Measuring Attention Control Abilities with a Gaze Following Antisaccade Paradigm

Jade Yonehiro (jade.yonehiro@asu.edu)
Arizona State University, 4701 W Thunderbird Road
Glendale, AZ 85306 USA

Nicholas D. Duran (nicholas.duran@asu.edu)
Arizona State University, 4701 W Thunderbird Road
Glendale, AZ 85306 USA

Abstract

Social gaze-following consists of both reflexive and volitional control mechanisms of saccades, similar to those evaluated in the antisaccade task. This similarity makes gaze-following an ideal medium for studying attention in a social context. The present study seeks to utilize reflexive gaze-following to develop a social paradigm for measuring attention control. We evaluate two gaze-following variations of the antisaccade task. In version 1, participants are cued with still images of a social partner looking either left or right. In version 2, participants are cued with videos of a social partner shifting their gaze to the left or right. As with the traditional antisaccade task, participants were required to look in the opposite direction of the target stimuli (i.e., gaze cues). Performance on the new gaze-following antisaccade tasks is compared to the traditional antisaccade task and the highly related ability of working memory.

Keywords: gaze-following; social cues; attention control; antisaccade; working memory

At any given moment, our environment is filled with far more information than we can observe at once. With a seemingly infinite number of incoming signals, we need some way to decide what we should pay attention to. To this end, attention control allows us to selectively attend to stimuli in the environment (Posner & Rothbart, 2007). In cognitive psychology, attention is typically studied by measuring a person's ability to orient attention "at will" in the face of a distracting stimuli (Unsworth, Schrock, & Engle, 2004). To date, the use of simple stimuli (e.g., flashes of light or basic geometric shapes) to capture attention has dominated the field of attention research; however, the generalizability of such stimuli has been the subject of some critique (Kingstone, Laidlaw, Nasiopoulos, & Risko, 2016). Joint attention, specifically the tendency to reflexively align one's attention with another person via gaze-following, may provide a unique opportunity to measure attention control in a more complex social context. Despite its potential, little is known about how joint attention abilities fit into current models of attention. The present study aims to bring together research on gaze-following and traditional models of attention control to evaluate the potential of using gaze cues as stimuli for measuring attention control.

Attention Control

Two contrasting processes drive attention control: bottom-up and top-down selection. Bottom-up or stimulus-driven selection refers to the passive and involuntary orienting of attention to salient and potentially important stimuli in the environment (Connor, Egeth, & Yantis, 2004). Top-down or goal driven selection refers to the volitional orienting of attention to stimuli that is relevant to a person's current behavior or intentions (Theeuwes, 2010). Although top-down selection is typically associated with attention control, both play important roles in the way we study attention.

Bottom-up selection is responsible for orienting attention to salient stimuli regardless of the intentions of the observer (Connor et al., 2004). For example, if there is a sudden flash of light while you are reading, you would automatically orient towards the source of the flash. This behavior has a significant survival purpose. Salient features such as stark color and geometric contrast could be a food source, while sudden movement or sounds could indicate a predator attack (Connor et al., 2004). For the modern-day human, however, salient bottom-up distractors can lead to difficulties with maintaining attention on important tasks (van Zoest & Donk, 2003).

Despite their automatic nature, bottom-up processes are not in complete control of our attention. Top-down processes allow us to orient attention "at will" to stimuli that are relevant to our current goals or behaviors (Theeuwes, 2010). Suppose the flash of light from the previous example came from an unimportant source like a camera flash. Top-down selection would allow you to ignore successive flashes and return your focus on your reading. It is generally believed that top-down selection occurs after bottom-up selection. This is because top-down selection requires recurrent feedback processes to modulate selection – a process reliant on working memory (Shipstead, Harrison, & Engle, 2015; Theeuwes, 2010).

Working Memory and Attention

Without the ability to hold our goals in mind, we would not be able to orient attention in a way that helps us achieve them. Working memory, the ability to temporarily maintain and manipulate goal-relevant information, is responsible for biasing top-down attention towards goal relevant stimuli through the maintenance of attentional priorities (Shipstead et al., 2015). Additionally, working memory is responsible for minimizing the effects of goal irrelevant stimuli, allowing us to maintain our attention and prevent disengagement (Heitz & Engle, 2007). But this relationship is not merely a one-way street. Just as attention needs working memory to help us select what to focus on, working memory needs attention to continually provide goal-relevant information and feedback (Conway, Kane, Bunting, Hambrick, Wilhelm, Engle, 2005). Because of this close relationship, attention and working memory are often studied in parallel. This has resulted in a multitude of working memory tasks, known as span tasks, which also tap into attention control abilities (Engle, 2002; Heitz & Engle, 2007; Unsworth et al., 2004).

A person's attention control abilities are usually measured with tasks that pit bottom-up and top-down selection against each other. These paradigms require a person to override a reflexive orienting response (bottom-up selection) and allocate attention to an alternative goal-related location (top-down selection via working memory goal maintenance) (Heitz & Engle, 2007; Posner & Rothbart, 2007). Working memory has been found to be vital to performance on these tasks. Individuals with high working memory ability resolve this competition quickly (Heitz & Engle, 2007; Shipstead et al., 2015; Theeuwes, 2010). On the other hand, individuals who score poorly on measures of working memory often have difficulty resisting bottom-up selection (Unsworth et al., 2004). They tend to make more errors and display slower response times on attention control tasks than those with higher working memory abilities (Conway et al., 2005; Conway, Kane, & Engle, 2003). It is our position then, that any new or modified attention task should be evaluated in relation to working memory abilities. Doing so will help elucidate the relationship between working memory and attentional control across a range of bottom-up and top-down constraints.

Social Attention

Recently, researchers have begun to question the generalizability of traditional cognitive tasks that use abstract stimuli (e.g., flashes of light or basic geometric shapes) to elicit bottom-up attention (Driver et al., 1999; Frischen, Bayliss, & Tipper, 2007; Friesen, Ristic, & Kingstone, 2004; Langton, Watt, & Bruce, 2000; Risko, Laidlaw, Freeth, Foulsmham, & Kingstone, 2012). Such stimuli are considered to be removed from the more real-world domains where attention is routinely employed, namely in social contexts. In response to this critique, many researchers have begun investigating how social cues influence the allocation of attention. Joint attention, the ability to share attention with another person, has become a popular medium for such investigations.

Joint attention, more specifically the tendency for people to involuntarily follow the gaze shifts or cues of another person, has engendered a long line of studies investigating how another person's gaze captures attention. These studies modify traditional attention paradigms (e.g., the Posner cueing task) to include some form of gaze stimuli. These stimuli range from schematic-static eyes (sketches of eyes looking left or right) to dynamic real faces (videos of real people's gaze shifts), with some even displaying various emotions. To date, most gaze-cueing research has focused on identifying whether or not various gaze-stimuli trigger reflexive bottom-up orienting. This is no small task, as even traditional stimuli range in their effectiveness. For example, a sudden onset peripheral cue, like a flash of color in your periphery will elicit reflexive orienting while centrally presented directional cues, like an arrow, do not. (Langton et al., 2000).

Researchers have repeatedly found gaze cues to reliably elicit bottom-up orienting in a way that closely resembles traditional attention cues, namely peripheral sudden onset cues (Frischen et al., 2007; Friesen et al., 2004; Risko et al., 2012). A few researchers have even found evidence that gaze-cues may be a stronger bottom-up stimulus than centrally presented directional cues. For example, Friesen et al. (2004) evaluated the bottom-up orienting strength of gaze and arrow cues by modifying the Posner cueing paradigm. They found that participants would orient in the direction of gaze, but not arrows cues, when the cues were counterpredictive to a target's location. They posited that, although both cue types can be used to direct attention, only gaze cues do so reflexively when presented centrally. These findings, and others like them (see Frischen et al., 2007 and Langton, Watt, & Bruce, 2000 for review; also, Friesen & Kingstone, 1998; Mundy & Jarrold, 2010), repeatedly demonstrate that gaze cues can be used to trigger bottom-up selection in a similar manner to traditionally used stimuli (e.g., peripheral flashes, etc.). This suggests that gaze cues are an effective medium for studying attention; however, more research is needed to evaluate how variations in gaze stimuli modulate the way people allocate their attention.

Despite robust evidence for the reliability of gaze cues to involuntarily orient attention, variations in gaze stimuli can have major impacts on this effect. Risko and colleagues' (2012) review of social stimuli demonstrated that changes in the "realness" of stimuli greatly impacts its bottom-up orienting strength. For instance, schematic faces elicit a larger orienting effect than real faces and dynamic gaze cues elicit stronger orienting responses that static cues. These findings suggest that not all gaze stimuli are created equal; however, little is known about the effect of using such stimuli for psychometric purposes. More research is needed to evaluate the merit of using gaze-stimuli to measure attention control. In addition, research on gaze cues has largely ignored the boarder literature on attention control. Critically, it has left the relationship between attention control and working memory largely unexplored. The present study aims to shed further light on these issues.

The Current Study

We aim to evaluate the potential of using gaze cues to measure attention control. We extend previous research on gaze stimuli in three ways. First, we have modified a traditional attention control task, the antisaccade, to make the bottom-up stimuli more social in nature. Specifically, we require participants to override the reflex to look in the direction of another's eye gaze and intentionally look to an alternative location. We use both still images (i.e., static stimuli) and videos (i.e., dynamic stimuli) of a real person's gaze shifts. Second, performance on the gaze-following paradigms will be directly compared to the original antisaccade task where the bottom-up stimuli are a simple flash. Third, we administer measures of working memory to probe the degree to which working memory ability supports top-down control in resisting distraction from increasingly complex and social bottom-up stimuli.

Hypotheses

Humans tend to prioritize and orient more reliably to social stimuli than abstract stimuli (Friesen et al., 2004). Furthermore, dynamic gaze stimuli have been found to elicit stronger orienting than static gaze stimuli (Risko et al., 2012). Thus, we predict that the dynamic gaze-following AST (antisaccade task) will be more difficult to perform than the static and traditional AST. We predict that accuracy rates will be lower and response times will be slower on the gaze-following AST than the traditional AST. We further predict that accuracy rates will be the lowest and response times will be the longest in the dynamic gaze-following AST.

Working memory is responsible for biasing top-down attention towards goal relevant stimuli and minimizing the effects of goal irrelevant stimuli (Heitz & Engle, 2007). As such, individual differences in working memory ability can be used to predict performance on attention control tasks (Conway et al., 2005). Individuals who score poorly on measures of working memory have more difficulty resisting bottom-up selection and tend to make more errors and display slower response times on attention control tasks (Conway et al., 2005; Conway et al., 2003; Unsworth et al., 2004). We hypothesize that working memory scores will predict performance on all three of the ASTs. Specifically, we expect to find that individuals with higher working memory scores will have higher accuracy rates and faster response times than those with lower scores.

Participants

142 undergraduate students were recruited from Arizona State University's subject pool. Five were removed for not following instructions and 13 were removed due to a computer error, resulting in a final sample of 124. There were 99 females, 24 males, and one participant who did not wish to provide a gender identification. Their mean age was 22.24 years (SD = 3.60). Participants were compensated with either a $15 gift card or credit towards course requirements.

Procedure

Participants were randomly assigned to one of two gaze-following groups: static (n = 59) or dynamic (n = 64) gaze cues. Due to concerns about practice effects in antisaccade tasks, assignment to gaze-following groups was between-subjects (Unsworth et al., 2004). After completing the gaze-following AST, participants then completed two working memory tasks (Operation Span and Symmetry Span tasks), and the traditional AST.

Tasks

Traditional Antisaccade Task In the traditional AST (Kane, Bleckley, Conway, & Engle, 2001), participants complete two consecutive trial types: pro- and anti-saccade trials. In the prosaccade trials (Figure 1a) a stimulus is flashed in the participant's peripheral vision on either side of a screen. Participants look at the side of the screen where the stimulus flashed. A target letter (P, B, or R) appears briefly on the same side as the flash and participants record which letter they saw. Prosaccade trials are easy to complete, as the tendency to look towards the flashed stimuli is reflexively driven by bottom-up selection (Unsworth et al., 2004). Researchers have demonstrated that high- and low-working memory individuals score equally well in the prosaccade trials (Conway et al., 2003; Unsworth et al., 2004).

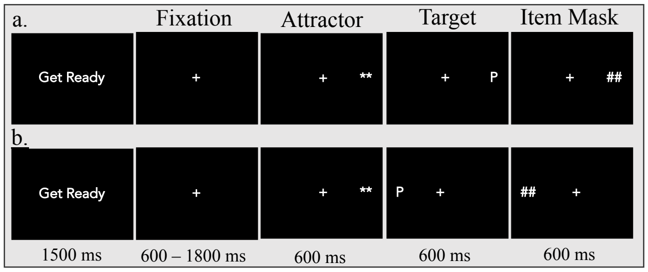

Figure 1: Procedure for gaze-following antisaccade trials.

In the antisaccade trials the same flash appears; however, the target letter appears on the opposite side as the flash (Figure 1b). Participants are instructed to suppress the automatic response to look at the flashed stimulus and instead look to the opposite side of the screen and report the letter they see. Thus, the antisaccade trials provide the competition between bottom-up and top-down selection required to measure individual differences in attention control (Unsworth et al., 2004). Individuals with low-working memory show more difficulty with the task, demonstrating slower response times and making more incorrect responses than individuals with high-working memory (Conway et al., 2003; Unsworth et al., 2004). Participants completed 70 antisaccade trials.

Gaze-cueing Antisaccade Task We developed two gaze cueing versions of the AST, which we refer to as the static-gaze and dynamic-gaze AST. Both versions were identical to the original tasks except for the stimuli used for the fixation and attractor screens. In the static-gaze version, the fixation screen (Figure 2) was replaced with a photo of a woman looking straight ahead. The attractor screen was replaced with an image of the woman looking either left or right. As with the original task, the direction of the gaze was counterbalanced and randomized across trials. In the dynamic-gaze version, the fixation screen was also replaced with a photo of a woman looking straight ahead. However, the attractor was replaced with a video of the woman's eyes shifting to the left or right. Participants completed 70 gaze-cueing antisaccade trials.

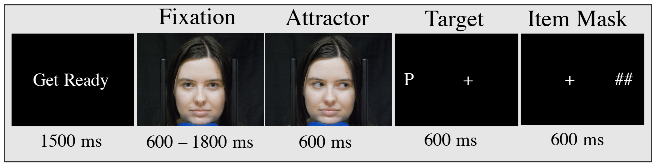

Figure 2: Procedure for gaze-following AST trials.

Operation Span Task In the Operation Span task (OSpan), participants must remember a series of letters while solving math equations (Unsworth et al., 2004). A to-be remembered letter is presented for 800 ms, followed by a math equation. Participants must identify if the solution provided for the math equation is true or false before they can move on to the next letter. Each block of trials randomly displays 3-7 to-be-remembered letters. At the end of the trial, participants must identify the letters they saw in the order in which they appeared using a 3x4 letter array (Figure 3). OSpan performance is assessed by totaling the number of letters correctly identified for trials with at least 80% accuracy on the trial math equations. Participants completed 10 OSpan blocks.

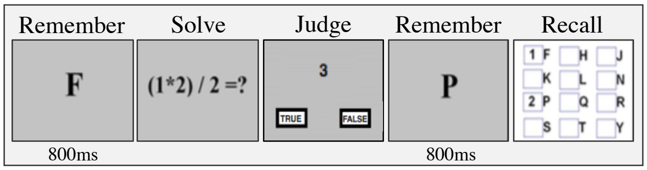

Figure 3: OSpan task example trial, image not to scale.

Symmetry Span Task In the Symmetry Span (SSpan), participants are presented with a 4 x 4 grid with a random red colored square. Next, participants must judge if a shape is symmetrical along the vertical axis. Each block of trials randomly displays 3-5 to-be-remembered red boxes with symmetry judgments made between each presentation. At the end of the trial, participants must identify the location of the red squares they saw in the order in which they appeared on a 4x4 grid (Figure 4). SSpan performance is assessed by totaling the number of letters correctly identified in order (Kane, et al., 2004). Participants completed 8 SSpan blocks.

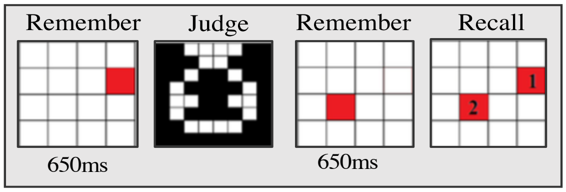

Figure 4: SSpan task example trial.

Results

Traditional and Gaze-cueing Antisaccade Tasks

AST difficulty was assessed using participant's accuracy rates and response time, such that lower accuracy rates and longer response times indicate greater task difficulty (Heitz & Engle, 2007; Shipstead et al., 2015; Theeuwes, 2010). We created two linear mixed-effects models to evaluate differences between the AST. The first model compared response times between (1) the traditional and the static AST, (2) the traditional and the dynamic AST, and (3) the static and dynamic AST relative to their respective performance on the traditional task. The second model compared accuracy rates in the same sets. For all models, a random intercept for dyad was included. Table 1 shows overall descriptives for performance on the gaze-following (static and dynamic) and traditional AST, and Table 2 provides a summary of model results.

Table 1: Observed mean and standard error for accuracy (ACC) and response times (RT) on the antisaccade tasks for the static and dynamic gaze-following groups.

Type	*Mean ACC (%)*	*SE*	*Mean RT (ms)*	*SE*
Static Group				
Traditional	58.99	0.82	747.54	6.21
Gaze-following	87.62	0.55	661.47	5.51
Dynamic Group				
Traditional	59.04	0.81	722.76	6.20
Gaze-following	75.58	0.70	727.28	5.95

Response Time The overall response time model was significantly different from the null model with only random effects (χ^2 (2,9) = 129.87, $p < .001$, $R^2 = 0.21$). Participants displayed faster response times on the static AST than the traditional AST (B = 1.69, SE=7.62, p <.001), but there was no difference in response times between the dynamic and traditional AST. Furthermore, participants in the static gaze-following group displayed faster response times than participants in the dynamic gaze-following group (B = -90.93, SE = 10.71, p <.001).

Accuracy Rates The overall accuracy model was significantly different from the null model with only random effects (χ^2 (2,8) = 1089.2, $p < .001$, $R^2 = 0.19$). Accuracy rates were higher in the static AST compared to the traditional

AST (B = 1.69, SE = 0.06, *p* <.001), and higher in the dynamic AST compared to the traditional task (B = 0.85, SE = 0.05, *p* <.001). Finally, the accuracy rates were higher on the static AST as compared to the dynamic AST (B = 0.84, SE = 0.08, *p* <.001).

Table 2: Results from mixed effects models

Response Time	*B*	*SE*	*t*	*p*
Static x Traditional	-86.95	7.62	-11.41	< .001
Dynamic x Traditional	3.98	7.52	0.53	0.60
Static x Dynamic	-90.93	10.71	-8.49	< .001
Accuracy	*B*	*SE*	*t*	*p*
Static x Traditional	1.69	0.06	26.97	< .001
Dynamic x Traditional	0.85	0.05	15.83	< .001
Static x Dynamic	0.84	0.08	10.16	< .001

*Note: ** = p< .001; * = p< .05*

Working Memory and Gaze-cueing Antisaccade Tasks

We created a composite working memory score (WM Span) by averaging the participants' normalized scores on the Ospan (M = 34.51, SE = 0.06) and Sspan (M = 17.21, SE = 0.03) tasks. Simple linear regressions were calculated to predict gaze-cueing AST response times based on WM Span. We found that WM Span did not predict response times on the static AST (*F*(1,56) = 1.12, *p* =.29). However, there was a marginally significant effect for the relationship between response time and dynamic AST (*F*(1,61) = 3.73, *p* =.06), such that higher working memory scores were associated with faster response times (see Figure 5).

For accuracy, WM Span also failed to predict performance on the static AST (*F*(1,56) = 2.29, *p* =.14). However, we found that WM Span did predict performance on the dynamic version (F(1,61) = 9.89, *p* =.002), such that higher working memory scores were associated with greater accuracy (see Figure 5).

Discussion

It has been well established that gaze cues elicit reflexive bottom-up orienting; but, unlike traditional stimuli, orienting occurs even when gaze cues are presented centrally and counterpredictive of a target's location (Friesen & Kingstone, 1998; Friesen et al., 2004). Thus, we hypothesized that the gaze-following AST would be more difficult to perform than the traditional task. We anticipated lower accuracy rates and slower response times on the gaze-following AST than the traditional AST, with performance being the lowest in the dynamic gaze-following AST. Our results were unexpected and provide interesting insight into the complex nature of gaze stimuli.

Contrary to our expectations, participants displayed faster response times and higher accuracy rates in the static gaze-following AST than the traditional AST. Additionally, working memory was unrelated to static AST performance. These results suggest that the static gaze stimuli used in this study likely elicited minimal bottom-up demands on attention control. On the other hand, the dynamic AST was more aligned with our original predictions. Although accuracy rates were higher in the dynamic AST task than the traditional task, there was no difference in response time compared to the traditional AST and working memory span was related to the dynamic AST such that individual with higher working memory spans responded faster and more accurately than those with lower spans.

One interpretation of our results is that static, and to some extent dynamic, gaze-cues of a real face do not tap into attentional capacities as strongly as traditional peripheral stimuli. However, when limiting our evaluation of performance to just gaze-cue types, the difference between static and dynamic AST performance does reveal that increasing the complexity of gaze stimuli (from static to dynamic) requires greater top-down control to override bottom-up facilitation.

The working memory results also provide some additional insight into the utility of gaze-cueing for measuring attention

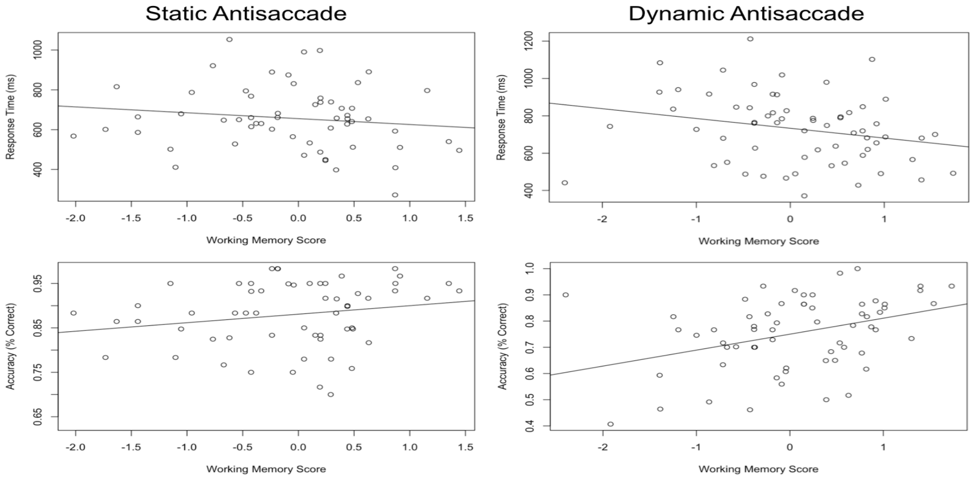

Figure 5: Response times and accuracy rates by WM Span score for the Static and Dynamic ASTs

control. Given individual differences in working memory ability have been shown to be highly related to attention control performance (Unsworth et al., 2004), it is not too surprising that there was no relationship with the static eye-gaze stimuli for this study. But as the stimuli being processed increases attentional demands, as with the dynamic gaze cues, we would expect working memory ability to predict performance. Indeed, this was the case.

Future Directions

Similar to Risko et al. (2012) we advocate for the need to systematically compare social stimuli that range in their approximation to real interaction. We also argue that it is critical to evaluate social stimuli within the framework of traditional theories and models of cognition. Although basic gaze-stimuli are thought to have a similar influence as stimuli used in traditional peripheral attention control task, when systematically compared to traditional tasks, this assumption might need further evaluation.

Acknowledgments

We would like to thank Gene Brewer and Randy Engle for providing the working memory and traditional antisaccade tasks. Funding by ASU's Professional Student Association's JumpStart Grant Program is gratefully acknowledged.

References

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: Bottom-up versus top-down. *Current Biology, 14*(19), 850-852.

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences, 7*(12), 547-552.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, Z. D., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769-786.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., Baron-Cohen, S. (1999). Gaze Perception triggers reflexive visuospatial orienting. *Visual Cognition, 6*(5), 509-540.

Engle, R. W., (2002). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science, 11*(1), 19-23.

Friesen, C.K. & Kingstone, A., (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze *Psychonomic Bulletin & Review*, 5(3),490–495.

Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology Human Perception and Performance, 30*(2), 319-329.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze Cueing of Attention: Visual Attention, Social Cognition, and Individual Differences. *Psychological Bulletin, 133*(4), 694-724.

Heitz, R. P., & Engle, R. W. (2007). Focusing the spotlight: Individual differences in visual attention control. *Journal of Experimental Psychology: General, 136*(2), 217-240.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130*(2), 169-183.

Kingstone, A., Laidlaw, K. E., Nasiopoulos, E., & Risko, E. F. (2016). Cognitive ethology and social attention. In M. Tibayrenc, & J. Francisco, On Human Nature (1 ed.). Elsevier Inc.

Langton, S. R. H., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? cues to the direction of social attention. *Trends in Cognitive Sciences, 4*(2), 50-58.

Mundy, P., & Newell, L. (2007). Attention, Joint Attention, and Social Cognition. Current Directions in Psychological Science, 16(5), 269-274.

Posner, M., & Rothbart, M. (2007). Research on Attention Networks as a Model for the Integration of Psychological Science. *Annual Review of Psychology, 58*, 1-23.

Risko, E. F., Laidlaw, K. E., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: towards an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience, 6*.

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2015). Working memory capacity and the scope and control of attention. *Attention, Perception, & Psychophysics, 77*(6), 1863-1880.

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica , 135*(2), 77-99.

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(6), 1302-1321.

Wolohan, F., & Crawford, D. (2012). The anti-orienting phenomenon revisited: Effects of gaze cues on antisaccade performance. *Experimental Brain Research, 221*(4),385-3.