possible explanation for this discrepancy is that some designs may have allowed for subjects to switch attention from the cued to the uncued location within a trial. In the present study we used a post-mask to prevent attentional switching. In a semantic judgement task, we found that when words appeared at the uncued location (20% of trials), performance was at chance. Such chance performance is consistent with the all-or-none serial model for selective attention. For masked displays, only one word can be processed at a time.

Our environment is full of visual information, and selective attention allows one to select important information while filtering out distractions. The partially valid cueing paradigm is one way to measure selective attention for visual stimuli (Posner, 1980). In partially valid cueing, a visual cue indicates where a target stimulus is most likely to appear, and participants are instructed to detect the target. For valid cues, the target appears at the likely, cued location. For invalid cues, the target appears at the unlikely, uncued location. Performance is better for valid cues than invalid cues, which is called a cueing effect.

Prior work has investigated the mechanisms underlying cueing effects, and the extent to which uncued information is accessed. For tasks involving detection of simple features, such as a coarse discrimination task using Gabor patches, cueing effects are consistent with a weighted parallel model (Johnson, Palmer, Moore, & Boynton, 2020) in which information from the cued and uncued locations are processed in parallel, but the cued location is given priority in decision making. Weighted parallel models have been investigated extensively in the cueing literature (for further discussion, see Kinchla, Chen, & Evert, 1995; Shimozaki, Eckstein, & Abbey, 2003).

Although cueing effects for simple features are consistent with a weighted parallel model, it is unclear whether cueing effects for more complex stimuli, such as words, are consistent with a parallel model or a model that assumes serial processing. McCann, Folk, and Johnston (1992) conducted a series of experiments where participants made lexical decisions on stimuli that appeared at cued or uncued locations. Stimuli were words that varied in frequency or nonwords. The cue was 80% valid, meaning that the target appeared at the cued location on 80% of trials, and at the uncued location on 20% of trials. Results showed that participants were faster at responding to the words when they appeared at the cued location versus when they appeared at the uncued location.

McCann et al. (1992) proposed two models that might account for their results. The first assumes that attentional selection is allocated to the cued location, and on invalid trials, when the target does not appear at that location, the selection mechanism must switch to the uncued location to process the target and make a response. This is a serial switching model, in which attentional selection is allocated to different locations over time. The second model is a weighted parallel model (Kinchla et al., 1995) which assumes that information from the cued and uncued locations are processed in parallel, however cued information is weighted more in decision making. McCann et al.'s results are consistent with both types because both models predict faster responses for to stimuli at cued location than uncued locations. One way to distinguish these two models is to limit attentional switching between the cued and uncued locations, resulting in all-or-none selection within a trial for a serial model. The all-or-none serial model is a special case of serial switching, in which only one stimulus can be accessed.

Prior work in divided attention has tested whether attention to words is consistent with a parallel model, or one that assumes all-or-none serial processing. White, Palmer, and Boynton (2018, 2019, 2020) tasked participants with making semantic judgements about words that appeared to either side of fixation. In the single-task condition, participants were cued to a single spatial location where a target word always appeared. In the dual-task condition, participants were cued to two spatial locations where targets always appeared. In both conditions, words appeared at two locations around fixation (above and below, or to the left and right of fixation). The authors also included a control task in which participants judged the color of the target word. White et al. (2018, 2020) used a post-mask to minimize attentional switching between objects. Divided attention effects for this semantic judgement task were consistent with the all-or-none serial model which proposes attentional selection can only be allocated to one word at a time.

Additionally, a correct response for one side in the dual-task condition was associated with an incorrect response on the other side (White et al. 2018, 2020). The negative correlation between responses was consistent with participants judging only one word in their dual-task condition, and guessing on a second word. In contrast, dual-task performance for the color task was consistent with a nearly unlimited-capacity parallel model, suggesting that participants can judge the color of two words simultaneously. Although this study was testing capacity limits in divided attention, and not selective attention, they suggest that for semantic judgement tasks that include post-masks, attentional selection can be allocated to only one word at a time.

The goal of the current study is to test whether cueing effects are consistent with the all-or-none serial model, or the weighted parallel model. To distinguish predictions for these two hypotheses, we use a semantic categorization task with partially valid cues. To make distinctions for an all-or-none serial model distinct from those of a weighted parallel model, we use a post-mask to minimize attentional switching (White, Palmer, & Boynton, 2018).

The all-or-none serial model predicts a cueing effect. Importantly, it predicts chance-performance for invalidly cued trials. In contrast, a weighted parallel model where attentional selection is allocated to both cued and uncued locations predicts above chance performance for invalidly cued trials. For the parallel model to result in chance performance at invalidly cued locations, it must behave like a serial model and fully weight the cued word over the uncued word.

Experiment

Method

Participants

There were 11 paid participants. All participants had normal or corrected-to-normal acuity. All gave written and informed consent in accord with the human subjects Institutional Review Board at the University of Washington, in adherence with the Declaration of Helsinki. To determine the appropriate sample size, we examined data from a previously conducted partially-valid cueing experiment (Johnson, Palmer, Moore and Boynton, 2020). In it, participants (n = 13) detected Gabor patches with similar methods as the current study. Based on the variability observed in that study (standard deviation of cueing effect = 7%), we conducted a power analysis to determine the sample size needed to detect a cueing effect the same size as found for the simultaneous display of that experiment (8%). Our calculations assumed an alpha error of 0.05 and a power of 95% (beta error of 0.05). The estimated minimum sample size was 10. To be conservative, we used 11 participants.

Apparatus

Displays were presented on a linearized CRT monitor (Sony GDM-FW900) with resolution 1024 by 640 pixels refreshing at 120 Hz. The monitor was viewed from 60 cm and the middle-gray background used in the experiment had a mean luminance of 56 cd/m2. Stimuli were created with MATLAB (MathWorks) and Psychophysics Toolbox (Brainard, 1997). Gaze position was monitored for all trials using an EyeLink 1000 (SR Research) and the Eyelink toolbox (Cornelissen, Peters, & Palmer, 2002). Trials without good fixation were excluded from analysis. Such excluded trials were infrequent; across participants, broken fixations or blinks occurred on 1.9 ± 0.6% of trials.

Stimuli

Participants categorized words that appeared above or below fixation. The letter strings were presented in nearly 100% contrast black Courier font (24 pt) against a gray background.

The word set used in this task was taken from White, Palmer, and Boynton (2018). Words were drawn from 12 semantic categories: animals, anatomy, clothing, food, professions, transport, plants, buildings, music, household, environment, and materials. Each category consisted of 35 nouns ranging from four to six characters in length. The median lexical frequency of the words used was 6.4 per million, according to the Clearpond database (Marian, Bartolli, Chabal, & Shook, 2012). A post-mask was used that consisted of random consonants. The number of characters in the mask matched the number of characters in each word presented within a trial.

Task

In this forced-choice discrimination task, participants were assigned two target categories from which target words were drawn (e.g. clothing and transportation), and the target categories were the same throughout the duration of the experiment. Non-target distractor words were drawn from any of the remaining 10 word categories. An example is shown in Figure 1 where the target categories were clothing and transportation, and the cue color was blue. In the leftmost column of Figure 1, the target was "belt," an example of clothing, and the distractor was "sofa," an example of furniture. The correct response was the keypress corresponding to "clothing."

Procedure

The stimulus configuration for four example trials is shown in Figure 1. Each trial began with a fixation marker for 1500 ms. The fixation marker was a black square at the center of the display that was 0.4 degrees of visual angle. A spatial cue was then shown for 500 ms. The cue consisted of two vertical lines above and below fixation. Participants were assigned a cue color of red or blue. They were told that the cue color indicated which side of fixation (above or below) a target word was most likely to appear. To make the cue endogenous, the uncued side was assigned the other color. In the example shown in Figure 1, which has a cue color of blue, participants were presented a blue vertical line on the cued side of fixation, and a red vertical line on the uncued side of fixation. The probability of the target appearing at the cued location was .8, and for the uncued location was .2.

Following the cue was a delay of 1000 ms where only the fixation marker was visible. During the delay, two 500 Hz tones were played for 250 ms each, with a 250 ms delay between each tone. Following the delay, words were presented above and below fixation for 15 ms. On valid trials, the target appeared at the cued location and the distractor appeared at the uncued location. On invalid trials, the distractor appeared at the cued location and the target appeared at the uncued location. Beginning at the same time as the stimulus, a 750 Hz tone was played for 250 ms.

Following the stimulus was an inter-stimulus interval (ISI). The ISI was followed by a post-mask of random consonants shown for 250 ms. The difficulty of the task was manipulated by varying this ISI such that performance in the valid cue condition was 70-80% correct. The average ISI across participants was 43.5 ms (range 17 to 104 ms). Following the mask was a delay of 500 ms where only the fixation marker was on the screen. Participants were then prompted to respond with a button-press indicating which of the two target categories the target word was drawn from, and were given as much time as needed to do so.

Each experimental and control session consisted of 80 trials. This design made for a total of 800 trials in the experimental condition and 240 trials in the control condition (described below). All participant completed at least two training sessions, followed by 10 experimental sessions and 3 control sessions.

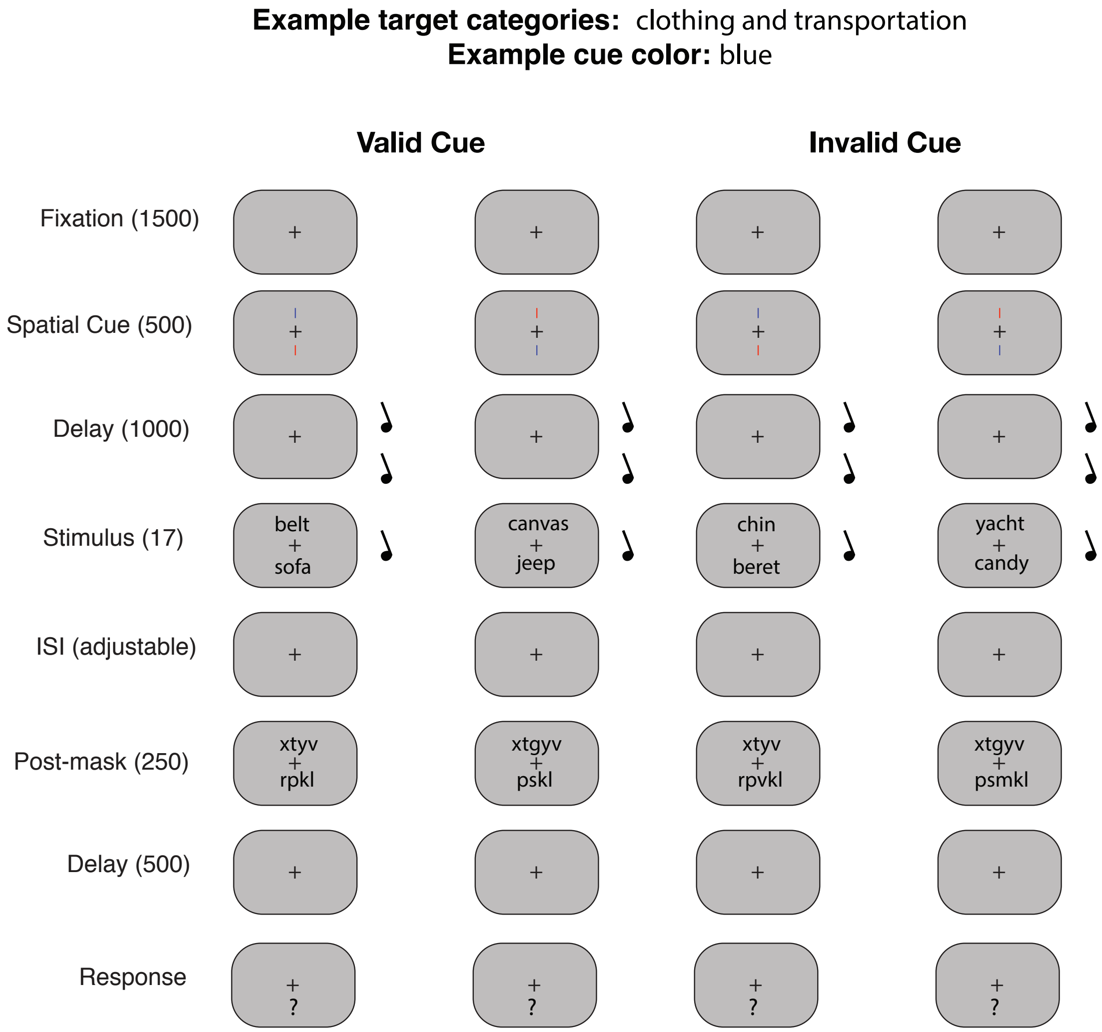

Figure 1. Schematic of four possible trial sequences for valid and invalid cues. Display durations are shown in milliseconds. In the example sequence, the cue color was blue, and the target categories were clothing and transportation. Two 500 Hz warning tones were played during the delay interval before the target, and a single 750 Hz tone was played during the target interval.

Control condition without masks

In addition to the experimental sessions, there were 3 control sessions where the ISI before the mask was 1000 ms for 10 participants, and 600 ms for one early participant. The purpose of this control condition was to confirm that errors in the task were due to the post-mask, and not characteristics of the stimulus, such as difficulty discriminating the words.

Results

In Figure 2, percent correct is shown for both valid and invalid cues. Chance performance for the categorization task was 50% correct. Performance was better when the cue was valid versus when it was invalid. The cueing effect was 27 ± 3%, which is a reliable difference, $t(10) = 10.13$, $p < .001$. Critically, performance in the invalid condition was not different from chance, 51 ± 1% $t(10) = 0.91$, $p > .05$.

Control Condition Results

In the control condition, performance in the valid cue condition was 96 ± 1%, and in the invalid condition was 78 ± 6%. The cueing effect was 11 ± 6%, which was reliable, $t(10) = 3.78$, $p < .001$. Performance in the valid condition without the mask was nearly perfect. Thus, performance was limited by the post mask rather than something like the discriminability of the target.

General Discussion

The current study tested whether selective attention to words is consistent with a weighted parallel model or an all-or-none serial model. Results showed a large cueing effect with chance performance for categorizing words at an uncued location. This is consistent with the all-or-none serial model. Our results allow us to reject a weighted parallel model where some

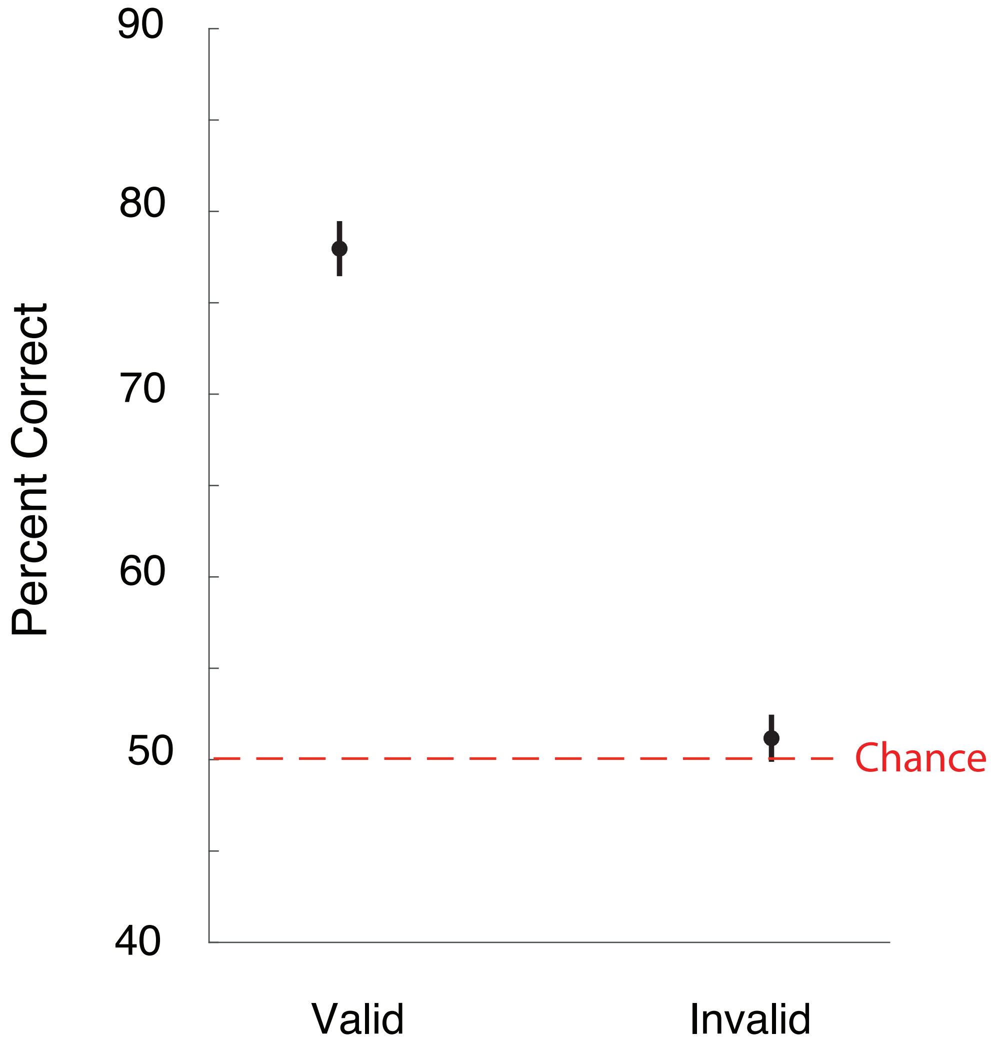

Figure 2. Percent correct for valid and invalid cues. Error bars represent the standard error of the mean. Performance for invalid cues was at chance.

amount of an attentional resource is given to uncued locations. Our results do not rule out a version of the weighted parallel model where the parallel model acts like a serial model and process only one word. However, this model would no longer be parallel and would instead be the same as an all-or-none process because participants can only process one word.

Relation to prior studies of words and partially valid cueing

The cueing effect found in the current study is larger than that found in prior partially valid cueing work using word stimuli. One of the first studies to measure cueing effects for words was McCann et al. (1992). Invalid cue performance in McCann et al. (1992) was not at chance. In that study, the difference in reaction time for valid and invalid trials was 57 ms, and errors were relatively few with a difference of 5%. There are a number of differences between the current study and McCann et al. (1992), and it is worth considering whether these differences explain the chance performance for invalid cues found in the current study.

Potentially the most important difference between the current study and McCann et al. (1992) is the use of post-masks. Post-masks are necessary to prevent attentional switching and test the all-or-none serial model (see Smith, 2000 for a review). Without a post-mask, participants can switch attention between the cued and uncued locations, allowing them to select words at uncued locations. The study of using post masks to prevent switching has a long history (Shiu & Pashler, 1994; White et al., 2018, 2020). We argue that the difference in invalid performance found in the current study and in other prior word cueing studies such as McCann et al. (1992), is due to our use of post-masks. Because McCann et al. did not use a post-masks, participants could have switched attention between cued and uncued locations.

McCann et al. measured reaction time for valid and invalid cues, and the current study measured percent correct. However, prior work using a lexical decision task similar to that in

McCann et al (1992), without the use of masks, measured performance as percent correct and invalid cue performance was 70% correct, which is well above chance (Cristescu & Nobre, 2008). Thus, it is unlikely that chance performance for invalid cues in the current study is solely due to percent correct being measured.

Another difference is that McCann et al. (1992) presented a single target stimulus with no distractors during each trial, where the current study presented both a target and a distractor. The presence of a distractor stimulus might make it more difficult to select the target than if the target were the only stimulus presented. Cristecu and Nobre (2008) conducted a partially valid cueing experiment using an unmasked lexical decision task with a distractor presented at the other location. Invalid performance was at 71% correct, which is much higher than the chance-level performance found for invalid cues in the current study. Thus, it is unlikely that using distractors alone resulted in chance performance on invalid trials. Additionally, the control condition in the current study had performance well above chance for invalid cues. Because of these results along with those of prior work, we argue that the addition of masks was critical to reducing invalid cue performance to chance.

Relation to prior studies of divided attention to words

The results observed in the current study are consistent with White et al. (2020), which found that in a divided attention tasks requiring semantic judgements of words, participants could judge only one word and were at chance for a second word. In contrast, when tasked with naming the color of the word, participants could judge two words almost as well as they could judge one. In both the semantic judgement task and color naming task, post-masks followed the target stimulus, however only the semantic judgement task produced divided attention effects that were consistent with all-or-none serial processing. These results suggest that it was the combination of the semantic judgement task and post-mask that produced the all-or-none serial result, rather than just the post-mask. Consistent with the findings of White et al. (2018), the current study found that in our selective attention task, categorization of uncued words were at chance.

Our results are also consistent with prior work using the redundant target paradigm that found results consistent with the serial process of words (Mullin & Egeth, 1989). In the redundant target paradigm (van der Heijden, 1975), one or two stimuli are presented that are either targets or distractors. The special property is that for the two-stimulus condition, the stimuli are either both targets or both distractors. The task is to respond "yes" when a target is present anywhere in the display and "no" otherwise. A self-terminating serial model predicts equal performance for the single- and two-stimulus condition while many parallel models predict a redundancy gain for the two-stimulus condition. Mullen and Egeth (1989) applied this paradigm to the semantic judgment of words and found no redundancy gain and interpreted their results as consistent with the serial processing of words.

Relation to prior studies of selective attention to words

Results from prior work using the Stroop paradigm has been used to argue that processing of words occurs automatically, without attentional selection (e.g. Brown, Gore, & Carr, 2002). In the standard Stroop paradigm, a color name is presented in an ink color that might or might not correspond (e.g. the word "RED" printed in green ink). The task is to report the ink color but under some conditions the word meaning can affect the response to the ink color (for a review see Macleod, 1991). A common finding is that participants are more accurate when the color of the ink matches the identity of the word (e.g. the word ‘Blue’ presented in blue ink) than when it does not. When there is a mismatch between color name and the color that the text is presented in, participants report the color name, which suggests that processing of the semantic meaning of the word is automatic and occurs in parallel with unlimited capacity. However, findings from more recent research indicates that standard Stroop paradigm does not manipulate attentional selection carefully enough, and that attentional selection is allocated towards identifying the word. Studies that more carefully manipulated the allocation of attention found that Stroop effects are diminished (Robidoux & Besner, 2015; see Besner et al., 2016 for a review). The findings of the current study suggest that word processing does require selective attention, and that the allocation of selective attention to words is serial and no automatic.

Conclusion

The present study used partially valid cueing to test whether selective attention to words can be allocated in parallel to multiple stimuli, or if it can be allocated to only one stimulus at a time. In an experiment using masked words to prevent attentional switching, results showed a large cueing effect and performance at the uncued location was at chance. This result is consistent with the all-or-none serial model. This adds to previous evidence from dual task and redundant target paradigms that words are processed one at a time.

Authorship

All authors developed the design. Data collection and analysis was done by Johnson, and she drafted the manuscript. All authors contributed to its revisions and approved the final version.

Author Note

Thanks go to Alex White for insightful suggestions about the study. This research was partially supported by a NIH grant EY1295 to G.M. Boynton.

Open Practices Statement

This study was not preregistered. Data from the study is available at the OSF data suppository: (under construction).

References

Brown, T. L., Gore, C. L., & Carr, T. H. (2002). Visual Attention and Word Recognition in Stroop Color Naming : Is Word Recognition " Automatic "? *Journal of Experimental Psychology: General*, *131*(2), 220–240. https://doi.org/10.1037//0096-3445.131.2.220

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox : Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavioral Research Methods, Instruments, & Computers*, *34*(4), 613–617.

Cristescu, T. C., & Nobre, A. C. (2008). Differential Modulation of Word Recognition by Semantic and Spatial Orienting of Attention. *Journal of Cognitive Neuroscience*, *20*(5), 787–801.

Ducrot, S. D., & Grainger, J. (2007). Deployment of spatial attention to words in. *Perception & Psychophysics*, *69*(4), 578–590.

Kinchla, R. a, Chen, Z., & Evert, D. (1995). Precue effects in visual search: data or resource limited? *Perception & Psychophysics*, *57*(4), 441–450. https://doi.org/10.3758/BF03213070

Macleod, C. M. (1991). Half a Century of Research on the Stroop Effect : An Integrative Review. *Psychonomic Bulletin*, *109*(2), 163–203.

McCann, R. S., Folk, C. L., & Johnston, J. C. (1992). The role of spatial attention in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1015–1029.

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

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

Shimozaki, S. S., Eckstein, M. P., & Abbey, C. K. (2003). Comparison of two weighted integration models for the cueing task: linear and likelihood. *Journal of Vision*, *3*(3), 209–229. https://doi.org/10.1167/3.3.3

Shiu, L., & Pashler, H. (1994). Negligible Effect of Spatial Precuing on Identification of Single Digits. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1037–1054.

Smith, P. L. (2000). Attention and Luminance Detection : Effects of Cues , Masks , and Pedestals. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(4), 1401–1420. https://doi.org/10.1037//0096-I523.26.4.H01

White, A. L., Palmer, J., & Boynton, G. M. (2018). Evidence of Serial Processing in Visual Word Recognition. *Psychological Science*, *29*(7), 1062–1071. https://doi.org/10.1177/0956797617751898

White, A. L., Palmer, J., & Boynton, G. M. (2020). Visual word recognition : Evidence for a serial bottleneck in lexical access. *Attention, Perception, and Psychophysics*, *82*, 2000–2017.

CHAPTER FOUR

Investigating attentional theories of multiple object tracking using sparse displays

Miranda Johnson

John Palmer

Geoffrey M. Boynton

Key words: attention, divided attention, multiple object tracking

Abstract

Tracking one slow-moving object is easy, but as the number of objects increases, our ability to track deteriorates. We investigate two competing attentional theories for the limits on tracking multiple objects. Switching theory proposes that attention is switched from one object to the next during tracking. We focus on the more specific all-or-none serial model, which assumes that participants can only track one object, and performance is at chance for tracking a second object. Resource theory proposes an attentional resource that is distributed across all tracked objects in parallel; the more objects one tracks, the less resource for each object. We focus on the more specific fixed-capacity parallel model, where the a representation of the stimulus is formed through sampling, and more targets means each target is sampled less. The current study distinguishes these two models using a dual-task design with sparse displays to control the contribution of visual crowding. Performance was compared when participants tracked one (single-task) or two (dual-task) targets moving in separate regions of the visual field. The all-or-none serial model and fixed-capacity parallel model predict dual-task deficits of differing magnitudes. Additionally, the all-or-none serial model predicts a negative correlation between dual-task responses, while the fixed-capacity parallel model predicts no correlation. Results show a dual-task deficit that is consistent with the all-or-none serial model, but no negative correlation. We discuss alternative models that can account for these results.

Whether driving on a busy street or supervising children on a crowded playground, the ability to track moving objects is important in a dynamic environment. Despite the importance of this task, there are limits to how many objects can be tracked at once. Our ability to track moving objects is often studied using the multiple object tracking paradigm (Pylyshyn & Storm, 1988). In multiple object tracking, a display is shown with some number of moving objects, and a subset of those objects are marked as targets. Participants track the targets for some set length of time, and then are typically asked to either select all target objects, or they are probed with an object and asked whether the probed object is a target or not. Performance has generally been found to decrease with the number of objects tracked (Alvarez & Franconeri, 2007; Pylyshyn & Storm, 1988). This article pursues the interpretation of this set-size effect. More specifically, we test two attentional hypotheses that describe set-size effects in multiple object tracking. But first, we introduce an important non-attentional limit on performance in multiple object tracking.

Crowding Theory

One phenomenon that has been found to influence performance in multiple object tracking is visual crowding. Visual crowding occurs when objects presented close in space interfere with one another. Assuming a display with fixed width and height, as the number of objects in a display increases, the spacing between objects decreases, leading to a higher likelihood of crowding. One measure of the stimulus conditions under which crowding is most likely to occur is referred to as Bouma's Law (Bouma, 1970). Under Bouma's Law, each stimulus within a display has a crowding window surrounding it, and additional stimuli placed within that window result in perceptual crowding. The size of the crowding window is roughly equal to half of the object's eccentricity. For example, an object at nine degrees eccentricity has a crowding window surrounding it of approximately 4.5 degrees of visual angle. This rule captures the well-known result that crowding increases with eccentricity.

Most studies of visual crowding use task involving discrimination to show its effects (Ester, Clee, & Awh, 2013; Levi & Carney, 2009; Palomares, Pelli, & Majaj, 2001). Visual crowding also influences performance in multiple object tracking. In experiments where the spacing between moving objects has been manipulated, performance has been found to be worse when spacing is small versus when it is large (Franconeri, Lin, Pylyshyn, Fisher, & Enns, 2008; Shim, Alvarez, & Jiang, 2008). Thus, it is clear that crowding influences our ability to track moving objects, and in experiments where object spacing is not controlled, it is difficult to distinguish crowding effects from divided attention effects. Crowding theory proposes that for uncrowded displays, there is no attentional effect on performance in multiple object tracking. Thus, one way to test for attention effects in multiple object tracking is to use sparse, uncrowded displays, and measure whether divided attention effects occur.

Serial Switching Theory

The current experiment uses sparse displays to investigate alternative models of attention in multiple object tracking. Specifically, we test two attentional models for set-size effects in tracking: serial switching and resource theory. Under the serial switching hypothesis, participants attend to one object at a time, and must switch attention to track multiple objects. An example of a serial switching model is one in which the locations of objects are recorded and updated over time (Holcombe & Chen, 2013). By this model, when a target object is attended, the location of that target is recorded before attentional selection switches to a different object. When a participant switches their attention back to a given target, they return to its most recently recorded location. If the target is still near that location, participants can select it and update their record of that object's spatial position. However, if the target has moved far away from its most recently recorded location, or if a distractor has moved near to the recorded location, the target is lost. By this hypothesis, set-size effects occur because increasing the number of targets increases the amount of time until selection returns to a given object and updates its location (Holcombe & Chen, 2013). Thus, increasing numbers of targets is associated with worse performance (i.e. a set size effect).

An important variable that influences performance in MOT is speed. As speed increases, performance declines. For the serial switching model, the effect of speed can be understood by the related idea of spatial frequency (Holcombe & Chen, 2013). In this context, temporal frequency is the rate at which objects pass through a given spatial location. For a circular trajectory with N objects and a fixed object speed, each point along the trajectory has a frequency at which an object passes through it. More objects along the trajectory results in a higher temporal frequency. Under the switching hypothesis, a higher temporal frequency leads to a higher likelihood of the target being lost because there is less time until a distractor occupies the location where the target was most recently selected. Thus, the serial switching hypothesis predicts that performance decreases with increasing temporal frequency.

The serial switching model is difficult to distinguish from models that assume limited capacity parallel processing. Both models predict set-size effects. To make predictions for serial switching distinct, we study conditions where there is little time to switch attention. Such conditions might result in a specific version of the serial switching hypothesis, called the all-or-none serial model. Under the all-or-none serial model, attentional switching is not possible. Instead, participants choose one target to track and stick with it through the duration of the trial without switching attention. Performance is predicted to be at chance for the unattended target.

Additionally, this model predicts that there is a negative correlation between responses in the dual-task condition. This extreme version of serial switching makes distinct predictions for divided attention effects across a variety of tasks. In other domains, evidence of all-or-none serial switching has been found in tasks where attention is divided across words separated in space (White, Palmer, & Boynton, 2019), visual search tasks that require different stimulus-response mappings (Sperling & Melchner, 1978) and tasks where attention is divided between different features of different objects (Bonnel & Prinzmetal, 1998).

Resource Theory

The second general hypotheses for why there are set-size effects in multiple object tracking is resource theory. Resource theory posits that a limited attentional resource is shared between attended objects, and the more objects that are tracked, the less of the resource is dedicated to each object. One way to implement the idea of a limited attentional resource is to assume that the speed at which objects can be tracked depends on how much of the resource is allocated to each object (Alvarez & Franconeri, 2007). As set size increases, the amount of resources allocated to each target decreases, resulting in worse performance.

To make the predictions of resource theory more concrete, we focus on a specific version of resource theory, called the fixed-capacity parallel model (Shaw, 1980). Fixed-capacity refers to extracting a constant amount of position information from the display per unit time. When estimates of a target object’s position become sufficiently noisy, the target is lost. One way to implement this abstract idea is to assume that each target’s representation is formed through a process of sampling, and the total number of samples is fixed (the sample size model; Horowitz & Cohen, 2010; Miller & Bonnel, 1994; Smith, Lilburn, Corbett, Sewell, & Kyllingsbæk, 2016). For multiple stimuli, equal numbers of samples are drawn from each object in parallel. Each object representation can be thought of as having an associated sampling distribution, and the standard deviation of the distribution is smaller with increasing numbers of samples, yielding a more accurate stimulus representation. The more stimuli that are attended, the fewer samples that are drawn from each distribution, which results in a sampling distribution with a larger standard deviation and thus a less accurate representation of the stimulus. Prior work in multiple object tracking has found set-size effects that are consistent with the fixed-capacity parallel model (Horowitz & Cohen, 2010), making it an obvious version of resource theory to test.

The Current Study

Predictions for the all-or-none serial model and fixed-capacity parallel model can be distinguished using an attentional operating characteristic (AOC; Sperling & Melchner, 1978). AOCs allows us to compare predictions for our two attentional models and predictions of crowding theory, which assumes that in sparse displays with no visual crowding, there are no attentional effects. The AOC method is commonly used in dual tasks to measure divided attention effects. Participants complete either one task (single-task condition), or two tasks simultaneously (dual-task condition). If the two tasks are independent, performance in the dual-task condition for each task is equal to performance in the single-task condition. If the two tasks are dependent in some way, dual-task performance is worse than single-task performance. AOCs have been used in prior work to measure dual-task deficits in multiple object tracking (Alvarez, Horowitz, Arsenio, DiMase, & Wolfe, 2005), however they have not yet been used to distinguish predictions of attentional theories in a task with sparse displays.

The current study uses a dual task where participants tracked either one or two targets that appeared above or below fixation. To make distinct the predictions for crowding theory, discs were widely spaced such that perceptual crowding was unlikely (Bouma, 1970). To distinguish the predictions for switching, fast disc motion was used so that attentional switching would be unlikely, which leaves one with the all-or-none serial model (Holcombe & Chen, 2013). We also focus on the fixed-capacity parallel model, a resource theory model that has been found in prior work to account for set-size effects in multiple object tracking (Horowitz & Cohen, 2010).

The three models described above can be distinguished by measuring the magnitude of the dual-task deficit and the correlation between accuracy for responses in the dual-task condition. The all-or-none serial model predicts a large dual-task deficit and a negative correlation between accuracy for the top and bottom responses in the dual-task condition. The fixed-capacity parallel model predicts a dual-task deficit that is smaller in magnitude than that predicted by the all-or-none serial model, and a zero correlation between the two responses in the dual-task condition. For our sparse displays, crowding theory predicts little or no dual-task deficit, and a zero correlation for accuracy in the dual-task condition.

Experiment

Method

Participants

There were 11 paid participants. All participants had normal or corrected-to-normal acuity. All gave written and informed consent in accord with the human subjects Institutional Review Board at the University of Washington, in adherence with the Declaration of Helsinki.

To determine the number of participants, we used pilot data from an unpublished pilot study. Participants (N = 6) each completed a multiple object tracking task with similar methods. A dual-task deficit of 30% was observed with a standard deviation of 6%, and the correlation between accuracy for each side in the dual-task condition was r = -.05 with a standard deviation of .12. Our goal was to distinguish the fixed-capacity parallel model and the all-or-none serial model. For the dual-task condition, we need to discriminate deficits of 23% and 11%. A power analysis with 80% power suggested a minimum of 4 participants. For the response correlation, we need to discriminate correlations of -.10 and 0. A power analysis using a one-tailed test with 80% power suggests a minimum of 11 participants.

Apparatus

Displays were presented on a linearized CRT monitor (Sony GDM-FW900) with resolution 1024 by 640 pixels refreshing at 120 Hz. The monitor was viewed from 60 cm and the middle-gray background used in the experiment had a mean luminance of 56 cd/m2. Stimuli were created with MATLAB (MathWorks) and Psychophysics Toolbox (Brainard, 1997). Gaze position was monitored for all trials using an EyeLink 1000 (SR Research) and the Eyelink toolbox (Cornelissen, Peters, & Palmer, 2002). Trials containing blinks or broken fixations were excluded from analysis. Such excluded trials were infrequent; across participants, blinks occurred on only 3 ± 2% of trials, and broken fixations occurred on 3 ± 1% of trials.

Stimuli

As illustrated in Figure 1, participants were presented with six black discs that were one degree of visual angle in diameter. Three discs appeared above fixation, and three appeared below fixation. The discs were positioned along invisible circular paths that were centered 6 degrees above and below fixation. The diameter of the circular path was 6 degrees so that the furthest point on the path was 9 degrees above fixation, and the closest point on the path was 3 degrees above fixation. Each disc was equally spaced around the circular trajectory at an angle of 120 degrees around the circle. The linear distance between each disc on a given trajectory was approximately 5.2 degrees in visual angle, which is larger than maximum crowding window as

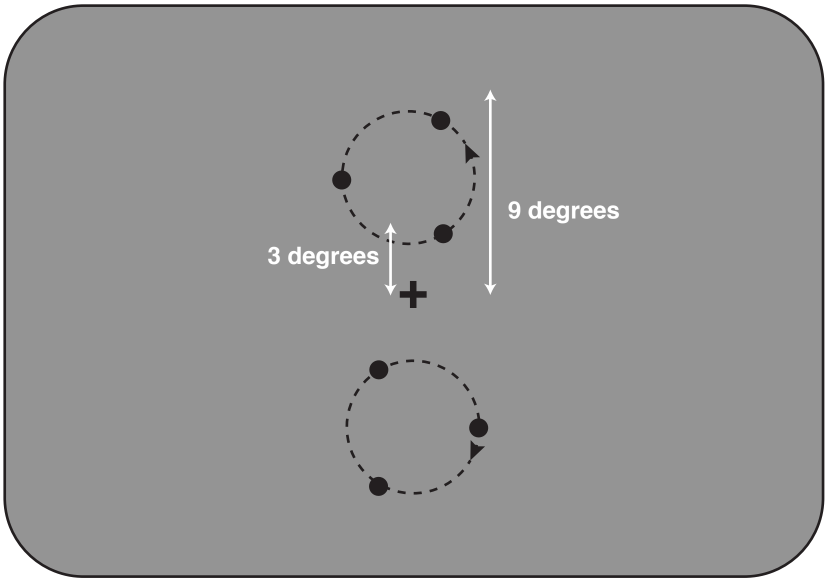

Figure 1. Stimulus object spacing. Discs were positioned along invisible circular trajectories, which are represented here by dashed circles, but were not visible in the experiment. Discs were evenly spaced along the circular trajectory. The circular trajectories were centered at six degrees in visual angle above and below fixation. The furthest points along the trajectory were 9 degrees eccentricity, and the closest points were 3 degrees eccentricity.

estimated by Bouma's law (Bouma, 1970).

The difficulty of the task was manipulated for each participant to maintain average single-task performance between 70-80% correct. Task difficulty was controlled by changing the speed of the disc motion. We varied rotational speeds between 1 and 2.2 rps. The maximum speed was limited to 2.2 rps to maintain the appearance of continuous motion. The average disc speed needed to obtain single-task performance between 70-80% correct was 1.6 rps (range 1.25 to 1.95 rps). This is equivalent to a linear speed of 29.5 degrees per second, or about 6 pixels per frame (120 Hz).

Procedure

The single- and dual-task conditions are shown schematically in Figure 2. In the single-task condition, participants tracked a single target that appeared above or below fixation. Each trial began with a blank screen for 1.5 seconds, and participants were told that they should use this blank period to blink as much as necessary and then not blink during the moving display. Following the blank period, the cue was shown for 1.5 seconds, during which six discs appeared on the screen, with three above and three below fixation. The single target disc was displayed in red, and all other discs were black. The cue was incorporated into the fixation point such that the top half of the stem on the fixation cross appeared in blue when the top half of the display was cued, and the bottom half of the stem on the fixation cross appeared blue when the bottom half of the display was cued.

The target then changed to black to appear identical to the distractors, and each set of three discs immediately began moving along an invisible circular trajectory for 4 seconds. During the four seconds of disc motion, each set of discs reversed direction three times, and when those reversals could occur was determined pseudo-randomly and independently for each side of fixation. For one set of discs, an opportunity for reversal occurred every 0.5 seconds, and for the other set of discs, it occurred every 0.6 seconds. This difference in when reversals could occur made it such that the two sets of discs never reversed at exactly the same time. Which side reversed at 0.5 or 0.6 seconds was counterbalanced. During the 4 seconds of disc motion, participants were instructed not to blink, and trials where participants blinked were not included in the final analysis.

Following the disc motion, participants were prompted to select the target with a mouse-click, and they were given as much time as needed to do so. The response prompt was incorporated into the fixation cross and was identical to the fixation cue shown at the start of the trial. Mouse clicks that did not correspond with any of the three discs on the cued side resulted in a 500 Hz tone being played, after which participants were given another chance to respond. Following response, feedback was shown at fixation for 2 seconds. Feedback was incorporated into the fixation cross in the same manner as the cue, with green indicating a correct response and red indicating an incorrect response.

In the dual-task condition, participants were instructed to track two targets, one above and one below fixation. The trial sequence for the dual-task condition was similar to that for the single-task condition. The key differences were that instead of a single target that appeared in red above or below fixation, there were two targets in red, one above and one below fixation. Additionally, the cue at the start of the experiment indicated that both sides were relevant, thus the full stem of the fixation cross was blue. After the 4 seconds period of disc motion, participants were prompted by a response cue to either select the bottom target first and then the

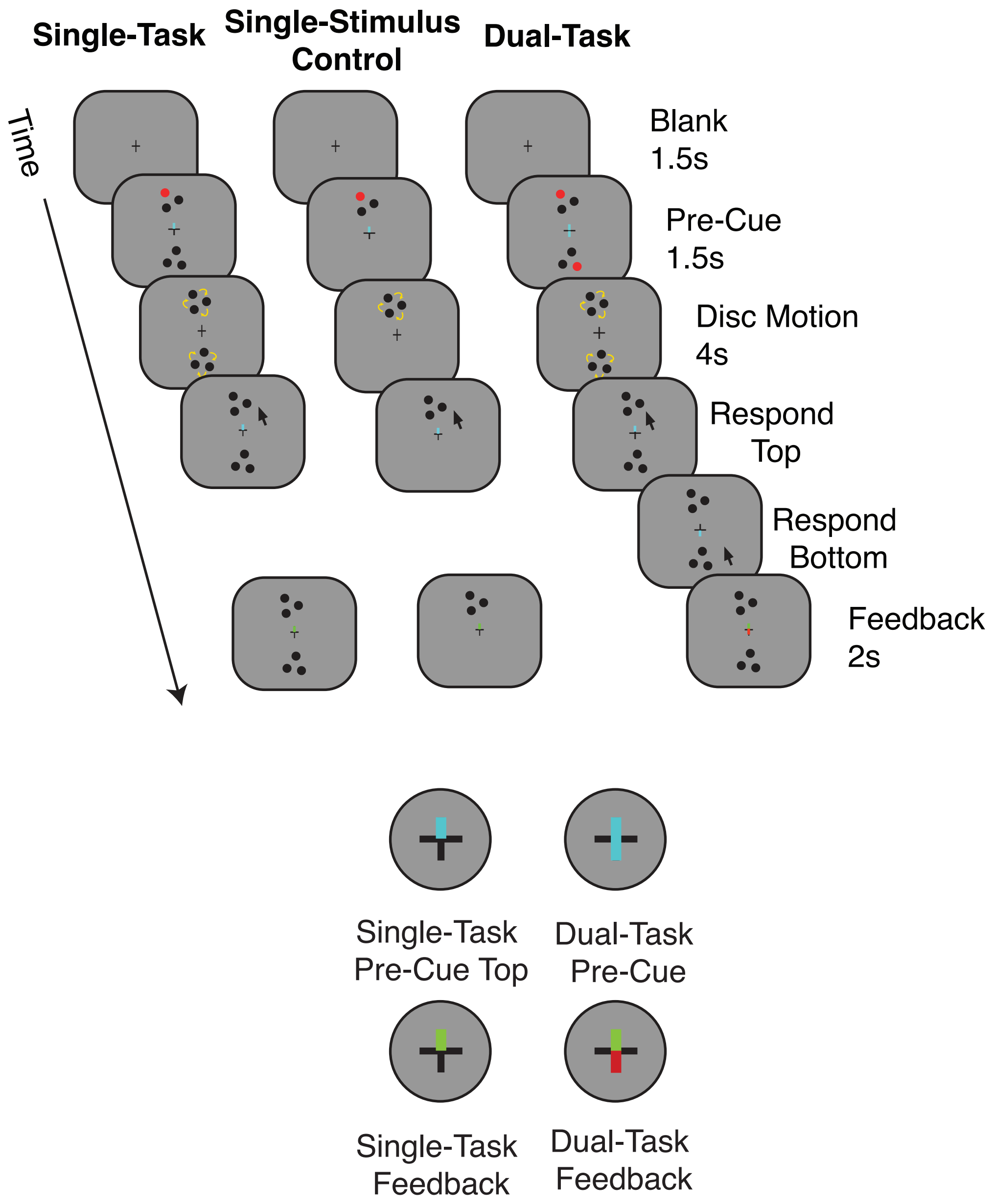

Figure 2. Trial sequence for the single-task, single-stimulus control, and dual-taks conditions. In this example of the single-task condition, the top half of the display is being cued and a target is shown above fixation in red. In the dual-task condition, both sides are cued, and two targets are shown in red, one above fixation and one below fixation. The response order in the dual-task condition is counterbalanced. Feedback was incorporated into the cue. Green indicated a correct response, and red indicated an incorrect response.

top target, or vice versa, and the order of response was counterbalanced. Feedback was shown simultaneously after both responses for 2 seconds, and was incorporated into the fixation cross in the same manner as the cue.

In addition to the single- and dual-task conditions, there was a single-stimulus control condition. For these trials, three discs appeared on the cued side of the display, and the uncued side of the display was blank. The trial sequence, stimulus, and response were otherwise identical to that of the single-stimulus condition. This condition was included as a test of crowding phenomenon.

Prior to the experiment, participants completed 2-3 training sessions, during which they learned to use the cues and perform the task. Participants then completed 20 experimental sessions, which took between fifteen to twenty hours, completed across several weeks. Each session consisted of 8 blocks of 12 trials, making 96 trials per session and 1920 trials per participant. Within a session of 96 trials, there were 24 single-task trials, 24 single-task control trials, and 38 dual-task trials. A mixed design was used such that the three conditions were randomly intermixed throughout a session of 96 trials.

Primary Results

Dual-task deficit

Performance in the single-task condition was 72 ± 2%, and performance in the dual-task condition was 52 ± 2% (chance was 33.3%). The difference is a large dual-task deficit of 20 ± 2%. To test how crowding contributes to this dual-task deficit, we compared performance in our single-task and single-stimulus control conditions. The mean difference in performance between these two conditions, which we call the dual-stimulus deficit, was 2 ± 1%. This difference was

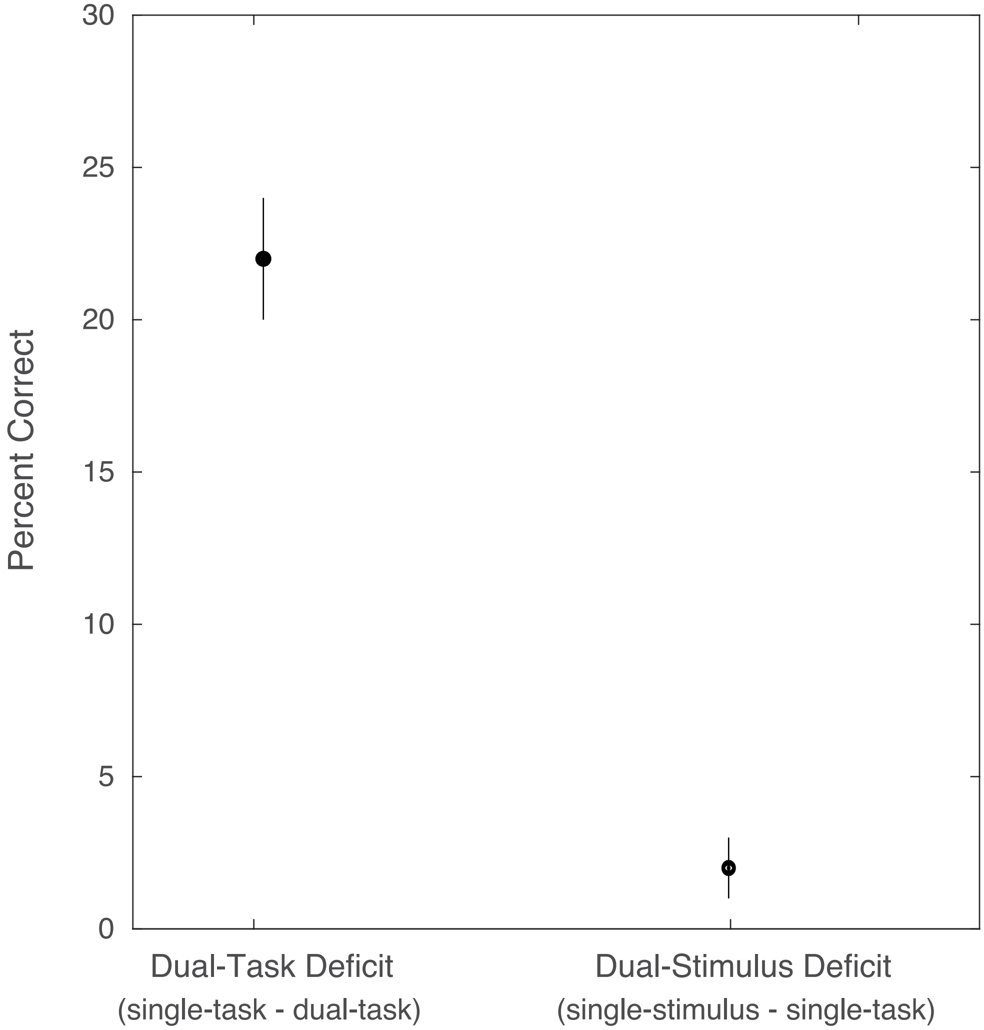

Figure 3. Comparison of the dual-task deficit and the dual-stimulus deficit. The dual-task deficit is the difference in percent correct for the single-task and dual-task conditions. The dual-stimulus deficit is the difference in percent correct for the single-stimulus control condition and the single-task condition. Error bars represent the standard error of the mean.

marginally reliable t(5) = 2.47, p = .06. While this difference suggests a small crowding effect, it is small compared to the dual-task deficit. The magnitude of the dual-task and dual-stimulus deficits are plotted side-by-side in Figure 3. The y-axis represents the difference in performance between the single-task and dual-task conditions (i.e. the dual-task deficit), and the difference in performance between the single-stimulus control and single-task conditions (i.e. the dual-stimulus deficit). The dual-stimulus deficit is a small fraction of the dual-task deficit. Thus, the residual crowding effect measured by the dual-stimulus deficits cannot account for the observed dual-task deficit.

Figure 4 shows the results compared to model predictions using the attentional operating characteristic (AOC; Sperling & Melchner, 1978). The y-axis shows performance when the target is on the top, and the x-axis shows performance when the target is on the bottom. Both axes go from chance performance (approximately 33% correct) to perfect performance (100% correct). The solid lines represent the prediction for crowding theory: if our ability to track multiple objects is only limited by perceptual crowding, then there should be no divided attention effect for this sparse display: accuracy for each of the two targets in the dual-task condition should be equal to that of the single-task condition.

The dashed diagonal line represents the prediction for the all-or-none serial model: if one can track only one target object at a time, and must guess on the location of a second target, there should be a large divided attention effect. The accuracy for the two sides trades off linearly.

The dotted curved line represents a prediction for the fixed-capacity parallel model, where processing for the two sides occurs in parallel, but is limited in capacity. Assuming signal detection theory and independent samples of the position information, one can calculate the predicted magnitude of the dual-task deficit for this model.

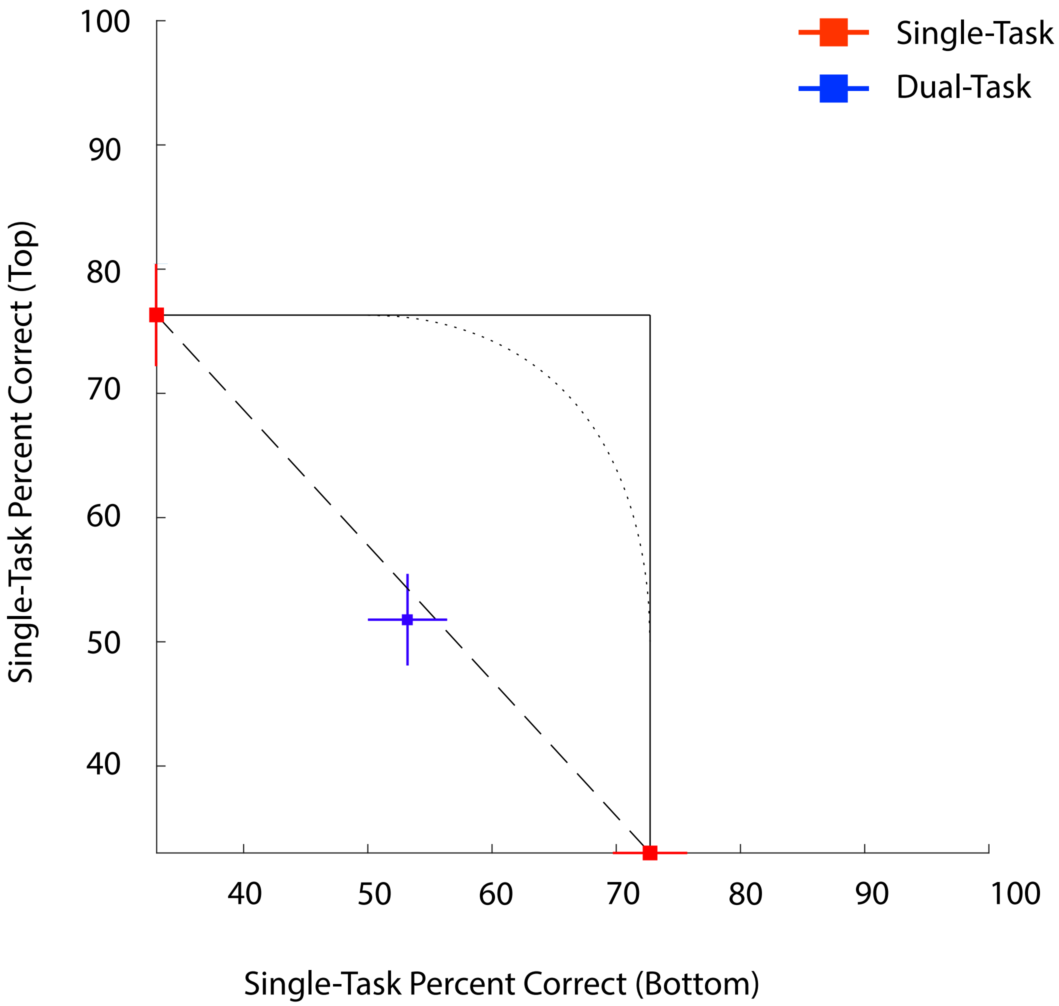

Figure 4. Percent correct for the single- (red) and dual-task (blue) conditions. Error bars represent the standard error of the mean. Model predictions are also shown. The solid lines represent the prediction for crowding theory. The dotted curve represents the prediction for the fixed-capacity parallel model. The dashed diagonal represents the prediction for the all or-none serial model.

Percent correct for the single-task condition is shown for the top (y-axis) and bottom (x-axis) responses. Single-task performance was 76 ± 4% for responses on the top, and 73 ± 3% for responses on the bottom. The dual-task deficit is plotted as a point, where the x-value represents dual-task performance for the top target, and the y-value represents dual-task performance for the bottom target. Performance for the top in the dual-task condition was 52 ± 4%, and performance for the bottom was 54 ± 3%. Plotting the dual-task deficit on the AOC reveals that the magnitude of the deficit is consistent with the all-or-none serial prediction, and much larger than predicted by the fixed-capacity parallel model.

Correlation between responses in the dual-task condition

Figure 5 shows the observed correlation along with the predictions of the three hypotheses. When performance is at chance, all three hypotheses predict a correlation of zero in accuracy between the top and bottom responses. As dual-task performance increases, the all-or-none serial model predicts a negative correlation in the accuracy between top and bottom targets. For example, for dual-task percent correct of 50%, this model predicts a negative correlation of $r = -.10$. By this model, the maximum dual-task performance is 67% correct for this 3-choice task. This model assumes that participants can only track one target in the dual-task condition, meaning that a correct response for the top side is associated with an incorrect response for the bottom side, and vice versa.

Both crowding theory and the fixed-capacity parallel model predict no correlation in dual-task accuracy across all ranges of dual-task performance. Under crowding theory, responses on each side are independent, and therefore there should not be a correlation. Under the fixed-capacity parallel model, because the limited resource is shared equally between targets in parallel, no correlation is predicted.

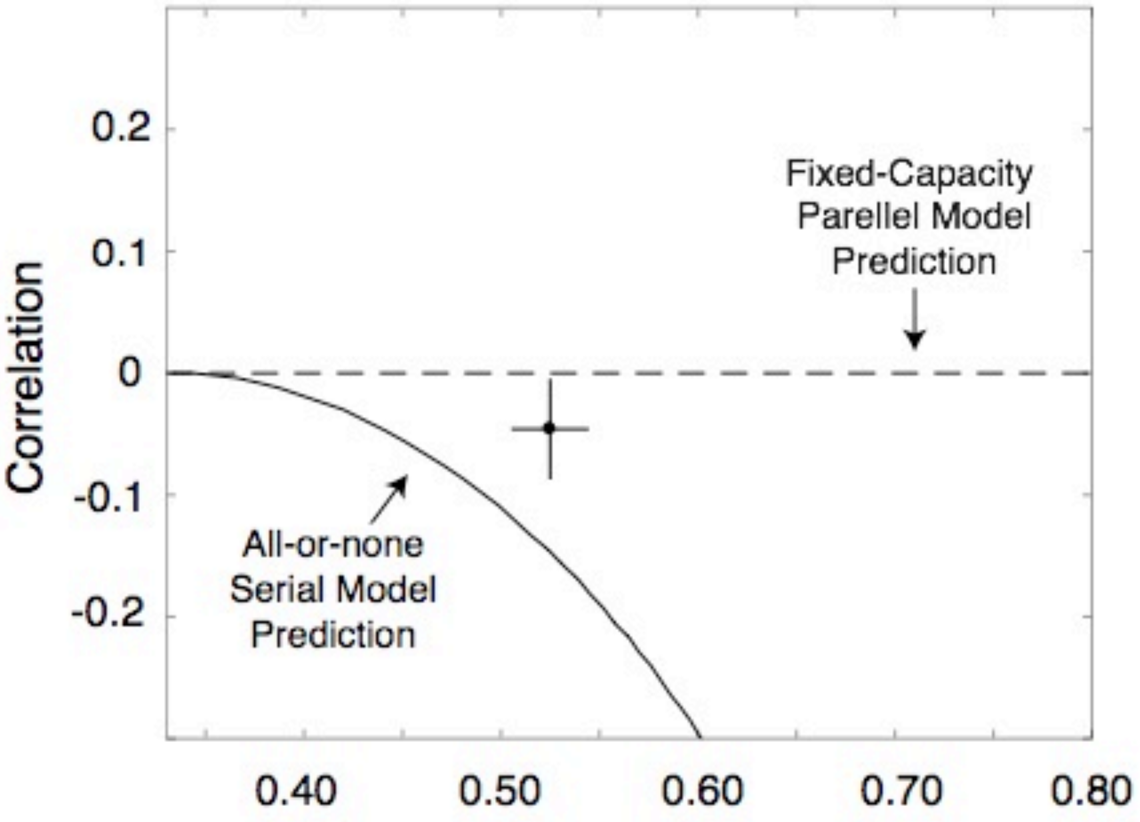

Figure 5. Predicted and observed correlations plotted as a function of dual-task percent correct. The dashed line represents predicted zero correlations under crowding theory and the fixed-capacity parallel model. The curve represents predicted negative correlations under the all-or-none serial model. The closed circle represents the average correlation, and error bars represent the standard error of the mean.

The correlation between dual-task accuracy for the top and bottom was r = -.05 ± .04 (Figure 5). This result was quite different from the r ≅ -.15 predicted by the all-or-none model for the observed dual-task performance.

Secondary Results

We tested for order effects in response accuracy for our dual-task condition by comparing performance for first responses and second responses. The difference in performance for first and second responses in the dual-task condition was -1 ± 1, with second responses being slightly better, however this difference was not reliable, t(5) = -1.94, p = .11. Thus, there is no evidence of memory or response interference for the second response.

General Discussion

Results show a large dual-task deficit, the magnitude of which is consistent with the all-or-none serial model. The observed dual-task deficit allows us to reject the fixed-capacity parallel model. Additionally, this dual-task deficit with our sparse displays provides evidence against crowding theory, which predicts no set-size effects in multiple object tracking when displays are uncrowded. The observed dual-task deficit measured in this study is consistent with those used in prior work that measured AOCs using a dual-task multiple object tracking design (Alvarez et al., 2005). Although the magnitude of the dual-task deficit was consistent with the all-or-none serial model, there was no correlation between accuracy for responses on the top and bottom in the dual-task condition, which is not consistent with the all-or-none serial model. We next put our results in the context of the larger literature, and consider alternative models that can account for this combination of results.

Implications for Switching Theory

The key prediction of the all-or-none serial model is that there is a negative correlation between responses in the dual-task condition. The current study is the first to test for negative correlations between dual-task responses using multiple object tracking. Large divided attention effects with negative correlations have been found for other tasks using a dual-task design (Bonnel & Prinzmetal, 1998; White et al., 2019), and these results provide strong evidence consistent with serial processing. Results of the current study showed no negative correlation between dual-task responses, which is not consistent with the serial switching hypothesis.

The all-or-none serial switching hypothesis assumes that participants do not switch attention, but instead track only one target through the duration of the trial. One possibility for the lack of a negative correlation is that participants tried to switch attention between targets even though it was advantageous to track only one target. If participants attempt to switch attention in conditions where switching is difficult (e.g. when fast speeds result in short temporal frequencies), dual-task performance decreases over time, but no negative correlation is predicted because both sides get a chance to be tracked. Such a scenario predicts results consistent with those found in the current study. This scenario is also consistent with the findings of Holcombe & Chen (2013), where performance was worse than a model that assumes participants can only track one object, and must guess on a second object.

Results consistent with the all-or-none serial model have been found in prior work using multiple object tracking. Holcombe & Chen (2012) tasked participants with tracking one or two targets that moved along a circular trajectory that was centered at fixation. Tracking speeds ranged from 0.7 to 1.9 rps, and psychometric functions were fitted to data for each participant. Speed limits (the speed at which participants were 68% correct for tracking one or two targets) were estimated for each participant. Results showed that speed limits were much lower for tracking two targets versus tracking one. The speed limit for the track-two condition was consistent with the all-or-none serial model.

In a follow-up study, Holcombe and Chen (2013) tasked participants with tracking one to three targets among distractors. The main manipulation was of temporal frequency, the rate at which objects passed through a given spatial location. For a circular trajectory with N objects and a fixed object speed, each point along the trajectory has a frequency at which an object passes through it. More objects along the trajectory result in a higher temporal frequency. Additionally, faster disc speeds lead to a higher temporal frequency. Under the switching hypothesis, a higher temporal frequency leads to a higher likelihood of the target being lost because there is less time until a distractor occupies the location where the target was most recently selected. Temporal frequency was manipulated by varying the number of objects on a given trajectory. On each trial, there were either 3, 6, 9, or 12 objects per trajectory. Speed thresholds, the speed at which performance fell midway between ceiling and chance, were measured for each of the tracking conditions.

The results of Holcombe and Chen (2013) showed that speed thresholds decreased with increasing numbers of target. Speed thresholds also decreased with increasing numbers of objects on each trajectory. Importantly, when speed thresholds were converted to temporal frequencies, there was no difference in temporal frequencies for the six, nine, and twelve object conditions. These results indicate that it is not speed that led to a difference in performance across object conditions, but temporal frequency. These set-size and temporal frequency effects are consistent with the more general serial switching model. Additionally, speed thresholds were worse than what is predicted by an all-or-none-type model that assumes participants can only track one target and must guess on the location of a second target. The authors proposed that performance can be worse than the all-or-none serial prediction if participants attempt to switch attention and track multiple targets, rather than giving up on switching and tracking a single target while guessing on additional targets. The results of this study can be described by the switching model, but it is unclear whether these results rule out a resource theory-like model.

Implications for Resource Theory

Prior research in multiple object tracking has attributed set-size effects and speed limits in multiple object tracking to an attentional resource that is shared across targets in parallel (Alvarez & Franconeri, 2007; Chen, Howe, & Holcombe, 2013; Holcombe & Chen, 2012). It is proposed that fast object speeds exhaust attentional resources, making it more difficult for participants to track targets. Additionally, with increasing set size, the amount of the attentional resource given to each target decreases, meaning that participants require slower speeds to track larger numbers of targets.

Results consistent with resource theory have been found in experiments using the simultaneous-sequential manipulation (Shiffrin & Gardner, 1972). Howe et al. (2010) used the simultaneous-sequential manipulation in a multiple object tracking task where target objects either move simultaneously during a single time interval, or sequentially over multiple time intervals. A model assuming unlimited attentional resources would predict no advantage for sequentially-presented stimuli over simultaneously-presented stimuli. However, a model assuming limited, parallel attentional resources would predict a sequential advantage because each time interval contains fewer moving targets to be tracked, and thus more resources are given to each target per time step. In one experiment included in this study, participants were more accurate in tracking targets in the sequential condition than the simultaneous condition, which is consistent with resource theory.

Although prior work has found results that are consistent with a limited attentional resource, discussions of resource theory are often vague about the specific characteristics of the resource and the mechanisms through which performance in multiple object tracking is influenced. To make the predictions of resource theory concrete, we focused on a specific version of a resource theory, the fixed-capacity parallel model. This model can be conceptualized using the sample size model (Bonnel & Miller, 1994; Smith et al., 2018), where a fixed number of samples is drawn from the stimulus and shared between targets, and the resolution of each target's stimulus representation decreases with increasing numbers of targets. Set-size effects that are consistent with the sample size model have been found in prior work using a multiple object tracking task where participants reported the direction of motion for target objects (Horowitz & Cohen, 2010). However, the dual-task deficit observed in the current study is larger than what is predicted by models that assume continuous, parallel sampling of target position. Our results can instead be predicted by a model that assumes repeated discrete, parallel sampling.

Like the fixed-capacity parallel model, a sampled discrete parallel model assumes that samples are shared between objects. However, sampled information is lost over time in a discrete manner, meaning that information about targets is either maintained at a given time, or completely lost. The loss of information compounds error possibilities over time and predicts that dual-task performance decreases over time. For the 4 second trials used in this experiment, the model can predict a dual-task deficit that is of a similar magnitude as (or even larger than) that predicted by the all-or-none serial model. Critically, it also predicts a zero correlation. Thus, this version of resource theory is consistent with the current results.

Implications for Crowding and Spatial Interference Theory

Although our sparse displays allow us to rule out crowding theory, we cannot rule out the more general spatial interference theory (Franconeri, Jonathan, and Scimeca, 2010). Spatial interference theory proposes that attentional selection of multiple targets is influenced by spatial interactions between targets (Shim et al., 2008). These interactions can occur at spatial distances larger than those associated with crowding. One way that this idea can be conceptualized is to assume that locations selected in space have a suppressive surround similar to the center-surround receptive fields found in brain areas associated with visual processing. Because object processing occurs in higher-level visual processing areas, the size of the suppressive surround is thought to span the full visual field in a manner similar to receptive fields for these brain areas. When there is only a single target, it is able to be selected and tracked. However, when there are multiple targets, there is competition between targets that result in interactions between selective regions and suppressive surrounds for each target.

The influence of target-target interactions on multiple object tracking performance can be measured by varying the spatial distance between targets and testing whether performance changes. Shim et al. (2008) conducted a series of experiments to test the influence of target-target spacing on tracking performance. In the first experiment, both target-target and target-distractor spacing were varied, and translational disc motion was used. Target-target spacing ranged from 0.45 to 2.91 degrees of visual angle, and target-distractor spacing ranged from 0.5 to 3 degrees of visual angle. Both target-target and target-distractor spacing influenced performance, with larger spacing being associated with better performance.

In a second experiment, a quadrant design was used to control target-target spacing (Shim et al., 2008). Participants tracked either one or two targets. In the two-target condition, the targets appeared either in the same quadrant, or in different quadrants. Performance was better for trials where targets were presented in different quadrants compared to those where they were presented in the same quadrant, indicating that greater target-to-target distances were associated with better performance. A third experiment used a circular display that was divided into 8 sections, and again, performance was better when the two targets were presented in separate sections versus when they were presented in the same section. Additionally, in trials where targets were presented in separate sections, larger distances between sections was associated with better performance.

The long-range spatial interactions described by spatial interference theory predict dual-task deficits for targets that are widely spaced, therefore such a model cannot be ruled out using the sparse displays such as those in the current study. One way to rule out such a model is to manipulate the perceptual organization of the stimulus using grouping. Grouping occurs when individual objects are made to appear as though they are components of a larger perceptual object. Grouping has been found to influence performance in multiple object tracking (Erlikhman, Keane, Mettler, Horowitz, & Kellman, 2014; Keane, Mettler, Tsoi, & Kellman, 2011; Yantis, 1992). If the selection mechanism described by spatial interference theory is object-based, grouping two targets together would allow both targets to fall within the selective region and neither target would fall within the suppressive surround, thus reducing the amount of competition between them. It would be therefore be predicted that the dual-task deficits would be smaller for targets that are grouped versus those that are not grouped.

Conclusion

The current study tested two broad hypotheses for set-size effects in multiple object tracking: serial switching and resource theory. We focused on specific versions of each hypothesis, the all-or-none serial model and the fixed-capacity parallel model. Sparse displays were used to control contributions from visual crowding. Results of the current study show large dual-task deficits for a tracking task where participants track either one or two targets. Our results are not consistent with the all-or-none serial model or the fixed-capacity parallel model. However, these results can be counted for by other models that fall within in the more general categories of serial switching and resource theory.

References

Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track ?: Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, *7*(4), 1–10. https://doi.org/10.1167/7.13.14.Introduction

Alvarez, G. A., Horowitz, T. S., Arsenio, H. C., DiMase, J. S., & Wolfe, J. M. (2005). Do Multielement Visual Tracking and Visual Search Draw Continuously on the Same Visual Attention Resources ? *Journal of Experimental Psychology: Human Perception and Performance*, *31*(4), 643–667. https://doi.org/10.1037/0096-1523.31.4.643

Bonnel, A., & Prinzmetal, W. (1998). Dividing attention between the color and the shape of objects. *Perception & Psychophysics*, *60*(1), 113–124.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(5241).

Chen, W. Y., Howe, P. D., & Holcombe, A. O. (2013). Resource demands of object tracking and differential allocation of the resource. *Attention, Perception, and Psychophysics*, *75*(4), 710–725. https://doi.org/10.3758/s13414-013-0425-1

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox : Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavioral Research Methods, Instruments, & Computers*, *34*(4), 613–617.

Erlikhman, G., Keane, B. P., Mettler, E., Horowitz, T. S., & Kellman, P. J. (2014). Automatic feature-based grouping during multiple object tracking. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(6), 1625–1637. https://doi.org/10.1037/a0031750.Automatic

Franconeri, S. L., Lin, J. Y., Pylyshyn, Z. W., Fisher, B., & Enns, J. T. (2008). Evidence against a speed limit in multiple-object tracking. *Psychonomic Bulletin and Review*, *15*(4), 802–

808. https://doi.org/10.3758/PBR.15.4.802

Holcombe, A. O., & Chen, W. (2012). Exhausting attentional tracking resources with a single fast-moving object. *Cognition*, *123*(2), 218–228. https://doi.org/10.1016/j.cognition.2011.10.003

Holcombe, A. O., & Chen, W. (2013). Splitting attention reduces temporal resolution from 7 Hz for tracking one object to , 3 Hz when tracking three. *Journal of Vision*, *13*, 1–19. https://doi.org/10.1167/13.1.12.Introduction

Horowitz, T. S., & Cohen, M. A. (2010). Direction information in multiple object tracking. *Attention, Perception, & Psychophysics*, *72*(7), 1765–1775. https://doi.org/10.3758/APP

Howe, P. D. L., Cohen, M. A., Pinto, Y., & Horowitz, T. S. (2010). Distinguishing between parallel and serial accounts of multiple object tracking. *Journal of Vision*, *10*(8), 1–13. https://doi.org/10.1167/10.8.11.Introduction

Keane, B. P., Mettler, E., Tsoi, V., & Kellman, P. J. (2011). Attentional Signatures of Perception : Multiple Object Tracking Reveals the Automaticity of Contour Interpolation. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(3), 685–698. https://doi.org/10.1037/a0020674

Miller, J., & Bonnel, A. (1994). Switching or sharing in dual-task line-length discrimination ? *Perception & Psychophysics*, *56*(4), 431–446.

Pylyshyn, Z. W., & Storm, R. O. N. W. (1988). Tracking multiple independent targets : Evidence for a parallel tracking mechanism *. *Spatial Vision*, *3*(3), 179–197.

Shaw, M. L. (1980). Identifying Attentional and Decision-Making Components in Information Processing. *Attention and Performance*, *8*, 277–296.

Shiffrin, R. M., & Gardner, G. T. (1972). Visual processing capacity and attentional control.

Journal of Experimental Psychology, *93*(1), 72–82. https://doi.org/10.1037/h0032453

Shim, W. M., Alvarez, G. A., & Jiang, Y. V. (2008). Spatial separation between targets constrains maintenance of attention on multiple objects. *Psychonomic Bulletin and Review*, *15*(2), 390–397. https://doi.org/10.3758/PBR.15.2.390

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. https://doi.org/10.1016/j.cogpsych.2016.07.002

Sperling, G., & Melchner, M. J. (1978). The Attention Operating Characteristic : Examples from Visual Search. *Science*, *202*(4365), 315–318.

White, A. L., Palmer, J., & Boynton, G. M. (2019). Visual word recognition : Evidence for a serial bottleneck in lexical access.

Yantis, S. (1992). Multielement Visual Tracking : Attention and Perceptual Organization. *Cognitive Psychology*, *24*, 295–340.

CHAPTER FIVE

Dissertation Discussion

The projects in this dissertation address limitations in attention to different types of stimuli. More specifically, are selective attention and divided attention consistent with models that assume serial processing, or those that assume parallel processing? We asked this question using Gabor patches and words for selective attention, and moving objects for divided attention. The results of each experiment indicate that the mechanisms underlying attentional processes vary depending on the type of stimulus being attended.

The first experiment asked whether selective attention to simple stimuli, such as Gabor patches, is consistent with selective perception or selective decision. We tested these two possibilities using the partially valid cueing paradigm with simultaneous and sequential displays where the cued and uncued locations were presented simultaneously, or sequentially over time. The SOA was long enough to make attentional switching possible in the sequential condition. Selective perception predicts no cueing effect for the sequential condition because participants can switch attention between cued and unced locations. Selective decision predicts no difference in the cueing effect between the simultaneous and sequential conditions. The results of this experiment were consistent with selective decision, which indicates that selection of simple stimuli like Gabors is consistent with selective decision and a parallel process.

The findings of the first experiment are consistent with prior work testing parallel decision models of selective attention (Kinchla, Chen, & Evert, 1995; Shimozaki, Schoonveld, & Eckstein, 2012), but were not consistent with other work using an RSVP sequence and varied SOAs (Correa, Lupiáñez, Madrid, & Tudela, 2006; Coull & Nobre, 1998). One potential reason for the differences in results here and in prior work is temporal uncertainty. In both selective attention studies presented in this dissertation, auditory tones were used to reduce temporal uncertainty. Pilot testing for my selective attention task using Gabors found that tones were needed to reduce changes in performance as a function of SOA. Additionally, some research has found that temporal uncertainty can decrease cueing effects (Gould, Wolfgang, & Smith, 2007).

The second experiment asked whether selective attention to more complex stimuli, such as words, is consistent with the all-or-none serial model or the weighted parallel model. These two models were tested using a partially valid cueing task where participants categorized words. The all-or-none serial model predicts chance performance for invalid cues. The weighted integration model predicts above-chance performance for invalid cues. The results of this experiment were consistent with the all-or-none serial model, indicating that selective attention to words is consistent with a serial process. Additionally, these results are inconsistent with selective decision and support selective perception as a serial process.

The results found in the second experiment differ from those found in prior research on selective attention to words . An experiment by McCann, Folk, and Johnston(1992) was among the first to study selective attention to words using cueing. Cueing effects in McCann et al. (1992) were much smaller than those found here. In that study, the difference in reaction time for valid and invalid trials was 57 ms, and the difference in error rates was 5%. It is worth considering whether differences between the second experiment here and McCann et al. (1992) might contribute to this difference in performance for invalid cues. One difference is that performance in the second experiment in this dissertation was measured as percent correct and not reaction time. However, prior work that used a lexical decision task and a partially valid cue much like McCann et al. (1992), but measured performance as percent correct, found cueing effect that were smaller than the ones found here and invalid performance well above chance (Cristescu & Nobre, 2008; Ducrot & Grainger, 2007).

Another difference is that McCann et al. (1992) presented a single target stimulus with no distractors during each trial, whereas in the second experiment in this dissertation, both a target and a distractor were presented within a trial. It is possible that the presence of a distractor stimulus might make it more difficult to select the target than if the target were the only stimulus presented. Cristecu and Nobre (2008) conducted a partially valid cueing experiment using an unmasked lexical decision task with a distractor presented at the other location. Invalid performance was at 71% correct, which is higher than the chance-level performance found for invalid cues in the second experiment of this dissertation. It is therefore unlikely that the invalid performance found here can be explained by the distractors alone.

Possibly the most important difference between the word cueing experiment in the current paper and that in McCann et al. (1992) is the use of post-masks. Post-masks are frequently used to prevent attentional switching and test the all-or-none serial model(Shiu & Pashler, 1994; see Smith, 2000 for a review; White, Palmer, & Boynton, 2018, 2020). Without a post-mask, participants can switch attention between the cued and uncued locations, allowing them to select words at both locations. We argue that the difference in invalid performance found here and in McCann et al. (1992) and other word cueing studies, is due to our use of post-masks. Because prior work with words did not use a post-masks, participants could have switched attention between cued and uncued locations.

The all-or-none serial result found in the second experiment of this dissertation is consistent with prior research in divided attention to masked words (White, Palmer, & Boynton, 2018; 2020) as well as prior research using the redundant target paradigm (Mullin & Egeth, 1989). Taken with the findings of prior research using masked words, these results suggest that selective attention to words is consistent with an all-or-none serial process.

The third experiment of this dissertation asked whether divided attention to moving discs is consisted with a serial model, such as the all-or-none serial model, or a parallel resource theory model, such as the fixed capacity parallel model. These two models were tested using the MOT paradigm where participants tracked either one or two moving targets. The all-or-none serial model proposes that the selection mechanism tracks only a single target at a time. The fixed capacity parallel model proposes that the selection mechanisms can be distributed to both targets, but with less efficiency because the attentional resource is shared between the two targets.
That there was a divided attention affect with the sparse displays used in the third experiment allows us to reject crowding theory. However, the results of this experiment were not consistent with either the all-or-none serial model or the fixed capacity parallel model. Although there was a dual-task deficit that was consistent with the all-or-none serial model, there was not a negative correlation between responses on each side, which is a key prediction of the all-or-none serial model. These findings have implications for each of the hypotheses tested here.

The all-or-none serial model assumes that participants do not switch attention, but instead track only one target through the duration of the trial. It is possible that participants tried to switch attention between targets even though it was advantageous to track only one target. If participants attempt to switch attention in conditions where switching is difficult (e.g. when fast disc speeds result in short temporal frequencies), dual-task performance is predicted to decrease, but there would be no negative correlation because both sides are tracked. This scenario would predict results that are similar to those found in the third experiment, and would be consistent with the findings of Holcombe & Chen (2013).

The fixed capacity parallel model assumes that a fixed number of samples is drawn from the stimulus and shared between targets, and the resolution of each target's stimulus representation decreases with increasing numbers of targets (Miller & Bonnel, 1994; Smith, Lilburn, Corbett, Sewell, & Kyllingsbæk, 2016). However, the dual-task deficit observed here is larger than what would be predicted by the fixed capacity parallel model. This large dual-task deficit can be predicted by a simple discrete parallel model where sampled information is lost over time in a discrete manner. The loss of information is compounded at each time point, meaning that the dual-task deficit increases with increasing time. The trials used in the third experiment were 4 seconds long, and a simple discrete parallel model could predict the observed dual-task deficit and no negative correlation.

The experiments outlined in this dissertation also allow us to understand more about the locus of selective and divided attention. Specifically, where in the processing stream is unattended information filtered out? Under early selection, such as the all-or-none serial model, unattended stimuli are filtered out in perception, before later stages of processing such as decision making and memory. For example, unattended stimuli might be filtered out in early visual cortex. Under late selection, such as the fixed-capacity parallel model, information is filtered after perception, in a later stage of processing such as decision making. The results of this dissertation suggest that masked words are filtered early in perception, whereas Gabor patches are filtered in decision making. For divided attention to moving stimuli, filtering of unattended stimuli might occur at various late stages, such as in decision making, memory, and response.

Each of the three experiments in this dissertation revealed different limitations underlying attentional selection. The diversity in results allows us to reject certain simple views of attention where selection is either always serial or always parallel. The mechanisms underlying attentional selection vary depending on the stimulus. More specific theories are needed to predict which types of stimuli will be consistent with serial processing, and which will be consistent with parallel processing.

References

Correa, Á., Lupiáñez, J., Madrid, E., & Tudela, P. (2006). Temporal attention enhances early visual processing : A review and new evidence from event-related potentials. *Brain Research*, *1076*, 116–128. https://doi.org/10.1016/j.brainres.2005.11.074

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *Journal of Neuroscience*, *18*(18), 7426–7435. https://doi.org/0270-6474/98/187426-10$05.00/0

Cristescu, T. C., & Nobre, A. C. (2008). Differential Modulation of Word Recognition by Semantic and Spatial Orienting of Attention. *Journal of Cognitive Neuroscience*, *20*(5), 787–801.

Gould, I. C., Wolfgang, B. J., & Smith, P. L. (2007). Spatial uncertainty explains exogenous and endogenous attentional cuing effects in visual signal detection. *Journal of Vision*, *7*(13), 1–17. https://doi.org/10.1167/7.13.4.Introduction

Holcombe, A. O., & Chen, W. (2013). Splitting attention reduces temporal resolution from 7 Hz for tracking one object to , 3 Hz when tracking three. *Journal of Vision*, *13*, 1–19. https://doi.org/10.1167/13.1.12.Introduction

Kinchla, R. a, Chen, Z., & Evert, D. (1995). Precue effects in visual search: data or resource limited? *Perception & Psychophysics*, *57*(4), 441–450. https://doi.org/10.3758/BF03213070

McCann, R. S., Folk, C. L., & Johnston, J. C. (1992). The role of spatial attention in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1015–1029.

Miller, J., & Bonnel, A. (1994). Switching or sharing in dual-task line-length discrimination ?

Perception & Psychophysics, *56*(4), 431–446.

Mullin, P. A., & Egeth, H. E. (1989). Capacity Limitations in Visual Word Processing. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(1), 111–123.

Shimozaki, S. S., Schoonveld, W. A., & Eckstein, M. P. (2012). A unified Bayesian observer analysis for set size and cueing effects on perceptual decisions and saccades. *Journal of Vision*, *12*(6), 1–26. https://doi.org/10.1167/12.6.27.Introduction

Shiu, L., & Pashler, H. (1994). Negligible Effect of Spatial Precuing on Identification of Single Digits. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1037–1054.

Smith, P. L. (2000). Attention and Luminance Detection : Effects of Cues , Masks , and Pedestals. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(4), 1401–1420. https://doi.org/10.1037//0096-I523.26.4.H01

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. https://doi.org/10.1016/j.cogpsych.2016.07.002

White, A. L., Palmer, J., & Boynton, G. M. (2018). Evidence of Serial Processing in Visual Word Recognition. *Psychological Science*, *29*(7), 1062–1071. https://doi.org/10.1177/0956797617751898

White, A. L., Palmer, J., & Boynton, G. M. (2020). Visual word recognition : Evidence for a serial bottleneck in lexical access. *Attention, Perception, and Psychophysics*, *82*, 2000–2017.