Title
The Influence of the Attention Set on Exogenous Orienting

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN
1069-7977

Authors
Lim, Ahnate
Sinnett, Scott

Publication Date
2013

Peer reviewed

The Influence of the Attention Set on Exogenous Orienting

Ahnate Lim (ahnate@hawaii.edu)
Department of Psychology, University of Hawaii at Manoa
2530 Dole Street, Honolulu, HI 96822 USA

Scott Sinnett (ssinnett@hawaii.edu)
Department of Psychology, University of Hawaii at Manoa
2530 Dole Street, Honolulu, HI 96822 USA

Abstract

The processing of incoming sensory information relies on interacting mechanisms of sustained attention (the ability to focus attention and ignore irrelevant stimuli) and attentional capture (the ability of certain stimuli to reflexively attract one's attention). Being able to precisely predict what can capture attention when it is engaged in a demanding task is important both for understanding the nature of attention as a cognitive system and also for practical applications. While evidence indicates that exogenous capture, a mechanism previously understood to be automatic, can be eliminated while concurrently performing a demanding task, we reframe this phenomena within the theoretical framework of the 'attention set' (Most et al., 2005). Consequently, the specific prediction that cuing effects should reappear when dimensions of the cue overlap with those in the attention set (i.e., elements of the demanding task) was empirically tested and confirmed. Suggestions for further theoretical refinement and empirical testing are discussed.

Keywords: Theories of attention; attention set; exogenous cuing; orienting; attentional capture; perceptual load

Introduction

As an information processing mechanism, one of the characteristic dichotomies of attention is how it must have the capacity to be both focused and distractible at the same time. The ability to ignore irrelevant stimuli and closely attend to a specific task at hand is fundamental to goal directed behavior. Conversely, the ability to be distracted by potentially dangerous events or to be drawn towards relevant information outside the current task or area of focus can be crucial for avoiding harm and responding effectively to the environment. In fact, neurological evidence has demonstrated that these dissociable mechanisms are underpinned by distinct and interactive neural networks (Corbetta, Patel, & Shulman, 2008; Corbetta & Shulman, 2002).

Many of the attentional mechanisms that we use on a daily basis can be characterized by the way in which they enable goal-directed and top-down control of behavior. Indeed, top-down attentional control has been observed to play a role in visual search (Wolfe, 2007), endogenous (participant directed) orienting of attention to spatial locations (Posner, 1980), and even feature integration (Treisman & Gelade, 1980), to name but a few. Although the environment may contain stimuli that actively compete for and capture attention, what ultimately becomes selected for subsequent processing can be influenced by "top-down signals" that filter for behaviorally relevant objects (Desimone & Duncan, 1995).

Considering the importance of responding effectively to changes in the environment, the attentional system allows for stimuli to 'reflexively' capture attention, in a 'bottom-up' environmentally triggered fashion. It has been shown that this aspect of attention can be dependent on the particular nature of the stimuli and environmental circumstances at hand. Such factors may include the role of stimulus saliency (Jonides & Yantis, 1988) and relevancy to behavioral goals (Yantis & Egeth, 1999), for example.

Of direct interest to the current study, recent research has suggested that the reflexive orienting of attention can, at times, be interrupted when an observer is undergoing a difficult and demanding task (Santangelo, Belardinelli, & Spence, 2007). In other words, where such exogenous, or stimulus-driven, mechanisms were previously thought to be automatic (Müller & Rabbitt, 1989), more recent evidence has suggested that these effects may be eliminated in a state of focused attention. For example, several recent studies have demonstrated that requiring participants to perform a concurrent demanding task can effectively eliminate the ability of exogenous cues to capture attention (Santangelo et al., 2007; Santangelo, Finoia, Raffone, Olivetti Belardinelli, & Spence, 2008; Santangelo & Spence, 2008; Theeuwes, 1991; Yantis & Jonides, 1990). Two of these studies, for instance, employed central-arrows as 100% predictive cues in a target detection task, while also deploying abrupt visual onsets as exogenous cues (Theeuwes, 1991; Yantis & Jonides, 1990), and found that the abrupt visual onsets had no effect on performance. Yet a different study by Van der Lubbe and Postma (2005) used more eccentric (peripheral) exogenous cues and obtained evidence to the contrary, where effects of the abrupt visual onsets were observed even when attention was engaged.

Thus, there appears to be evidence indicating that under some circumstances exogenous cuing effects remain, while under others these effects are eliminated. While these experiments often employ a demanding central task, the question remains as to why opposing findings have been observed. Importantly, the answer to this question should provide insight as to how the sustained attentional system interacts with the attentional capture system.

Although it is inherently difficult to formulate theories of attention that are both broad in scope (encompassing several classes of phenomena) while concurrently possessing predictive power for detailed behavioral outcomes, there are frameworks that could provide initial scaffolding towards such comprehensive theories. One such general framework for combining aspects of both inattentional blindness (i.e., an indirect measure of sustained focus) and attentional capture has been proposed by Most, Scholl, Clifford, and Simons (2005). Central to their theoretical framework, is the idea of an 'attention set' that is synonymous with the current task at hand or state of mind. The authors postulate that this 'attention set' should be the most influential factor in determining what captures attention. Incidentally, the idea that the current frame of mind determines how attention is allocated has also been proposed by Neisser's construct of the *perceptual cycle* (1976). While Most and colleagues' formulation provides an explanatory construct for both sustained attention and attentional capture, their emphasis on the attention set can be used to infer precise predictions. Specifically, Most et al. (2005, p. 218) proposed that:

"Although some stimulus properties (e.g., uniqueness) can affect noticing, to a larger extent the unexpected objects that people consciously see depend on the ways in which they 'tune' their attention for processing of specific types of stimuli—that is, on the attentional set *that they adopt."*

Consequently, this leads to the prediction that irrelevant events that are within the same attentional set should be capable of capturing attention (i.e., irrelevant events that are similar to the targets used in a separate and attended to task), whereas events that fall outside of the attention set should go unnoticed (e.g., a gorilla walking amidst a group of people passing a basketball while counting passes, see for example Simons & Chabris, 1999).

Most and colleagues' (2005; 2001) predictions regarding the influence of the attention set were supported by a series of empirical studies centered around a paradigm in which participants counted the number of bounces of a subset of items moving within a display. Crucially, an unexpected object entered the display after several trials and detection rates for these objects were used as a measure of attentional capture. In this way, Most et al. were able to manipulate the composition of the attention set (the items moving and bouncing within the display), and observe the subsequent effects on attentional capture. Of critical importance to their theory, the findings suggest that the capture of awareness is influenced both by top-down and bottom-up interactions, where the most influential factor is ultimately the attention set adopted (although certain bottom-up factors such as stimulus salience can increase the chance that objects will be noticed). In general, when unexpected items possessed features that overlapped with those in the attentional set, participants consistently noticed them, whereas when the items were outside the attention set, participants rarely noticed them. Bearing in mind that Most and colleagues' (2005; 2001) experiments were adaptations of an inattentional blindness paradigm where participants were tested on their awareness and processing of an unexpected event, the question remains as to whether the same predictions would generalize to a different task setting where attention is focused on a central area (rather than across the experimental display), and attentional capture is measured through exogenous cuing rather than conscious detection.

To recall a related example that was mentioned earlier in more detail, Santangelo and colleagues devised a paradigm involving both a demanding central task and an exogenously cued target detection task, and found that exogenous orienting does not capture attention in a mandatory fashion when undergoing a demanding central task (see Santangelo et al., 2007; Santangelo & Spence, 2007, 2008). That is, when one's attention is engaged in performing a perceptually or attentionally demanding task, the automatic effects of exogenous cues seemingly disappear. This finding is especially important considering that previous accounts of exogenous cuing suggest that peripheral cues automatically capture attention (Jonides, 1981; Müller & Rabbitt, 1989; Van der Lubbe & Postma, 2005).

While it is possible that the elimination of the cuing effect could be related to an increase in perceptual load and a concomitant reduction in available attentional resources, as suggested by Santangelo et al. (2007), Most et al.'s (2005) theoretical framework could equally predict the same result. That is, Most et al. would predict that the elimination of the cuing effect would be related to the fact that the peripheral cues were not contained in the 'attention set' (i.e., the cue was not a part of, nor was it related to, anything in the central task). This was precisely the case in the paradigm used by Santagelo and colleagues (2007; 2008). Specifically, participants were required to detect a number amongst a rapid serial presentation of letters and numbers, while the peripheral cue was a geometric shape (i.e., not a letter or number). Adopting Most et al.'s logic, the peripheral cue was task irrelevant and not related to anything in the attention set (letters or numbers), therefore it is not surprising that it failed to capture attention. Accordingly, one can predict that if the irrelevant peripheral cues were to be manipulated such that they overlapped with the current attention set (i.e., the peripheral cues and central targets come from the same category or share the same features), they should successfully capture attention despite being completely irrelevant to the task at hand. In the present study, our goal was to investigate the relationship between central processing and the peripheral capture of attention (exogenous orienting), and how this relationship is mediated by the attention set.

Using a within subjects design, participants performed a difficult central task requiring them to detect numbers that were presented within a stream of rapidly presented letters. On a subset of trials participants responded to the location of a peripherally presented target (above or below) that was orthogonally cued. Critically, we presented different types of peripheral cues to each participant, such that the cue was either of the same content as the central task or different. Note that the cue itself was completely irrelevant to the task and in theory would be outside of the attention set if it did not share any stimulus characteristics with items in the attention set (i.e., the central task in this case). If Most et al.'s (2005) prediction holds, exogenous cuing effects should be eliminated when peripheral cues are different from stimuli in the central stream. However, if peripheral cues are related to (or were even subsets of) the central task, then an exogenous cuing effect (i.e., attentional capture) should emerge.

Methods

Participants

Twenty-three participants (mean age = 22 ± 4; 13 females) were recruited from undergraduate courses at the University of Hawaii at Manoa, and offered course credit for their participation. All participants were naïve as to the purpose of the experiment and had normal or corrected to normal vision. Ethical approval was obtained from the University's Committee on Human Subjects.

Stimuli

All stimuli were presented on a 20", iMac using Bootcamp and DMDX software (Forster & Forster, 2003). Observers sat approximately 60 cm from the display. Stimuli in the central rapid serial visual presentation (RSVP) stream was constructed from randomly chosen non-repeated letters (11 selected from set of 17: B, C, D, E, F, J, K, L, M, N, P, R, S, T, Y, X, Z), each presented for 100 ms with a 16.7 ms inter-stimulus interval (ISI). For digit detection trials, numbers were selected from a set of six: 2, 3, 4, 5, 6, 9. Visual targets were black circles (subtending 2°) and cues were either black rectangles (2.5° x 1.7°) or numbers of comparable size (i.e., outside or in the attention set of the primary task, respectively; see Figures 1 and 2). Aside from the use of number cues on half of the trials, all stimuli, presentation times, and counterbalancing were constructed to be similar to the unimodal visual condition used in Santangelo et al.'s experiment (2007).

Procedure

All participants were presented with written instructions for the task on the computer screen. Next they were presented with practice trials and given accuracy and reaction time feedback after the end of each trial. The participants had the option of repeating the instructions, repeating the practice trials, or continuing with the experiment. The experimenter also monitored participants during the practice trials to ensure their understanding of the task.

For the actual task, participants were required to monitor the RSVP stream presented in the center of the display, and to respond to the occurrence of a numerical digit. A digit occurred on the majority of trials (67%). On the remaining trials (33%) the digit was not presented and instead, participants responded to the location of a spatial target that could have occurred in one of the four corners of the screen. A peripheral cue was presented on all trials, but was irrelevant to either task. The cue could have validly predicted the side of the spatial target or not (note, a spatial target was not present on digit trials). Responses were made using one of three keys following detection of either 1) a number, 2) an upward spatial target, or 3) a downward spatial target.

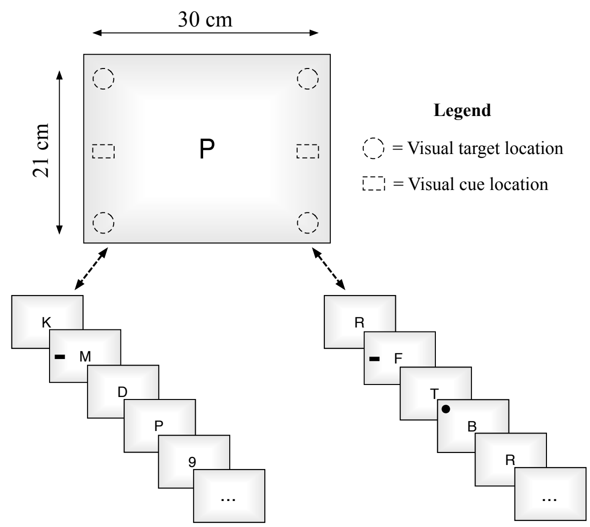

Figure 1: Schematic representation of the task. See text for details.

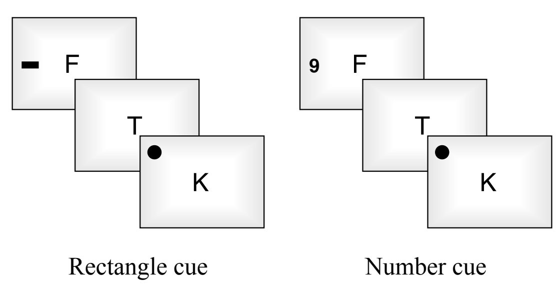

Figure 2: The two different cue types used in the task.

Each trial began with a fixation cross (1000 ms) followed by the RSVP stream of 11 items. On digit detection trials, the numbers randomly occurred in either the third, sixth, or ninth position in the stream (see also, Santangelo et al., 2007). A spatial cue was also presented on each trial (for 100 ms, identical to item duration), occurring in the third or sixth position on either the right or left side of the display equiprobably. When spatial targets occurred a number was not presented in the stream, and the spatial target appeared two positions after the cue (5^{th} or 8^{th} position). The two types of cues, rectangles or numbers, also occurred equiprobably (see Figure 2). Each experimental session consisted of 196 randomized trials, 132 of which were the digit detection task, and 64 of which were target detection (Santangelo et al., 2007). Cue combinations and trial repetitions were counterbalanced. Participants were instructed to respond as soon as targets were detected.

Results

Mean reaction times (RTs) and error rates were analyzed using three repeated measures ANOVAs (analysis of variance): one for the overall experiment and two separate ANOVAs for the digit and spatial target detection conditions. Assumptions of sphericity were tested on all analyses, with Huyn-Feldt corrections being applied to *p* values where appropriate.

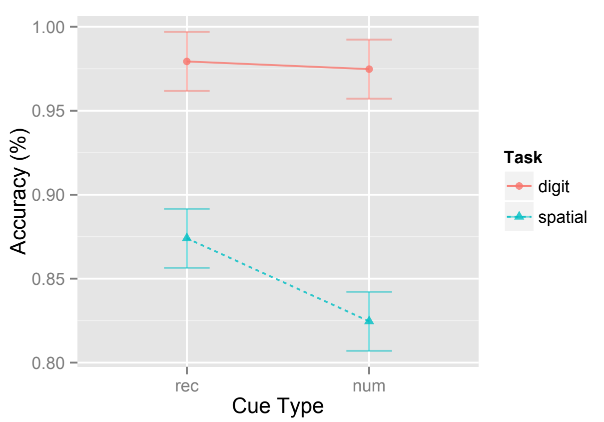

Figure 3: Mean error rates across tasks and cue types. Error bars indicate standard error values.

The first ANOVA was performed on the RT data with factors of task type (digit or target) and cue type (rectangle or number). There was no main effect of task type, $F(1, 22) = 1.2$, $p = .3$, indicating that there were no overall differences in RTs across digit ($M = 594$ ms) and target ($M = 556$ ms) detection tasks. There was, however, a main effect of cue type, $F(1, 22) = 8.4$, $p = .008$, indicating that, overall, RTs were slower when number cues ($M = 589$ ms) occurred compared to rectangle cues ($M = 563$ ms). There was no interaction between task and cue types, $F(1, 22) < 1$, *ns*, indicating no differences in RT patterns across the two tasks. In examining the error data, there was a main effect of task type, $F(1, 22) = 36.2$, $p < .001$, with lower error rates for the digit task (2%) compared to the target task (15%). Error rates were also higher on trials with number cues (10%) than on those with rectangle cues (7%), $F(1, 22) = 4.5$, $p = .045$, indicating that on average the task was more difficult when number cues were present. Notably, the analysis revealed a marginally significant interaction between task and cue types, $F(1, 22) = 3.5$, $p = .07$, indicating that number cues tended to be more distracting than rectangle cues (18% vs 13%, respectively) during spatial target detection, but not during digit detection (3% vs 2%, see Figure 3).

A second three way ANOVA performed on the digit detection condition with factors of digit position (3), cue position (2), and cue type (2) revealed that participants detected the digits significantly faster when they were presented in the ninth ($M = 455$ ms) position than when presented in the sixth ($M = 584$ ms) or third ($M = 736$ ms) positions respectively, $F(2, 44) = 25.6$, $p < .001$. Reaction times were also faster when cues were presented in the third position ($M = 574$ ms) than in the sixth position ($M = 613$ ms), $F(1, 22) = 13.0$, $p < .01$. There was also a significant interaction between digit position and cue position, $F(2, 44) = 5.4$, $p = .008$, suggesting that performance was worse when the cue occurred at the same time as the digit. Although there was no main effect in mean RTs between trials with number cues compared to trials with rectangle cues $F(1, 22) < 1$, *ns*, there was a significant three-way interaction, $F(2, 44) = 4.1$, $p = .04$, indicating a different pattern of RTs between rectangle and number cues. Specifically, when the target and cue both occur in the third position the number cue adversely affected performance whereas the rectangle cue did not (see Figure 4). No significant differences were found in the error rate data.

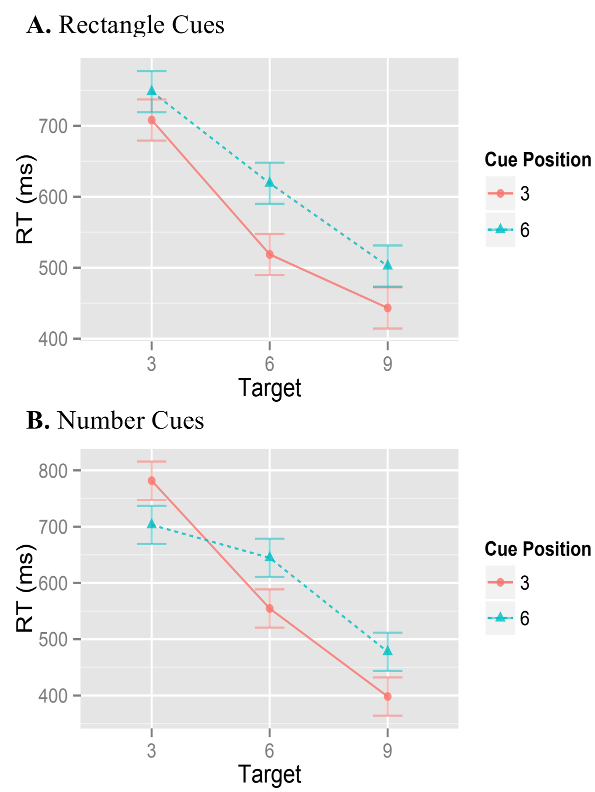

Figure 4: Interaction between target position, cue position in temporal stream, and cue type. Graph A shows trials with rectangle cues, whereas Graph B shows those with number cues.

The third, and most important, two way ANOVA was performed on the spatial target detection condition with factors of cue validity (2) and cue type (2). While there were no main effects of cue validity on reaction times, $F(1, 22) = 3.0$, $p = .1$, there was a main effect of cue type, $F(1, 22) = 11.0$, $p = .003$, where reaction times were slower when the target preceding cues were numbers (M = 570 ms) compared to when they were rectangles (M = 543 ms). Paramount to this study, the interaction between cue validity and cue type was also significant, $F(1, 22) = 4.7$, $p = .04$, indicating the presence of cuing effects for number cues on the one hand (554 ms for valid cues, and 587 ms for invalid cues), and the lack of cuing effects for rectangle cues on the other hand (539 ms for valid cues, and 547 ms for invalid cues, see Figure 5). No significant differences were found in the error rates across cue type or validity.

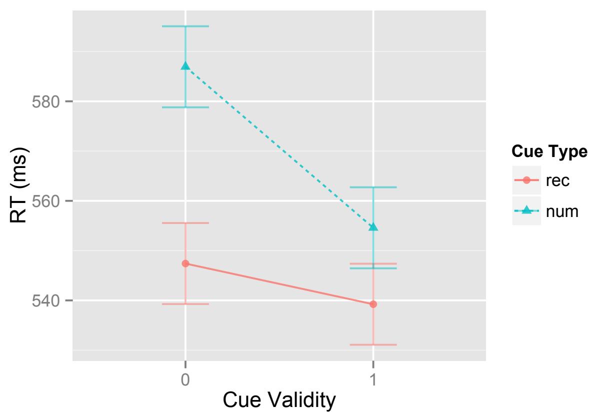

Figure 5: Interaction of cuing effects within the spatial target detection condition. Cue validity: 0 = invalid cue, and 1 = valid cue.

Discussion

The main purpose of this experiment was to test Most et al.'s (2005) theoretical framework on attention. To this end, our findings unequivocally support the prediction that irrelevant items that fall within the attentional set are capable of capturing attention while irrelevant items that fall outside of the attention set do not. That is, as was predicted, peripheral cues that were in the same category, or had overlapping features with the central task (numbers) had a cuing effect (33 ms) on spatial target detection, while items outside the category (rectangles) did not (i.e., the cuing effect was eliminated). Not only was a cueing effect observed for peripheral number distractors, but this type of cue also led to a general increase in RTs for spatial trials. This indicates that despite being irrelevant to the task, the number cues were nevertheless processed, and served as more effective distractors when compared with the rectangles.

Although valid number cues effectively captured attention, they did not facilitate overall faster reaction times. That is, the mean reaction time for trials with rectangle cues was in fact faster than for number cues, despite the lack of a cuing effect within this condition. This may possibly be due to the interaction of the number cues with the requirements of the central task. That is, any potential facilitating effects on performance of the valid number cues were probably offset by the overlap and interference with the digit detection task. The effects of this interference were also observed in the higher error rates for trials with number cues when compared to rectangle cues for spatial target detection (Figure 3).

Reaction times on the digit detection trials also point towards greater interference from the number cues. The interaction indicates that performance was worse on trials when the cue occurred at the same time as the digit, with more interference occurring from number cues when presented in the third frame. It is worth noting here that the lack of clearly distinguished differences between the effects of the rectangle and number cues on digit detection may be due to a more general distracting effect of the number cues. That is, the number cue may induce a distracting effect that generalizes beyond those particular trials to even cause the rectangle cues to become more distracting than they naturally would be. An effective way to test this theory would be to have participants also perform a task that consisted only of rectangle cues, and to then compare the pattern of results.

Aside from providing support for the theoretical position that the most influential factor for attentional capture is the 'attention set'—or the current items in focus—our findings also lend support to the notion that attention focuses on objects and features in addition to spatial location (e.g., Duncan, 1984; Egly, Driver, & Rafal, 1994; for a review, see Scholl, 2001). In refining our understanding of the attention set, it becomes imperative to more precisely define the attention set itself, for the reason that when one is engaged in a task, there are usually multiple objects or different classes of events to attend to. For example, in this experiment, we defined the attention set as being the digit detection task, due to the fact this occurred the majority (67%) of the time. The most important object in the central stream was the number, and accordingly the identity of peripheral cues was manipulated to be numbers on half of the trials.[1] The question remains however, as to what role the letters in the letter stream do play in the attention set.

Despite the fact that the letters within the RSVP stream are of a different category than the number targets, they are nevertheless processed by virtue of proximity to the number targets (both temporally and spatially) and the fact that the participants must monitor the stream in order to accurately detect the number amongst letters. We speculate that the letters should also be in the attention set, but is this assumption warranted? It is possible that the letters themselves may undergo some form on inhibition, and therefore a cue that is a letter might not capture attention. Even if this assumption were to be warranted, what would

[1] Note that on each trial the number cues were different than the numbers presented in the central RSVP task, thereby avoiding any potential confounds.

be the precise role of letters in the attention set? Would the letters be afforded equal roles to the numbers, or would their roles be lesser, perhaps even of an inhibitory nature?

Thus it is clear that although Most and colleagues' (2005) framework provides a constructive foundation to build upon, further theoretical refinement and specification through experimentation is needed. Given that many aspects of attention appear to operate in context dependent manners, exploring these contexts within the unifying framework of the attention set may prove to be an informative approach for understanding the mechanisms of attention.

References

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron, 58*(3), 306-324.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201-215.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of neuroscience, 18*(1), 193-222.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*(4), 501-517.

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123*(2), 161.

Forster, K., & Forster, J. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. Long & A. Baddeley (Eds.), *Attention and Performance IX* (pp. 187-203). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Attention, Perception, & Psychophysics, 43*(4), 346-354.

Most, S. B., Scholl, B. J., Clifford, E. R., & Simons, D. J. (2005). What you see is what you set: sustained inattentional blindness and the capture of awareness. *Psychological Review, 112*(1), 217.

Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2001). How not to be seen: The contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychological Science, 12*(1), 9-17.

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance, 15*(2), 315.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco: W. H. Freeman.

Posner, M. I. (1980). Orienting of attention. *Q J Exp Psychol, 32*(1), 3-25.

Santangelo, V., Belardinelli, M., & Spence, C. (2007). The suppression of reflexive visual and auditory orienting when attention is otherwise engaged. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 137-148.

Santangelo, V., Finoia, P., Raffone, A., Olivetti Belardinelli, M., & Spence, C. (2008). Perceptual load affects exogenous spatial orienting while working memory load does not. *Experimental brain research, 184*(3), 371-382.

Santangelo, V., & Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of Experimental Psychology: Human Perception and Performance, 33*(6), 1311-1321.

Santangelo, V., & Spence, C. (2008). Is the exogenous orienting of spatial attention truly automatic? Evidence from unimodal and multisensory studies. *Consciousness and cognition, 17*(3), 989-1015.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition, 80*(1-2), 1-46.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception-London, 28*(9), 1059-1074.

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Attention, Perception, & Psychophysics, 49*(1), 83-90.

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

Van der Lubbe, R. H. J., & Postma, A. (2005). Interruption from irrelevant auditory and visual onsets even when attention is in a focused state. *Experimental brain research, 164*(4), 464-471.

Wolfe, J. M. (2007). Guided Search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated Models of Cognitive Systems*. NY: Oxford.

Yantis, S., & Egeth, H. E. (1999). On the distinction between visual salience and stimulus-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 25*(3), 661.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance, 16*(1), 121.