The spatiotemporal dynamics of visual attention during real-world event perception

by

Ryan Ringer

M.S., Kansas State University, 2016

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Abstract

Everyday event perception requires us to perceive a nearly constant stream of dynamic information. Although we perceive these events as being continuous, there is ample evidence that we "chunk" our experiences into manageable bits (Zacks & Swallow, 2007). These chunks can occur at fine and coarse grains, with fine event segments being nested within coarse-grained segments. Individual differences in boundary detection are important predictors for subsequent memory encoding and retrieval and are relevant to both normative and pathological spectra of cognition. However, the nature of attention in relation to event structure is not yet well understood. Attention is the process which suppresses irrelevant information while facilitating the extraction of relevant information. Though attentional changes are known to occur around event boundaries, it is still not well understood *when* and *where* these changes occur. A newly developed method for measuring attention, the Gaze-Contingent Useful Field of View Task (GC-UFOV; Gaspar et al., 2016; Ringer, Throneburg, Johnson, Kramer, & Loschky, 2016; Ward et al., 2018) provides a means of measuring attention across the visual field (a) in simulated real-world environments and (b) independent of eccentricity-dependent visual constraints. To measure attention, participants performed the GC-UFOV task while watching pre-segmented videos of everyday activities (Eisenberg & Zacks, 2016; Sargent et al., 2013). Attention was probed from 4 seconds prior to 6 seconds after coarse, fine, and non-event boundaries. Afterward, participants' memories for objects and event order were tested, followed by event segmentation. Attention was predicted to either become impaired (*attentional impairment* hypothesis), or it was predicted to be broadly distributed at event boundaries and narrowed at event middles (the *ambient-to-focal shift* hypothesis). The results showed marginal evidence for

both attentional impairment and ambient-to-focal shift hypotheses, however model fitness was equal for both models. The results of this study were then used to develop a proposed program of research to further explore the nature of attention during event perception, as well as the ability of these two hypotheses to explain the relationship between attention and memory during real-world event perception.

The spatiotemporal dynamics of visual attention during real-world event perception

by

Ryan Ringer

M.S., Kansas State University, 2016

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Dr. Lester Loschky

Copyright

Abstract

Everyday event perception requires us to perceive a nearly constant stream of dynamic information. Although we perceive these events as being continuous, there is ample evidence that we "chunk" our experiences into manageable bits (Zacks & Swallow, 2007). These chunks can occur at fine and coarse grains, with fine event segments being nested within coarse-grained segments. Individual differences in boundary detection are important predictors for subsequent memory encoding and retrieval and are relevant to both normative and pathological spectra of cognition. However, the nature of attention in relation to event structure is not yet well understood. Attention is the process which suppresses irrelevant information while facilitating the extraction of relevant information. Though attentional changes are known to occur around event boundaries, it is still not well understood *when* and *where* these changes occur. A newly developed method for measuring attention, the Gaze-Contingent Useful Field of View Task (GC-UFOV; Gaspar et al., 2016; Ringer et al., 2016; Ward et al., 2018) provides a means of measuring attention across the visual field (a) in simulated real-world environments and (b) independent of eccentricity-dependent visual constraints. To measure attention, participants performed the GC-UFOV task while watching pre-segmented videos of everyday activities (Eisenberg & Zacks, 2016; Sargent et al., 2013). Attention was probed from 4 seconds prior to 6 seconds after coarse, fine, and non-event boundaries. Afterward, participants' memories for objects and event order were tested, followed by event segmentation. Attention was predicted to either become impaired (*attentional impairment* hypothesis), or it was predicted to be broadly distributed at event boundaries and narrowed at event middles (the *ambient-to-focal shift* hypothesis). The results showed marginal evidence for both attentional impairment and ambient-to-focal shift hypotheses, however model fitness was equal for both models. The results of this

study were then used to develop a proposed program of research to further explore the nature of attention during event perception, as well as the ability of these two hypotheses to explain the relationship between attention and memory during real-world event perception.

Table of Contents

List of Figures

List of Tables

Chapter 1 - Introduction

The reality which we perceive on a daily basis can be seen as somewhat of a façade, despite being rooted in the real world. While online conscious processing of the environment seems as if it is a continuous, unbroken, and information-rich experience, the observer is often missing a great deal of information due to cognitive and perceptual limitations. Despite the loss of information at both early and late stages of processing, our subjective experiences maintain a certain degree of coherence with the objective world. We are able to use attention to narrow the scope of information in the external environment, and prior experiences (i.e., memory) can be used heuristically to infer unknown information.

For instance, imagine watching a baseball game, where each moment can be broken down into a sequence of steps. As the batter steps to the plate, the pitcher prepares to throw the ball from the mound. The pitcher winds his arm back, lifts his leg, and throws the ball. At this point, several things can happen. The batter could swing or they could abstain. They could hit the ball, or they could miss and get a strike. At each event leading up to this period of uncertainty, the observer is able to guide attention to regions of the field that are relevant to the game's progression. Furthermore, the viewer's understanding of the game provides additional assistance to constrain the degree of uncertainty for the future and may guide attention prior to any changes occurring. But for now, those who are watching the game are less able to predict what will happen next.

The fact that such a complex interaction between perception, attention, and memory can occur is an incredible feat. However, the integration of old and new information allows us to efficiently comprehend a seemingly endless stream of familiar information with very little effort. Meanwhile, event representations in working memory (called an *event model*) guide the detection and extraction of new information so the observer may adapt to novel variations of a familiar situation (Kurby & Zacks, 2008). According to Event Segmentation Theory (EST; Zacks & Tversky, 2001), the moment at which a threshold for unpredictability (e.g., prediction error) is reached, an event boundary is created, marking a potential turning point in the narrative. However, the spatial and temporal changes of attention during these points of uncertainty remain an elusive and important question.

1.1 Attention

Attention is a varied subject, spanning much of the history of psychological science. It is an essential function which allows us to limit information processing to only useful information[1]. Over the extensive development of attentional research, many models have been proposed, including the attentional spotlight (Posner, 1980), the zoom-lens model (Eriksen & St. James, 1986; Eriksen & Yeh, 1985), the donut model of attention (M. M. Müller & Hübner, 2002), the multi-focal lens model (McMains & Somers, 2004; Somers & Sheremata, 2013), and the surround-suppression model of attention (de Haas, Schwarzkopf, Anderson, & Rees, 2014; Schwartz et al., 2005). A common theme of all these models of spatial attention is they assume that attention enhances a range of space (both in physical and cortical space) to be preferentially processed, while preventing capture by irrelevant or interfering information in surrounding locations. The mechanisms which guide attention can vary, however, depending on *where* and *when* useful information can provide an optimal answer on where to move the eyes next.

---

[1] While attention can be attributed to multiple elements of perception (e.g., feature and object-based attention), this literature review will focus on spatial attentional selection since this is the basis of the study in this dissertation. For an extensive review on feature and object-based attention and its relationship with spatial attention, see Treisman (2006).

The degree of attentional gain, namely the strength of the attentional boost, afforded to a target location is dependent upon a number of factors, including the strength of association between the cue and the target (Geng & Behrmann, 2005; Posner, 1980), task experience (Crundall, Underwood, & Chapman, 1999, 2002), the nature of the cue type (Posner, 1980), the amount of time between the attentional cue and the target (Nakayama & Mackeben, 1989), and the attentional capacity and processing speed of the individual (Ball, Beard, Roenker, Miller, & Griggs, 1988; Ball, Owsley, Sloane, Roenker, & Bruni, 1993). Allocating attention to a stimulus or space can result in shorter response latencies, improved contrast and orientation sensitivity, and improved memory for attended items (for review, see; Carrasco, 2011). However, cognitive and perceptual resources have limited capacity for how much information can be processed at a given moment in time (Wickens, 2001). Thus, the amount of attentional gain that can be devoted to a given stimulus is affected by dividing attention in space (Eriksen & St. James, 1986; Eriksen & Yeh, 1985; N. G. Müller, Bartelt, Donner, Villringer, & Brandt, 2003; Ringer et al., 2016; Williams, 1988, 1989) or among multiple tasks (Pashler, 1994). Additionally, dividing the central executive resources that are critical for maintaining information in working memory (e.g., inhibition) can also adversely affect attention (Logan & Gordon, 2001; Reimer, 2009).

However, this effect is indirect, and results in general interference with attention equivalently across the visual field and rather than affecting the spatial distribution of attention (Ringer et al., 2016). Nevertheless, sustained dual-task interference at the level of the central executive does seem to be more likely to result in impaired memory encoding between working memory and long-term memory (Ringer et al., 2016). Therefore, it may be the case that the broadening and narrowing of attention has a unique relationship with the amount of control being exerted over focal attention. However, it has not yet been verified if the shape of the window of attention is a direct consequence of executive control, or if it is mediated through the maintenance of working memory content.

1.2 Ambient versus Focal Processing Modes

Ambient processing typically relates to the broadened capture of general information for purposes of localization and orientation of one's self (or other objects) in the environment, whereas focal attention relates to the extraction of detailed identity information (Leibowitz & Post, 1982; Trevarthen, 1968). These contrasting processing systems are similar to the *zoom-lens* account of attention where, given a constant level of attentional resources, one can either attend broadly, with attentional resources sparsely distributed, or one can attend narrowly with attentional resources focused within a narrower range of space (Eriksen & Yeh, 1986; Eriksen & St. James, 1985). Ambient and focal processing are also mostly associated with peripheral versus central vision, respectively (Liebowitz & Post, 1982). However, these differences in attentional processing begin at the sensory level (i.e., the retina), whose central and peripheral projections extend into functional differences in the visual cortex. Visual acuity is best in central vision, with receptive field sizes in the fovea (the central 1^{o} of vision) being quite small and able to discriminate between fine details (≈ 44 cycles/degree; Loschky, McConkie, Yang, & Miller, 2005) during natural scene viewing. As retinal eccentricity increases, sensitivity to detailed information decreases exponentially until beginning to stabilize at approximately 10-15 degrees of vision, where only coarse details are identifiable (4 cycles/degree; Loschky et al., 2005).

Later in visual processing, central and peripheral vision generally project into the ventral (*what*) and dorsal (*where/how*) paths, respectively, which are two well-established neurophysiological structures that are critical for perception and working memory (Goodale & Milner, 1992). The fine-detailed ventral path and has been implicated with object recognition character and objects (Livingstone & Hubel, 1988), as well as face perception (Kanwisher, McDermott, & Chun, 1997). Conversely, peripheral vision feeds into the dorsal stream of visual processing and contributes to spatial orientation, motion detection (Livingstone & Hubel, 1988), and action perception (Goodale & Milner, 1992)[2]. However, these two separate information pathways operate together, in parallel. One important function of the central executive to connect these systems so that the spatially-dependent contextual information in the dorsal stream can constrain spatially-invariant information in the ventral stream. When frontal executive control is weakened or overridden by strong bottom-up features, the connection between dorsal and ventral stream information is weakened and potentially disrupted (Schulman et al., 2002). Thus, salient information previously outside of focal attention can be quickly encoded as the dorsal-ventral stream connection is reestablished. An important consideration is what connections are made, and when.

1.3 Coarse-to-Fine Processing in Real-World Situations

Coarse-to-fine processing entails that, during naturalistic viewing conditions, early perceptual processes sample from a broad range of space before narrowing to a more localized region in which useful information is likely to exist (Hegdé, 2008). Not surprisingly, this strategy for information extraction requires intricate coordination between ventral and dorsal stream systems optimize the deployment of attention. While the *ventral* and dorsal visual

---

[2] While there tends to be a consistent relationship between central versus peripheral vision and ventral versus dorsal visual streams, processing of global (i.e., central and peripheral) information does occur in specific areas of the ventral stream. Specifically, the parahippocampal place area processes the spatial relationships among scene global content (Aminoff, Kveraga, & Bar, 2013) while the superior medial temporal cortex (MT+) processes optic flow (Duffy & Wurz, 1997), as well as viewpoint-dependent relationships among scene objects (Shelton & McNamara, 2001).

streams seem to operate in parallel, they are by no means fully independent of each other. In fact, converging fMRI research seems to suggest that the top-down and bottom-up control of ventral stream activity is constrained by dorsal stream network activity (for review, see Corbetta & Schulman, 2002). These two pathways converge with frontocortical connections to provide an integrated, representation of the search space that includes both *what* and *where*. For instance, a number of studies have found evidence of a coarse-to-fine movement of spatial processing in natural scene recognition, with abstract "superordinate" categorization (e.g., indoor vs. outdoor) preceding basic level categorization (e.g., beach vs. desert)(Loschky & Larson, 2010). Such an integrated whole is known as the *gist*, or the earliest semantic representation of the scene as a whole that is achievable within the first fixation of an image. The gist of a scene activates a rich network of semantic information about the potential contents of the scene (and their likely locations) which can improve performance on tasks that require later focal processing, such as visual search (Eckstein, Drescher, & Shimozaki, 2006; Torralba, Oliva, Castelhano, & Henderson, 2006; Zelinsky & Schmidt, 2009) and object identification (Henderson, 1992). Although attentional cues can be utilized to improve both spatial and featural selectivity, there seems to be an earlier and stronger benefit for cuing spatial information (Liu, Stevens, & Carrasco, 2007). Thus, it is unsurprising that early scene gist information provides a cue for localizing where important information might be. For instance, when scenes were presented for a memory task, participants' first eye movements were often to regions that were rated as highly informative for the scene (Mackworth & Morandi, 1967). Later research manipulated the objects within these informative regions to be either consistent or inconsistent with the scene, finding that participants' eye movements were no more likely to land on the consistent versus inconsistent objects in the scene (Henderson, Weeks, & Hollingworth, 1999). However, when participants were asked to identify whether a specific object was present or absent, they fixated the consistent item faster than the scene inconsistent object. The authors concluded that the speed of finding these objects was a product of the relationship between the featural properties of that object and the likely location of that object within the scene (Henderson, Weeks, & Hollingworth, 1999). Thus, the gist of the scene guided attention at the earliest stages of processing when the nuances of that scene were relatively unknown.

Scene processing does not cease after the first fixation, however. Visual search often extends over several fixations and has a very stereotypical progression from ambient to focal processing. Early in scene processing, during ambient mode, fixations are short in duration while saccades are typically relatively long (Unema, Pannasch, Joos, & Velichkovsky, 2005). As the search progresses, viewers transition into the focal mode, fixation duration increases while saccade length decreases. The ambient mode of processing provides a cheap, coarse sample of the scene, whereas focal processing utilizes only a narrow subset of the scene space for more detailed (and costly) information extraction. The importance of the scene context in reducing the cost of focal processing cannot be overstated. Search behavior indicative of ambient-to-focal processing phases has been observed in macaques, but removing background scene information resulted in increases in fixation durations and decreases in saccade amplitude (Ito et al., 2017). Additionally, sudden changes to the scene in the middle of a visual search can cause temporary increases in fixation durations as the mind resets to ambient mode to reassess the global scenescape (Pannasch, Dornhoefer, Unema, & Velichkovsky, 2001). Likewise, the presence of a scene background shows some benefit to categorizing actions by an agent (Larson, Hendry, & Loschky, 2012). The fact that this ambient-to-focal pattern of information processing is so embedded in visual cognition for still images over the course of several fixations might also suggest that it is an inherent component to dynamic event perception.

During real-world event perception, a great deal of information is redundant regarding both perceptual and semantic characteristics. However, dynamic events are unique in that, while there is redundancy in some aspects of the scene from moment to moment (e.g., layout, objects, people, etc.) changes in these scene components require monitoring and updating. Furthermore, events contain a wide variety of changing types of information that may or may not be relevant to the event being perceived. Therefore, a theoretical framework for event structure will be a prerequisite for modeling changes in attention.

1.4 Event Segmentation Theory

Event Segmentation Theory (Zacks & Swallow, 2007) posits that in the midst of the constant flow of information during every-day experience, the ability to efficiently comprehend and remember our experiences hinges on breaking down the continuous stream of information into manageable “chunks” of information. These chunks can be fine-grained, meaning they are composed of small meaningful bits of information over a brief amount of time, or they can be coarse-grained, meaning that they reflect large bits of meaningful information over a longer period of time. Furthermore, these chunks are organized hierarchically, such that fine grained event segments are encapsulated in coarse grained events (Hard, Tversky, & Lang, 2006; Zacks & Tversky, 2001). By segmenting events, these bits of information can be isolated from semantically, perceptually, and temporally distinct bits of information, while also maintaining a sense of continuity from one event to the next. Thus, the uncertainty for what will happen next when a given fine-event ends is now limited to a small number of possibilities that are congruent with the coarse event. Limiting one’s expectations from one event to the next makes it possible to establish a temporal relationship between fine and coarse grained event segments, providing the means for guiding attention to relevant information (Zacks & Tversky, 2001).

The typical protocol for segmenting events (e.g., film or text) is for participants to make a response when a meaningful change in the event has ended and another has begun (Newtson, 1973) or when “a meaningful unit of event information has ended” (Eisenberg & Zacks, 2016, pg. 4). For fine event segments, participants are told to “respond to the smallest meaningful unit of change in the event has ended, and another has begun.” Likewise, for coarse segmentation, participants are told to respond to the largest units that are natural and meaningful to them (Speer, Swallow, & Zacks, 2003). Importantly, there seems to be a relationship between segmentation agreement and event memory. Participants who were more likely to segment at normed times (i.e., if they segment where others are more likely to segment) generally have better memory for those events tended to be more accurate with regard to free-recall of events, as well as new/old discrimination of still pictures of events (Sargent et al., 2013).

At the heart of event segmentation is the validity of the event model’s ability to predict event information (Reynolds, Zacks, & Braver, 2007; Zacks, Kurby, Eisenberg, & Haroutunian, 2011) which reduces as the number of aspects of the event that change increases (Huff, Meitz, & Papenmeier, 2014; Magliano, Kopp, McNerney, Radvansky, & Zacks, 2012; Zacks, Speer, Swallow, & Maley, 2010; Zwaan & Radvansky, 1998). When the assumptions or predictions for an event model are violated to a significant degree, error monitoring processes signal that the current event model is no longer adequate and must be updated to match the current situation (Kurby & Zacks, 2008; Reynolds et al., 2007; Zacks et al., 2011). Specifically, the event boundary is a time in which the gate between sensory information and the event model opens. Thus, the update to the event model will be a reflection of what is perceptually available (Kurby

& Zacks, 2008). Therefore, the perceptual information present at the boundary acts as a snapshot of the critical moment that distinguishes one event grain from the next.

The newly updated event model elicited at the boundary is not random, but dependent upon sensory information at the moment of the event boundary and prior experiences that predict what should occur next (Kurby & Zacks, 2008). Thus the breadth of predictions generated from long-term memory is limited (Hard et al., 2006) and the ability to narrow attention to quickly and efficiently represent sensory information can be improved.

1.5 Event Comprehension and Attention

Just as event models constrain information in working memory—they can also help to constrain what information is being captured from the environment. Evidence for the effects of expertise on segmentation have been shown with expert versus novice comparisons between novice and expert dancers (Blasing, 2015). The results showed that novices tended to segment more frequently and with less between-subject agreement, whereas experts tended to segment less often, but with a higher degree of agreement. Thus, if all event boundaries occur due to an error in predicting the event, then an observer with no event schema would constantly fail to predict future event information and would constantly be updating their event model (Gernsbacher, 1990).

Likewise, eye-movements of skilled athletes (i.e., tennis and soccer players) tend to fixate on the face and chest of other athletes in play, where shifts in behavior would be less frequent but more meaningful, requiring fewer shifts in attention (Piras, Lobeietti, & Squatrito, 2010). Thus, top-down influences of attention suppress the capture of repetitive (and likely less-meaningful) perceptual changes. Meanwhile, novices focused on the location and movement of the ball, again showing that novices' attention, guided by sparse representations of the event in memory, may be limited in their ability to determine what is meaningful for the event (Piras et al., 2010). Thus, a lack of event knowledge would likely produce constant updates to the event model, as is the case of novices segmenting much more frequently than experts (Blasing, 2015). However, poor executive control of working memory or attention could result in a number of problems, despite memory hierarchies being intact in long-term memory (Kurby & Zacks, 2008). Nevertheless, Event Segmentation Theory is but one framework for describing the mechanisms that allow for comprehension of dynamic events in real-time.

Hierarchical memory schemas provide a flexible template to on to which subtle variations for each unique episode can be integrated within the global event model (Hard et al., 2006). Inhibitory processes are necessary to reduce the activation of irrelevant information in working memory, while error monitoring processes are necessary to identify when an event model is no longer useful in predicting incoming event information. Likewise, selective attention is critical for inhibiting the capture of irrelevant information in working memory during narrative comprehension. The role of inhibition of sensory processing was originally suggested by the Structure Building Framework (Gernsbacher, 1990), however attention also serves a critical function in EST, where a sensory gate remains closed in the event middle, but opens at the event boundary (Kurby & Zacks, 2008; Reynolds et al., 2007). But, if sensory input is not optimally gated (either by what passes through the gate, or when the gate is opened or closed), it could translate into difficulties with maintaining information in working memory.

For instance, in one study, recognition memory and event order memory were tested between young and old adults (Kurby & Zacks, 2011). Older adults have a wealth of experience to guide attention, but younger adults tend to perform better in terms of alertness and processing speed. The results showed that although older adults showed more deficits in memory for event order compared to younger adults, fine segmentation agreement was the best predictor for event order memory regardless of age (Kurby & Zacks, 2011). Conversely, no main effect of coarse event segmentation agreement was found for event order, but it did interact with age, such that age related deficits in event comprehension may be rooted in an inability to integrate smaller (fine) units into the larger context of the coarse units of information (Kurby & Zacks, 2011). Long-term memory could be used to infer missing information (Graesser, Singer, & Trabasso, 1994), however such a heuristic approach risks creating false memories (Magliano, Kopp, Higgs, & Rapp, 2017) and in extreme cases may stagnate the ability to adapt to new situations. For accurate event comprehension, new perceptual input must be monitored to determine whether the event is compatible with the event model in working memory (Zacks & Swallow, 2007). and can be integrated with the current event structure. However, if prediction error outweighs the usefulness of the active event model in working memory, the event model must be updated (Reynolds et al., 2007).

Whether event boundaries cause a complete reset to the event model (i.e., a *global* update) or if only the event indices which changed are updated at the event boundary (i.e., an *incremental* update) is still a matter of debate. Currently, EST allows for only global updates to event models, and the only time in which perceptual input is able to be integrated with the current event model is when the sensory gate between new information and LTM is temporarily opened. However, there are several other models which allow incremental updating (e.g., Event Index Theory; Huff et al., 2014; Zwaan & Radvansky, 1998) or mapping of new information directly on to foundational event information (Gernsbacher, 1990). At first glance, the notion that event updates could be made incrementally seems a more intuitive and a more efficient process. Updating only the information that has changed during the event would maintain continuity over the course of the event and would be less taxing on brain areas responsible for processing unique event indices. Furthermore, evidence from Huff et al. (2014) has found that event updates operate on a continuum, with the number of situation changes predicting both the likelihood of segmentation at that moment in time and the number of prediction errors for what will occur next in the narrative. At the same time, requiring executive processes to maintain individualized control over multiple event domains is exceedingly complex, whereas perceiving the event holistically would be a much simpler process. Furthermore, computational models of event segmentation have found that events can spontaneously arise from simple prediction error detection to reset the event model based on perceptual input (Reynolds et al., 2007). Thus, if changes in the tuning of perceptual input at event boundaries are the critical factor behind transitioning from one event to the next, then more specific predictions about event perception can be made by integrating the rich bodies of knowledge of event perception and visual attention.

One model for event perception has begun to integrate several accounts of dynamic event perception, namely the Scene Perception and Event Comprehension Theory (SPECT; Loschky, Hutson, Smith, Smith, & Magliano, 2018). SPECT has been effective in explaining some limitations of several event perception theories, including Event Segmentation Theory (Zacks & Swallow, 2007), the Event Indexing model (Zwaan & Radvansky, 1998), and the Structure Building Framework (Gernsbacher, 1990). SPECT's advantage is its ability to connect low-level attentional selection and scene perception to higher order event and narrative comprehension. SPECT argues that *front-end* perceptual processes extract information on a fixation-by-fixation basis to generate complex information representations that can be utilized by *back-end* processes in working and long-term memory. Attention during each eye fixation modulates the way in which information is extracted from the environment. When broadly attending, attention captures the *gist* of the scene for coarse representations of information, which lays the foundation of the event (Gernsbacher, 1990; Loschky et al., 2018). Attending narrowly extracts specific features of the event, like characters or objects. Once the foundation has been laid, back-end processes maintain an ongoing representation of the event in working memory, onto which new information can be mapped. However the presence of an event model guides attention in a way that suppresses irrelevant or redundant information, limiting the degree to which working memory is affected by new information over time (Gernsbacher & Faust, 1991). SPECT, EST, Event Indexing Theory, and the Structure Building Framework all provide a mechanism for selectively integrating new information with existing event models. However, however unlike the theories listed above, SPECT makes explicit predictions for how the current event model in WM (back-end processes) influences attentional selection during eye fixations (front-end processes; Hutson, Smith, Magliano, & Loschky, 2017; Loschky et al., 2018; Loschky, Larson, Magliano, & Smith, 2015).

An important next step for event perception research is to observe and model the changes in covert attention in response to boundary and non-boundary time periods during event perception, since changes in covert attention can provide insights into internal cognitive and perceptual processing states. Specifically, it will be important to determine *where*, *how*, and *when* attention changes in response to an event boundary. Where and how attention is affected by an event boundary is of critical importance because differences in the distribution of attention over space may reflect unique effects of perceptual and cognitive loads (Ringer et al., 2016). Additionally, changes in attention that occur as a function of the temporal proximity to an event boundary may be uniquely affected by coarse and fine event boundaries. Different time courses for coarse versus fine events might also help to determine if changes in attention are discrete, as argued by EST (Kurby & Zacks, 2008) or incremental, as argued by the Event Indexing Model (Zwaan & Radvansky, 1998), or either discrete or incremental as is the case with the Structure Building Framework during mapping and shifting, respectively (Gernsbacher, 1990).

Overt attentional measures of attention during event perception have shown evidence of systematic fluctuations in visual attention as a function of temporal proximity to event boundaries (Eisenberg & Zacks, 2016). During passive viewing of events, saccade amplitudes increased prior to coarse event boundaries, and decreased following coarse event boundaries, while fixation duration decreased leading up to coarse event boundaries. This would suggest a shift from focal to ambient processing before the event boundary, meaning that attention broadens in anticipation of an event boundary and broadens immediately afterward. Therefore, the switch between broad and narrow spans of attention between respective boundary and non-boundary event components would suggest that attention is able to narrow in on specific information within the event but must sample broadly around the time at which the event model is updated. However, this effect was not consistent across experiments, with differences in groups or videos driving the lack of consistency between experiments. Nevertheless, estimating the span of covert attention through overt attentional measures comes with an important caveat: the absence of overt attention (e.g., an eye movement) is not necessarily the absence of covert attention, and the presence of an eye movement made to one location does not mean that attention was absent elsewhere. With these cautions in mind, it is now important to reevaluate the status of *covert* attention research within the context of event structure.

1.6 Measuring Covert Attention in Dynamic Environments

A number of measures of covert attention have been in use for decades—often relying on detection and discrimination tasks, where the observer is required to respond to the presence of a particular stimulus. Research from Karlene Ball (Ball et al., 1988; Ball, Edwards, & Ross, 2007; Ball et al., 1993) has utilized the effect of a foveal load on the expansion of the window of attention in time with the Useful Field of View (UFOV) task. Nevertheless, the UFOV task's design prevents it from being implemented in real-world contexts and so UFOV performance can only be correlated with attentional performance in other tasks (Ball et al., 2007).

In contrast to the UFOV's precise stimulus control, but inability to detect real-time changes in attention, the Peripheral Detection Task (PDT) was an early, but widely used method for measuring covert attention in simulated real-world environments. The PDT (Crundall et al., 1999, 2002; Jahn, Oehme, Krems, & Gelau, 2005; Patten, Kircher, Ostlund, & Nilsson, 2004) requires participants to watch a video of driving down a road, while using a brake pedal to avoid traffic hazards. To measure attention, participants must also respond when they detect a peripheral attentional probe, often times an LED or specified positioned in the margins of the video display. In one study (Crundall et al, 2002) measuring the effect of experience on the ability to detect road hazards, the results showed a decrease in the attentional probe detection rate as driving difficulty increased. However, this effect did not change as a function of eccentricity for either the experienced or novice driving group, thus ruling out tunnel vision (Crundall et al., 2002). The authors concluded that rapidly changing environments resulted in a *general interference* of attention. Although they manipulated the difficulty of the task with a visual load, the lack of a tunnel vision effect was attributed to the fact that hazards can appear at any eccentricity and therefore do not qualify as a foveal load that would bias attention to a narrow range of space. Thus, the fact that dual-task interference occurs at the perceptual stage is not sufficient to cause attentional tunneling (e.g., a sustained focal processing mode). The critical aspect (as per Williams, 1988; Williams, 1989) is that attention is preoccupied with foveal processing. Conversely, the unpredictability of the driving environment may have reduced viewers' ability to allocate attention to a particular region of the driving environment.

The above conclusion has been further solidified with later research by Ringer et al. (2016), which measured changes in attention that occur in the presence of an auditory or foveal load. This measure, called the Gaze-Contingent Useful Field of View (GC-UFOV) task required participants to discriminate the orientation of Gabor patches that were presented at different eccentricities from the fovea. Gabor patch discrimination served as a measure of attentional breadth by assuming that if a participant was able to discriminate between different Gabor orientations, then attention was likely allocated to that location. As the Gabor patches mimic the response properties of receptive fields in the visual cortex, attention increases contrast sensitivity of those receptive fields, thus improving the signal quality throughout the visual stream (Carrasco, Penpeci-Talgar, & Eckstein, 2000). By thresholding the amount of time needed to process the Gabor patches (a) at each retinal eccentricity, and (b) in single-task conditions before the onset of a filtered Gaussian noise mask, a baseline level of performance can be equalized across the visual field. Furthermore, the Gabor patches were size-scaled to become larger with increasing eccentricity to disentangle eccentricity-dependent, low-level visual resolution differences (e.g., cortical magnification; Daniel & Whitteridge, 1961; Rovamo & Virsu, 1979; Rovamo, Virsu, & Naesaenen, 1978; Virsu & Rovamo, 1979) from eccentricity-dependent changes in attention. Therefore, any changes in Gabor discrimination that exist between single and dual-task conditions can be attributed purely to changes in attention.

When the effects of a foveal (L vs. T discrimination) perceptual load and an auditory (N-Back) working memory load were compared, there was clear evidence that the foveal load produced both general interference of attention and attentional tunneling, whereas the auditory working memory load produced only general interference. An additional driving simulator study by Gaspar et al. (2016) found similar evidence of general interference on Gabor orientation discrimination with the auditory N-Back task in a driving simulator study. Therefore, the critical distinction between these two changes in attentional breadth came from different sources. For the foveal load, the participants' instructions were to prioritize the foveal L vs. T task over the peripheral Gabor task. Alternatively, the general interference effect brought on by the N-back task related to reduced control of central executive processing (e.g., inhibition; Kane, Conway, Miura, & Colflesh, 2007) over visual attention. Thus, the degree of attentional focus (i.e., whether an observer is attending narrowly or broadly) may be related to the connections between visuospatial bias (i.e., whether it is useful to attend narrowly or broadly) rather than overall attentional resources.

Since this new method of measuring covert attentional breadth has been validated for its sensitivity to perceptual and cognitive (WM) dual-task loads, it is clearly well-suited to address changes in attention that are specific only to attention, and not low-level changes in visual sensitivity due to cortical magnification. Furthermore, it is also sensitive to global fluctuations in attentional gain (i.e., global attentional facilitation or attentional impairment) that relate to available WM resources. Therefore, the GC-UFOV is well suited to measuring the cognitive and perceptual underpinnings of real-world event perception while participants view everyday events.

1.7 Hypotheses

Changes in attentional breadth can change in two primary ways: as a global attentional effect, where the overall level of attention in the visual field fluctuates, and as a change in the shape of the attentional window, by either broadening or tunneling. With the wealth of literature finding that memory for event details tends to be best around event boundaries (Huff et al., 2014; Zacks & Swallow, 2007), it seems as though it would be safe to assume that attention improves around event boundaries. Reynolds et al. (2007) originally argued that during event perception, the sensory gate remains closed, whereas at event boundaries the gate opens, allowing perceptual information to be incorporated into the event model in WM. At first glance, EST would predict attentional facilitation at event boundaries. However, this assumption has so far been empirically denied. The current empirical literature supports two categories of hypotheses regarding the effects of event boundaries on attention—a hypothesis which suggests that attention becomes impaired around event boundaries (Crundall et al., 2002; Huff, Papenmeier, & Zacks, 2012), and one in which attention shifts to an ambient mode of processing prior to the event boundary, while non-boundary periods facilitate focal visual processing (Eisenberg & Zacks, 2016; Smith, Whitwell, & Lee, 2006).

An attentional impairment effect would replicate two previous findings measuring covert attention at event boundaries from Crundall et al. (2002) and Huff et al. (2012), both of which used a Peripheral Detection Task (PDT). If attention is impaired at event boundaries, then a main effect of boundary type should find that coarse and fine boundary periods produce lower overall Gabor accuracy than non-boundaries. Furthermore, if attention becomes impaired at event boundaries, this effect should increase with temporal proximity to the boundary (Huff, Papenmeier, & Zacks, 2012). Therefore, an interaction between boundary type and the time around the event boundary would be expected, such that Gabor accuracy is worse as the temporal offset around the event boundary approaches zero. For non-boundaries, the opposite effect could occur, with Gabor accuracy improving as time around the event middle approaches zero. However, it is also possible that, if event models are stable at non-boundary periods, Gabor accuracy should be relatively flat except for the brief moment at which the event model is updated.

Conversely, if attentional breadth is controlled by the specificity of event representations in working memory, then changes in attention might be explained by a shift from ambient to focal processing modes. An interaction with eccentricity is the critical distinction between the *attentional impairment* hypothesis and the *ambient-to-focal* hypothesis, whether the interaction is with boundary type (fine and coarse vs. non-boundary) or boundary offset time. The point is that attention (operationalized as Gabor accuracy) changes as a function of eccentricity. When attention is narrowed, the Gabor accuracy slope by eccentricity would be negative (i.e., higher accuracy with smaller eccentricities) and when attention is broadened, the eccentricity slope will be neutral (or more positive, relative to a narrowed focus of attention). Changes in the size of attentional breadth would also suggest that there is likely no main effect of boundary time on Gabor accuracy since attention would be constant resource. Thus, the event model update would only affect *where* attention is allocated; not *how much* attention can be allocated.

The ambient-to-focal hypothesis would predict broadened attention in response to event boundaries, and narrowed attention at non-boundaries. Event models become less useful in predicting event information as the event boundary approaches, and thus the event model must be updated to increase accuracy in predicting future information (Huff et al., 2014; Reynolds et al., 2007). However, contrary to the *attentional impairment* hypothesis, the means of opening the sensory gate is to allow event information in from the entire visual field, rather than limiting information input to central vision. This explanation of the gating mechanism would also be in line with one of the critical aspects of SPECT (Scene Perception and Event Comprehension Theory; Loschky et al., 2018): that scene gist helps to guide the shift from one event to the next. Thus, if the event boundary represents a full reset of the event model, then broadening attention would capture the global features of the scene space to aid in that transition.

Ambient-to-focal processing of scenes suggests that the ambient mode is useful in extracting a coarse layout of the scene to provide contextual information before narrowing into specific regions of the visual field for more detailed information about objects (Pannasch, Helmert, Roth, Herbold, & Walter, 2008) and actions (Larson et al., 2012). Thus, support for the ambient-to-focal hypothesis will be found if the interaction between boundary offset time and eccentricity is such that Gabor discrimination accuracy differs least between foveal, parafoveal, and peripheral presentations when presented closest to the event boundary, while with increasing temporal distance from the boundary, Gabor discrimination decreases with increasing retinal eccentricity.

The ambient-to-focal hypothesis has been given some support through changes in overt attention (i.e., eye-movements) during dynamic event perception, but this effect seemed to differ as a function of the participants and/or event stimuli used. For instance, Eisenberg and Zacks (2016) found systematic increases in saccade amplitude and decreases in fixation duration (e.g., ambient attentional processing) at event boundaries for one set of stimuli but did not replicate it for another. The difficulty with finding consistency between these two studies illustrates the influence of stimulus-based factors in controlling eye-movements in a way that is not necessarily true for covert attention. Importantly, the GC-UFOV task is quite sensitive in distinguishing between perceptual loads and cognitive loads, when compared to eye-movements, and thus may be able to detect attentional changes where eye-movements could not (Ringer et al., 2016).

In contrast to the ambient-to-focal shift hypothesis, the *tunnel vision* hypothesis is also viable, though its support has been found in contexts outside of event perception. An interaction where information that falls further away from central vision is suppressed at the event boundary could still support an attentional impairment hypothesis. In this case, event boundaries would represent periods of uncertainty for future information (i.e., a foveal, perceptual load), and attentional impairment would be greatest in the periphery when Gabors appeared closer to the event boundary. Tunnel vision has typically been observed in reading, and is attributed to difficult, infrequent words or crowded spacing (for review, see Rayner, 1998). It has been previously argued that attention during natural scene viewing begins foveally and then expands globally within the first 150 ms (Rayner, Smith, Malcolm, & Henderson, 2009). If attention is constrained by a particularly difficult bit of information in the fovea, the viewer is unable to shift covert attention to the next peripheral target location (Nuthmann, Smith, Engbert, & Henderson, 2010). Not only does the foveal load reduce information extraction efficiency for the currently fixated target—it prevents peripheral pre-processing of the to-be-fixated target, as well. Thus, there is an additional working memory load for later information (Engbert, Nuthmann, Richter, & Kliegl, 2005).

An additional factor to consider is whether there are differences in attention between coarse and fine boundaries. Several studies have found a number of neurophysiological, cognitive, and behavioral differences between the two segmentation grains as they relate to the magnitude of responses to coarse versus fine event boundaries. Prior fMRI data has shown stronger overall BOLD (Blood-Oxygen Level Differences) signal changes for fine boundaries than coarse boundaries, while specific areas of the brain critical for error monitoring (e.g., the right anterior cingulate cortex) and spatial attentional selection (e.g., left intraparietal sulcus) showed stronger effects for coarse boundaries (Zacks et al., 2010).

Nevertheless, these effects have not been consistently found across studies, and thus may represent effects that are specific to different types of stimuli, rather than event perception as a whole. In a more recent study investigating changes in brain activity while viewing everyday tasks, stronger effects for coarse boundaries were found across the cortex, but especially in the cuneus, precuneus, and lingual gyrus (Kurby & Zacks, 2018). Another inconsistency between the Zacks et al. (2010) study and the majority of other neuroimaging results for event perception was the lack of was that motion sensitive areas of the cortex (e.g., MT+) were not found to be affected at event boundaries. For instance, one study using multiple object tracking found that, for fine boundaries, activity in bilateral MT+ increased prior to the event boundary and decreased after the boundary (Zacks, Swallow, Vettel, & McAvoy, 2006), whereas for coarse boundaries the opposite effect was found. Although this study did not measure the effects of naturalistic events, other studies have founding converging evidence on the importance of motion in the detection of event boundaries (Speer, Swallow, & Zacks, 2003; Zacks, Braver, & Sheridan, 2001). Taken together, these studies suggest that, by using videos of natural, everyday events, stronger attentional effects should occur at coarse boundaries.

Finally, it will be important to consider whether the effects of time, eccentricity, or boundary type are useful as random effects across subjects. Such random effects are useful in determining if a particular effect is consistent across all participants, or if additional flexibility is necessary to capture individual differences (Baayen, Davidson, & Bates, 2008). For instance, a random effect of boundary offset time (i.e., the temporal distance from the boundary) would suggest that there is indeed variability in terms of when people detect event boundaries. Likewise, a random effect of boundary type would indicate the degree of consistency in which individuals' attention was affected by a given boundary type. Likewise, in the event that random effect of the interaction between boundary offset time and eccentricity significantly improves model fitness, then it would suggest that there are important differences across individuals in the spatiotemporal dynamics of attention at event boundaries. Conversely, if a by-subject random effect interaction between boundary offset time and eccentricity does not improve model fitness, then it would suggest that the spatiotemporal dynamics of attention at event boundaries are relatively uniform across individuals.

Chapter 2 - Methods

2.1 Overview

The methodology for this experiment contained four phases (Fig. 1): (1) threshold estimation of Gabor target-to-mask stimulus onset asynchrony (SOA; i.e., processing time), (2) measures of Gabor discrimination accuracy while watching videos of real-world events, (3) measures of video memory, and (4) measures of event segmentation. In phase 1, each participant's SOA threshold was estimated for each of three Gabor eccentricities on a meaningless neutral image background (i.e., 1/*f* noise). The initial thresholding phase was critical for equalizing Gabor difficulty across the visual field when fully attended since phase 2 measured changes in attention due to perception of event content in the videos. In phase 2, Gabors were presented while participants viewed videos of real-world events. The times in the videos at which Gabors were presented focused on sampling around the times of coarse boundary, fine boundary, and non-boundary time periods, which were empirically derived by prior research (Eisenberg & Zacks, 2016; Sargent et al., 2013). In phase 3, participants' memory for objects and events were tested to determine if changes in visual attention predicted later memory for the contents of the videos. In phase 4, participants segmented events at coarse and fine grains to determine if changes in attention were well predicted by individual, conscious, subjective experience with the videos or if changes in attention were predicted better by the empirically derived, normative boundaries from prior research (Sargent et al., 2013).

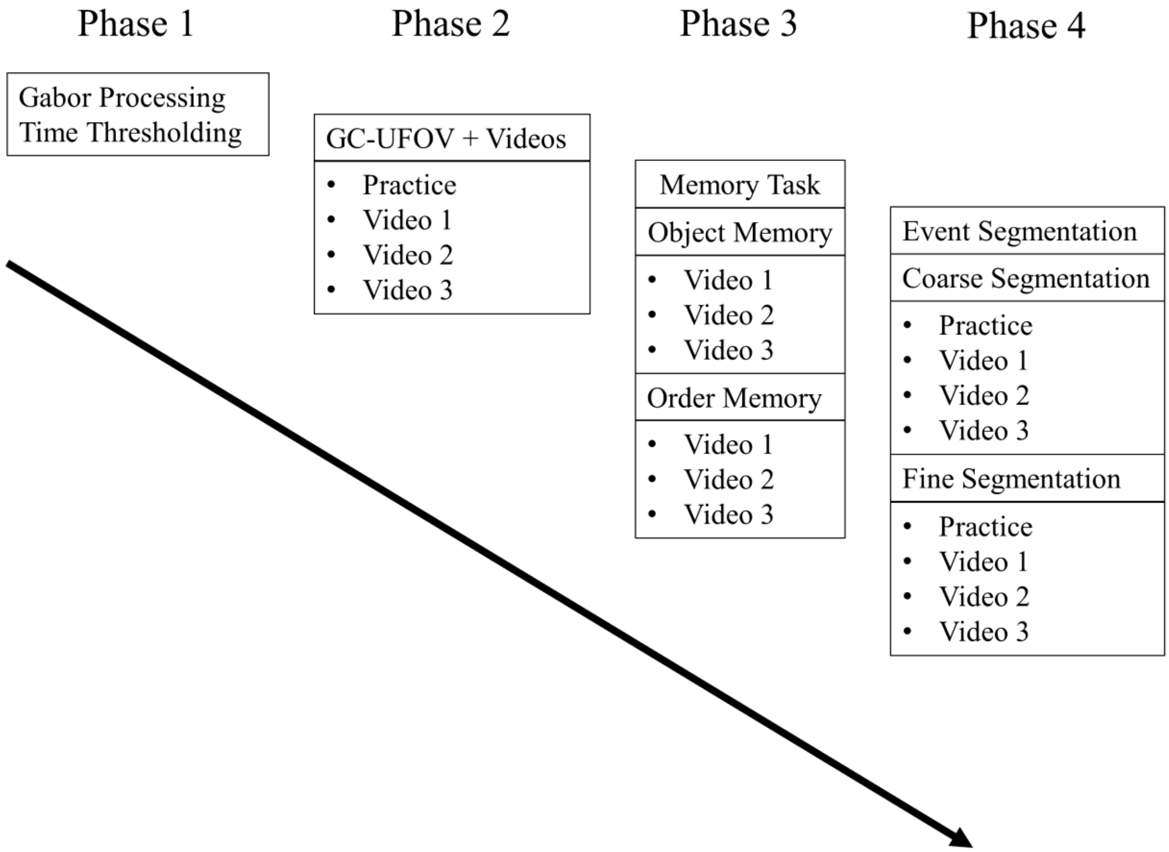

Figure 1. Experiment overview showing four separate phases: Phase 1: Gabor processing time thresholding; Phase 2: the GC-UFOV task during event perception; Phase 3: object and event memory testing; and Phase 4: event segmentation at fine and coarse boundaries. Note, however, that event segmentation grains were counterbalanced across participants.

2.2 Participants

A total of 107 undergraduate general psychology students, aged 18-25 years old (*M* = 19.5 years, *SD* = 1.33, 50 females) participated in a 2-hour study for course credit. Thirteen participants did not complete the study and were not included in the final analyses. Participants were screened for normal or corrected to normal vision ($\leq$ 20/30 Snellen acuity) using the Frieberg Acuity and Contrast Sensitivity Test (Bach, 2006). A preliminary examination of Gabor accuracy revealed that two participants' overall Gabor accuracy was at or less than chance

(< 50% correct), while another participant showed 100% accuracy. These subjects were removed from the data prior to additional analyses[3].

2.3 Materials

2.3.1 Apparatus.

The experiment was conducted on two custom PC computers, running Microsoft Windows 7, with an Intel Core i7 970 processor (3.2 GHz), 24 GB DDR3 RAM, a 256 GB Samsung 840 Pro series solid state hard drive (SSD), and a 2GB Radeon HD6950 video card. All tasks will be presented on a 19” ViewSonic CRT Monitor (model (G90fb) at 60 Hz, at a pixel resolution of 1024x768. Viewing angle was maintained at 65 cm from the monitor using a chin and forehead rest, providing a viewing angle of 31.4° x 23.89° of visual angle.

Gaze-position was measured monocularly by either an EyeLink 1000 or an EyeLink 1000+ eye-tracker (SR Research, Ontario, CA) at a sampling rate of 1000Hz. All manual responses were recorded by an 8-button Cedrus RB-844 response box having millisecond resolution.

2.3.2 Stimuli.

2.3.2.1 Gabor Stimuli. Gabor patches were generated in MatLab (version 2017a). The Gabor patches were adapted from Ringer et al. (2016). Patch sizes had been previously scaled in size using an adaptive threshold estimation algorithm that changed the size of the Gabor patches

[3] The two participants removed for below-threshold accuracy showed SOA thresholds were at floor-level (i.e., fastest possible) processing times, whereas the subject removed for having perfect accuracy showed SOAs that were at the maximum allowed processing speed (SOA = 500 ms). Thus, it is likely that thresholds for these participants were not effectively equalized across Gabor eccentricities during Phase 2 of the experiment. However, there was no meaningful effect of removing the three outliers on the subsequent results.

to adjust the difficulty of the task with unlimited processing time[4]. Thus, Gabor patch size increased with retinal eccentricity, such that the ability of participants to resolve the Gabor patches at each eccentricity would not be affected by low-level visual acuity or cortical magnification (for review, see Strasburger, Rentschler, & Jüttner, 2011). The Gabor patches were presented at 0°, 4.5°, and 9° eccentricity (Fig. 2) subtended 2.1°, 4.8°, and 8.4° degrees of visual angle, respectively. The Gabor discrimination task was a go-no-go task, in which participants had to determine whether the briefly presented Gabor patches were tilted to the right, with the target patches having been rotated 15 degrees clockwise (off-vertical) for valid trials, or counter-clockwise for catch trials. To respond to the Gabor discrimination task, participants were instructed to press a button marked "Patch" with their right thumb. To limit processing time, the Gabor patches were followed by a filtered Gaussian white noise mask (matching the mean spatial frequency, mean luminance, and mean Michelson contrast of the particular Gabors) after an SOA previously thresholded for that participant to yield 75% accuracy in Phase 1. The Gabor patches and masks were also surrounded by a neutral gray annulus (whose width = 10% of the Gabor radius) to ensure that the luminance and contrast variability of the Gabor patches relative to the scene background information did not laterally mask the patches (Ringer et al., 2016; Saylor & Olzack, 2006).

[4] The experiments in Ringer et al (2016) used Gabor patches at 0º, 3º, 6º, and 9º degrees, whereas this study presented patches at 0º, 4.5º, and 9º. Given that m-scaling follows a linear function (Strasburger et al., 2011) Gabor patch sizes for 3º and 6º were averaged to determine the 4.5º size.

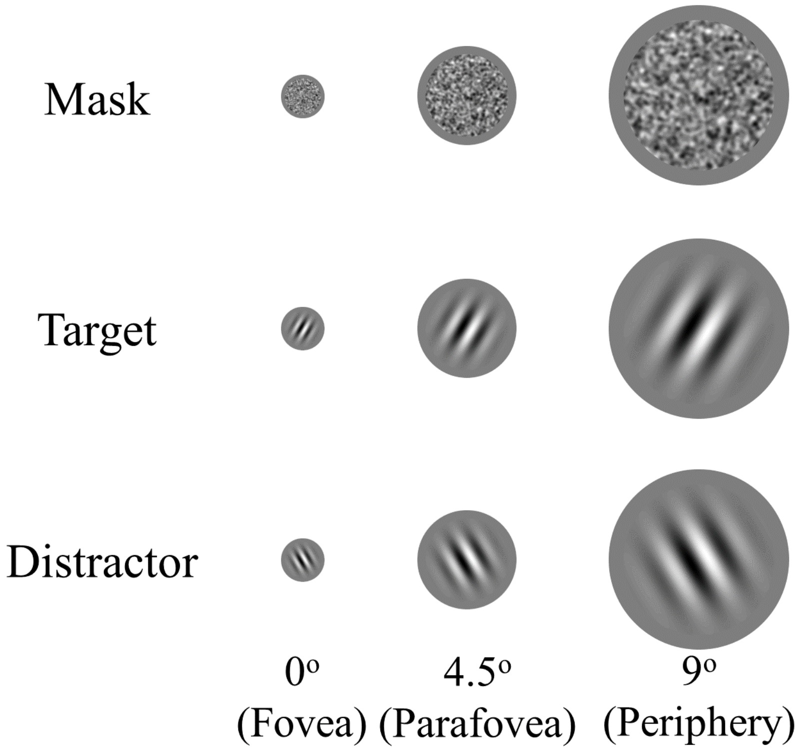

Figure 2. Gabor patches (middle and bottom rows) and masks (top row) used in the Gabor discrimination task. Participants were instructed to respond to the right-tilted (target) patches, and to not respond to the left-tilted (distractor) patches. Task difficulty was manipulated by a backward mask which covered the Gabor patches immediately following a pre-specified duration which produced 75% accuracy for the participant.

For Gabor presentations at 0°, only one patch was displayed. The Gabor patches presented at 4.5° and 9° from central vision contained four patches presented equidistant from each other. The reason for presenting multiple Gabor patches at 4.5° and 9° was that if only one patch were presented, it is possible that the Gabor patch would be presented outside of the viewable portion of the display if the participant were looking at the edge of the image. Therefore, the inclusion of four patches ensured that at least one patch would be visible during a fixation in which there was a Gabor presentation. The configuration of the patches also randomly rotated between 0° and 45° to prevent participants from anticipating the precise location of the Gabor patches across Gabor presentations.

Gabor patches were presented on two different types of image backgrounds: a 1/*f* noise background (Fig. 3) during Phase 1 SOA thresholding, and the real-world event video backgrounds during Phase 2 real-world event video viewing. During the SOA thresholding phase, the patches appeared in a neutral 1/*f* noise background while the Gabor presentation time and mask duration were manipulated using the Single Interval Adjustment Matrix (SIAM; Kaernbach, 1990). The 1/f noise background was used to emulate the power spectrum slope common to natural images, as well as the contrast sensitivity of the primary visual cortex (Field, 1987). Thus, the background of the training images was similar to the low-level features that would be present in any number of visual backgrounds without the presence of any objects on which to narrow the focus attention. To create relatively strong masking, a 2:1 mask-to-target duration ratio was used.

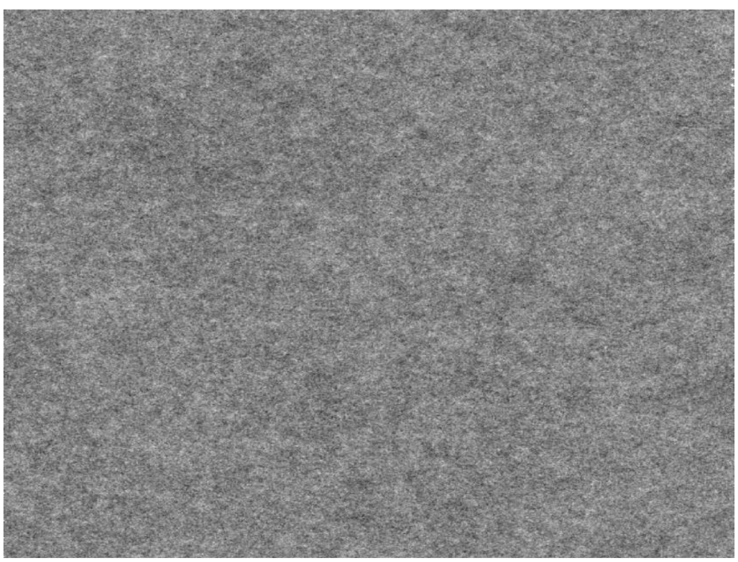

Figure 3. Sample 1/*f* noise image used as the background for Gabor SOA thresholding in Phase 1. While this image contains no recognizable objects or scene information, it's distribution of spatial frequencies is similar to that of natural scenes.

2.3.3.2 *Event Perception and Segmentation Stimuli.* While previous versions of the GC-UFOV task have utilized an artificial task to manipulate attention (e.g., the N-Back task; Gaspar et al., 2016; Ringer et al., 2014; Ringer et al., 2016; Ward et al., 2018), this experiment sought to use the structure of real world events to manipulate attention. In this study, critical moments in the event sequences (i.e., the event boundaries) were used as the attentional manipulation. In the second and fourth phases of the experiment, participants viewed one practice event video (folding laundry), followed by three experimental videos previously used by Sargent et al. (2013) and Eisenberg and Zacks (2016), each of which portrays an actor carrying out a specific task: making breakfast, setting up a dining room for a party, and transplanting garden plants into new containers. The event segmentation data from Sargent et al. (2013) were used to create triggers for the onset of Gabor patches for both fine and coarse events at time points that were above a threshold (probability density > 0.6) for an event boundary. Likewise, moments in the video which the normative event segmentation data suggested were very unlikely to be considered events (i.e., below the 0.6 threshold for an event boundary) were also included as non-event-boundary times. Events for which the inter-event interval was less than 12 seconds were not be used for the Gabor discrimination task to reduce the likelihood that a delay in a Gabor presentation would overlap in time with the next Gabor presentation.

Table 1

Number of Event Boundaries by Video

	Coarse	Fine	Non-Boundary
Breakfast	**10**	**12**	**4**
Garden	**6**	**15**	**5**
Party	**6**	**17**	**3**

Total number of event boundaries used for Gabor presentations in each video.

2.3.3.3 Object Memory Test Items. The object memory task consisted of a conceptual (verbal) recognition task, since this type of object memory has been shown to be better for objects which appear at event boundaries, and poorer for objects appearing at non-boundaries (Swallow, Zacks, & Abrams, 2009). Object memory was tested for all three videos, after which event order memory for all three videos was tested. The order for each block of memory tests matched the order in which they viewed each video. Items for the object recognition memory task consisted of names of objects which were either present or absent from each video. "Boundary" objects were objects which appeared or were manipulated at an event boundary, whereas the "non-boundary" objects were background objects that were identifiable but not relevant to the events[5]. The "absent" objects were foils, namely objects consistent with the scene or event but which were not actually present in the event videos. For example, in the making breakfast event video, "butter" and "spice rack" were both "present" objects, however only butter was coded as an event object because it was actively manipulated during an event boundary. Conversely, "whisk" was a foil object because it was not present in the video, though it was consistent with what one would expect in a video about making breakfast.

Both the present and absent objects were selected based on pilot surveys given to thirteen lab assistants who were naïve to the event videos. The objects selected for the memory task were those which matched the event or scene schemas for a portion of the surveys but were neither entirely present nor entirely absent from the survey responses. In other words, if an object was reported by all participants, it was not included as foil object. This was intended to include foil

[5] On average, boundary objects were manipulated by the actor 4.5 times, however non-boundary objects were interacted with an average of .33 times.

objects with a moderate degree of likelihood for being present in the event video. In total, there were 27 items per video, with nine objects for each of the absent, event, and non-event categories.

2.3.3.4 Event Order Memory Test Items. Event order memory was tested because previous research with Event Segmentation Theory has posited that segmentation ability is related to accurate memory for the temporal ordering of events. Therefore, event order memory was tested by providing participants with descriptions two event actions from which the participants selected the action which occurred first in the video. In the "Making Breakfast" video, for instance, participants were instructed to choose between "Putting the pan on the stove" and "Getting eggs from the refrigerator." The actions used in the task were derived from verbal recall data used by Sargent et al. (2013), which provided correlation scores between the ground-truth event order and participants' verbal recall for the actions. Thus, participants with high memory correlations were more consistent with the actual action sequences, whereas participants with low memory correlations were less accurate in recalling action order. The free-recall memories reported from the upper decile and not reported by the lower decile of the 234 participants were then used to target the action pairings that were most likely to discriminate between high and low event memory participants. 9 pairs per video were used.

2.4 Procedure

2.4.1 Precursory Information.

All participants gave their informed consent to participate in this study. Afterward, demographic information was collected, followed by Snellen acuity, and Weber contrast sensitivity.

2.4.2 Phase 1: Gabor SOA Thresholding.

In phase 1 of the experiment, Gabor thresholding consisted of 288 trials of the Gabor orientation discrimination task. Participants were first calibrated with the eye-tracker using a 9-point calibration grid. Calibrations were validated by having an average and maximum error of no more than 0.5^o and 1^o visual angle, respectively. Following calibration, participants were given instructions for the Gabor discrimination task, as well as the need to maintain fixation until the white noise mask appeared.

Trials began when participants focused their gaze on a centrally located small bull's eye pattern and pressed a button to present a fixation fail-safe dot in the same location. If gaze was still centered in the middle of the screen, the trial began. The trial began with a 1/*f* noise background image, followed by the immediate appearance of a set of Gabors at one of the three eccentricities (0^o, 4.5^o, 9^o), and then a mask. For the first 36 trials, the Gabor patches and masks appeared at a fixed duration of 150 and 300 ms, respectively, with no manipulation of the thresholding algorithm. After the first 36 trials, participants were provided accuracy feedback on their performance for each eccentricity. This first block of trials served as practice for the Gabor discrimination task and could alert the experimenter to correct any confusion on the part of the participant before the SOA was adjusted. After the practice set of 36 trials, the Gabor SOA was manipulated by a thresholding algorithm (SIAM; Kaernbach, 1991) to determine the processing time threshold needed to achieve 75% accuracy for each Gabor eccentricity. Feedback was given to the participant after every 36 trials to ensure they were not biasing attention to any particular retinal eccentricity. For instance, if a participant was scoring 100% accuracy in central vision, but was scoring 50% in the periphery, they could adjust their strategy to improve their performance in the periphery so that accuracy would be equalized across the visual field.

Trials in which the participant moved their eyes before the mask could appear (henceforth referred to as *nil patches*) did not affect the SIAM thresholding algorithm, nor were those trials included in the participants' accuracy feedback, since the duration of the retinal image activated by the Gabor could not be limited by the mask. After the 288 trials were completed, the SOA thresholds for each eccentricity were calculated and carried over to the GC-UFOV task in phase 2 of the experiment.

2.4.3 Phase 2: GC-UFOV Task and Event Perception.

In phase 2 of the experiment, participants were told that they would watch four videos of everyday events in preparation for a later memory task. They were also informed that they would be completing the same Gabor discrimination task as they did in the previous thresholding task. First, participants watched a 90 second practice video of a person folding laundry. The next three videos were recorded experimental trials in which Gabor patches were triggered to appear either before, during, or after an event boundary, or non-boundaries, namely times in which prior data from Sargent et al. (2013) suggested that participants would be least likely to detect an event boundary. However, the Gabors did not appear until the *next* fixation following the Gabor trigger (Fig. 4). After the Gabors appeared, the filtered Gaussian noise masks appeared on top of the Gabor locations to interrupt further processing of the patches. Participants had up until the next Gabor presentation to respond, thus reaction time was not measured in the results, rather processing time was limited through the backward masks.

The aim of this phase of the study was to measure changes in attention around event boundaries, and to determine if there is a distinct attentional effect for coarse versus fine boundaries and non-boundary locations. However, the GC-UFOV methodological framework recommends that attention probes be presented at the onset of a fixation, during post-saccadic suppression (Loschky & McConkie, 2002; Shioiri, 1993) to prevent attentional capture by the sudden onset of the gaze-contingent stimuli (Theeuwes, Kramer, Hahn, Irwin, & Zelinsky, 1999), and to ensure that the Gabor patches did not appear toward the end of a fixation, when an eye-movement is likely to occur (Matin, 1974). Therefore, Gabor presentations occurred at the onset of a fixation, with a maximally acceptable lag between the beginning of the fixation and the onset of the Gabors not exceeding 80 ms (Loschky & Wolverton, 2007).

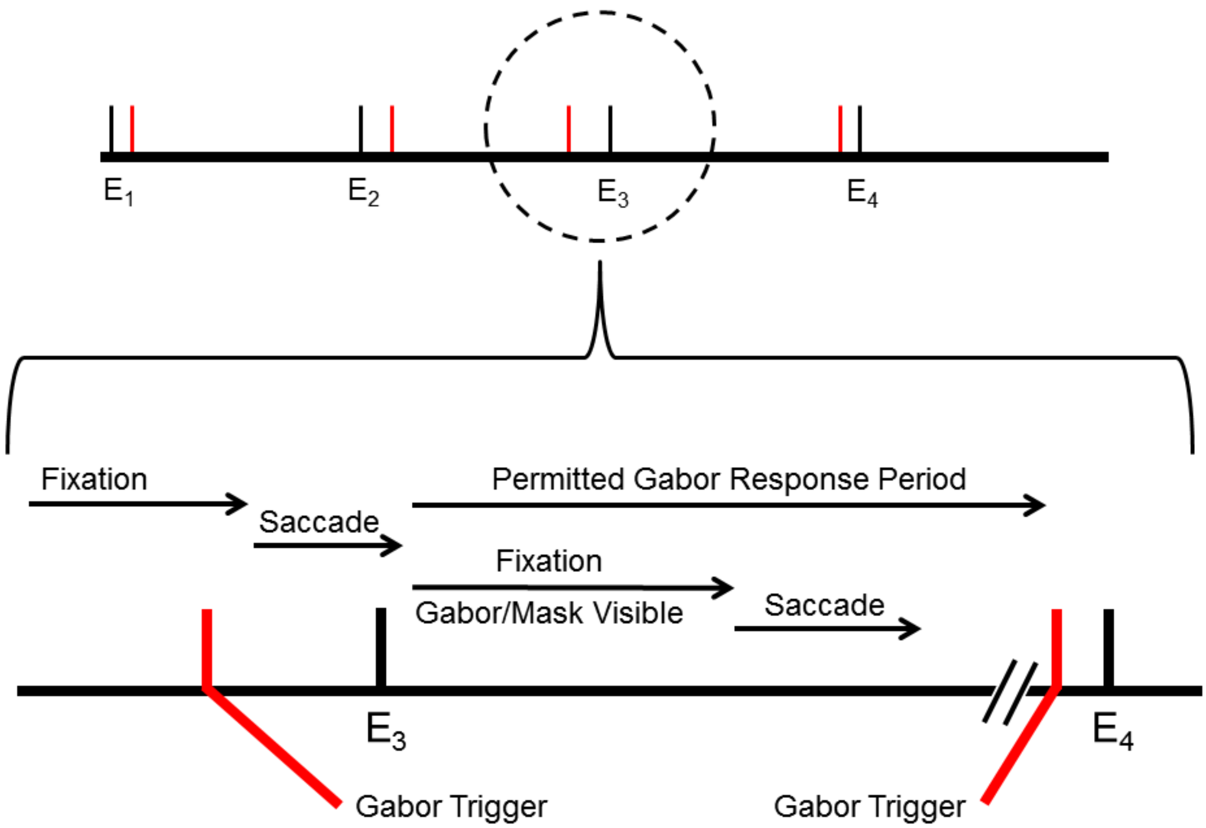

Figure 4. Gabor presentation schematic for the GC-UFOV task while viewing videos of events. Each Gabor presentation is time-locked to a time around an event boundary (i.e., E_1, E_2, E_3, E_4). Note that in this illustration, the fixation prior to the Gabor trigger extends past the trigger time, and so the presentation of the Gabor and mask must wait until the onset of the next fixation. Meanwhile, the time in which the participant may respond to a Gabor extends until near the next Gabor trigger time.

This method created a situation in which the minimum trigger-presentation offset in time between the critical video frame at which the Gabor *could* be presented and the video frame in which the Gabor was *actually* presented could be no less than zero, while the maximum trigger-presentation delay between the Gabor presentation trigger frame and the actual Gabor presentation could extend up to the next Gabor presentation frame[6]. Therefore, the non-negative lower boundary for Gabor presentation times from the Gabor trigger frame was likely to produce a certain degree of positive skew in the distribution of boundary offset times, or the difference in time between the Gabor presentation time and the event boundary. This was confirmed by pilot data (Fig. 5), with the mean shift in mean presentation times occurring 1 second after the event boundary. Thus, the Gabor trigger time offsets used to control the presentations of Gabor patches was subtracted by 1 sec from the desired target boundary times, such that the trigger times were -4 s, -2.5 s, -.5 s, 1 s, and 2.5 s from the event boundary. Additionally, combinations of boundary type, boundary offset trigger times, and Gabor eccentricity, Gabor identity (valid vs. catch trial) were counterbalanced to ensure even sampling across participants and videos.

[6] It is also important to note that the maximum lag between the critical frame in the video and the presentation of the Gabor must also include the presentation time of the Gabor, the mask, and an appropriate time to respond. Thus, the current program allowed for the maximum Gabor presentation offset to be up to two times the sum of the Gabor and mask presentation times before the critical frame of the next Gabor presentation.

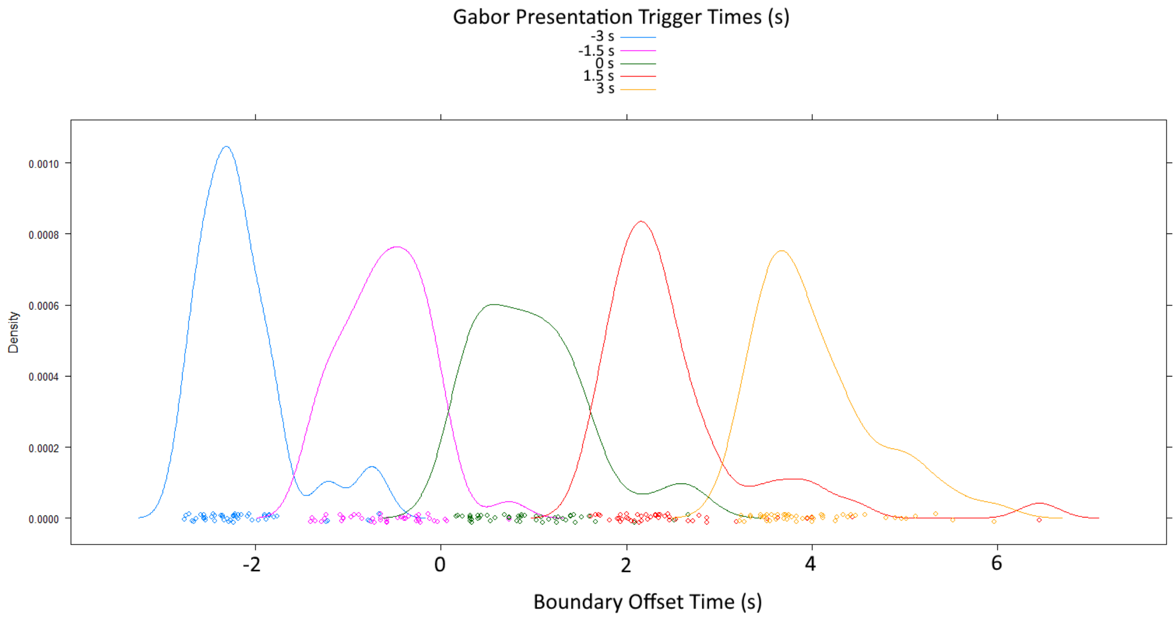

Figure 5. Pilot results from four subjects showing the distribution of actual Gabor presentation offset times in relation to the trigger times (i.e., the earliest times at which the Gabors could be presented). The positive skew in presentation distributions required that the Gabor trigger times be shifted 1 second to improve the symmetry of Gabor patch presentations before and after the event boundary during the main experiment.

2.4.4 Phase 3: Memory Tests.

In Phase 3, participants completed memory probes for object and event memory tests, each using a 2-alternative forced choice (2-AFC) recognition task. Participants first completed the object memory task for all videos, followed by the event order memory task. Memory probes were blocked by video, and order for the memory task matched the order of the videos in Phase 2. Thus, memory for each video was probed in the same order that the videos were originally viewed.

2.4.4.1 Object Memory.

Trials began when participants focused on a fixation dot and pressed a button on the response box. This caused a neutral gray screen to appear for 750 ms, followed by a screen containing a label for an object (Fig. 6). Participants responded by pressing a button to indicate whether the object was present or absent from the video, with the “present” and “absent” responses appearing on the left and right sides of the screen respectively, which corresponded with the button mapping of the response box.

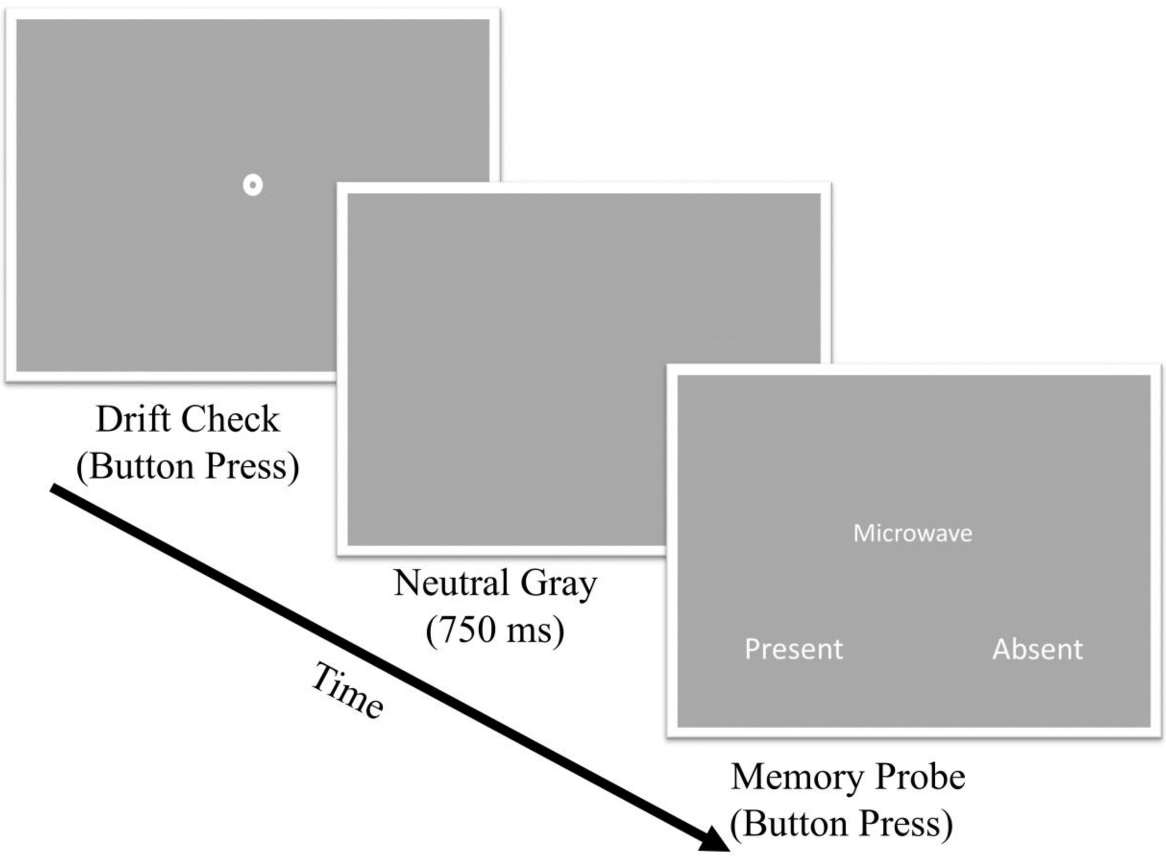

Figure 6. Trial schematic for object memory task.

2.4.4.2 Event Order Memory. Event order memory (Fig. 7) was probed with a two-alternative forced-choice task, where participants were given two different actions appearing on the right and left side of the screen. Participants responded by pressing a button to indicate which action occurred first in the video, with the button mapping for each response appearing in the lower corners of the screen.

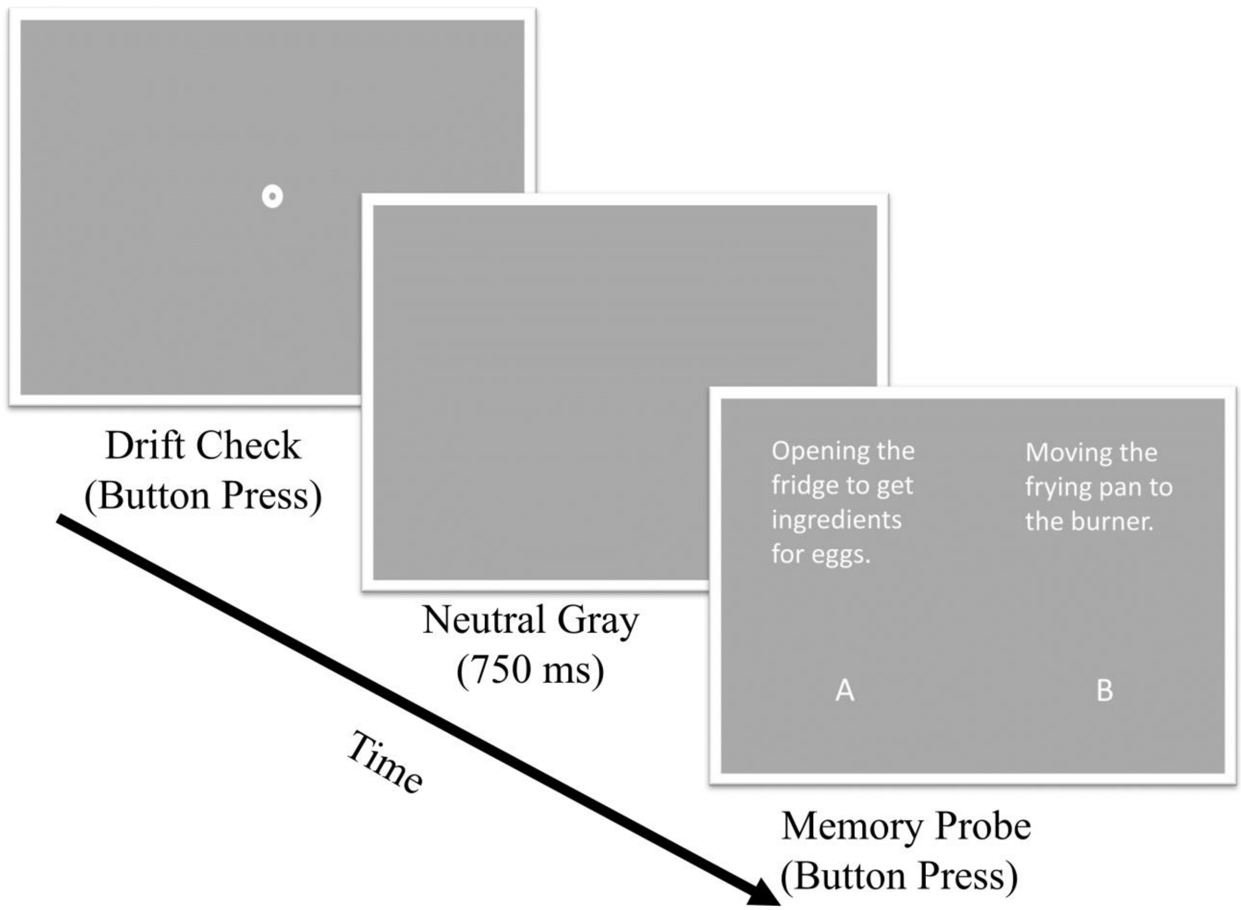

Figure 7. Trial schematic for event order memory task. Participants began the trial by focusing on a grey drift-check dot, which brings a neutral gray image, followed by the memory probes. Participants were instructed to indicate which of the two actions occurred first in the event video.

2.4.5 Phase 4: Event Segmentation. In Phase 4 of the experiment, participants watched the event videos from the GC-UFOV portion of the experiment and identified event boundaries in them. For coarse-grained segmentation, participants were instructed to "Press a button any time a large meaningful unit of information has ended," and for fine-grained segmentation, they were instructed to "Press a button any time that a small meaningful unit of information has ended" (Eisenberg & Zacks, 2016, pg. 4). Practice for event segmentation was performed on the

“Folding Laundry” video before continuing to the experimental event videos. After each video, including the practice, participants were given feedback on the number of events they detected. Trials were blocked between coarse and fine-grained event segmentation, and coarse and fine-grained segmentation was counter-balanced across participants. Video order within each coarse and fine block matched the original viewing order from Phase 2.

Chapter 3 - Results

3.1 Event Segmentation

Event segmentation results were analyzed to determine the location of normed event boundaries, as well as determining the degree of segmentation agreement among participants, namely how consistent the detection of these boundaries was across subjects (Hard et al., 2006; Zacks et al., 2010). Furthermore, it was important to determine whether the high degrees of boundary alignment between coarse and fine boundaries observed in previous studies (Hard et al., 2006; Kurby & Zacks, 2011; Zacks, Braver, et al., 2001; Zacks et al., 2010) occurred for the current experiment.

Analysis of event segmentation data was performed using a method adapted from Papenmeier (2016), which fits a probability density function to keypresses around clusters of frames from the event segmentation task to find the moments in the videos that were most likely to be considered event boundaries. Segmentation was analyzed separately for coarse and fine levels using a kernel gaussian smoothing algorithm (*R*, version 3.5.0). Consistent with Sargent et al. (2013) and Eisenberg and Zacks (2016), temporal bandwidths for the kernel smoother were set to 3 seconds, and cutoff values for the Gaussian smoother were set to a probability of .5. Furthermore, two participants did not complete all trials in the event segmentation task and were omitted from the results.

3.1.1 Segmentation Alignment. Alignment of event segmentation data between coarse and fine boundaries was assessed by creating two general linear models where the probability density for coarse event boundaries was predicted by the probability density for fine boundaries. In one model, the intercept for the probability density of a coarse boundary for each video was included as a by-item random effect, whereas a second model allowed fine boundary probability density to vary as a function of video. The results show that the second model demonstrated significantly better fit (BIC = -91484) compared to the first, simpler model (BIC = -91063). Thus, there was sufficient evidence to suggest that segmentation agreement differed as a function of video (Fig. 8). This model was then selected for fixed-effects evaluation. Fixed effects tests revealed significant agreement between coarse and fine segmentation (B = 0.576, SEM = 0.039, $t(64997)$ = 14.70, $p < .001$). When the inverse of the model was assessed (i.e., when coarse segmentation probability was used to predict fine segmentation probability), a similar result was found, with coarse segmentation probability significantly predicting fine segmentation probability (B = 0.777, SEM = 0.024, $t(64997)$ = 32.28, $p < .001$). Thus, overall fine boundary segmentation probability significantly predicted coarse segmentation probability, however coarse segmentation ability predicted fine segmentation to a stronger degree.

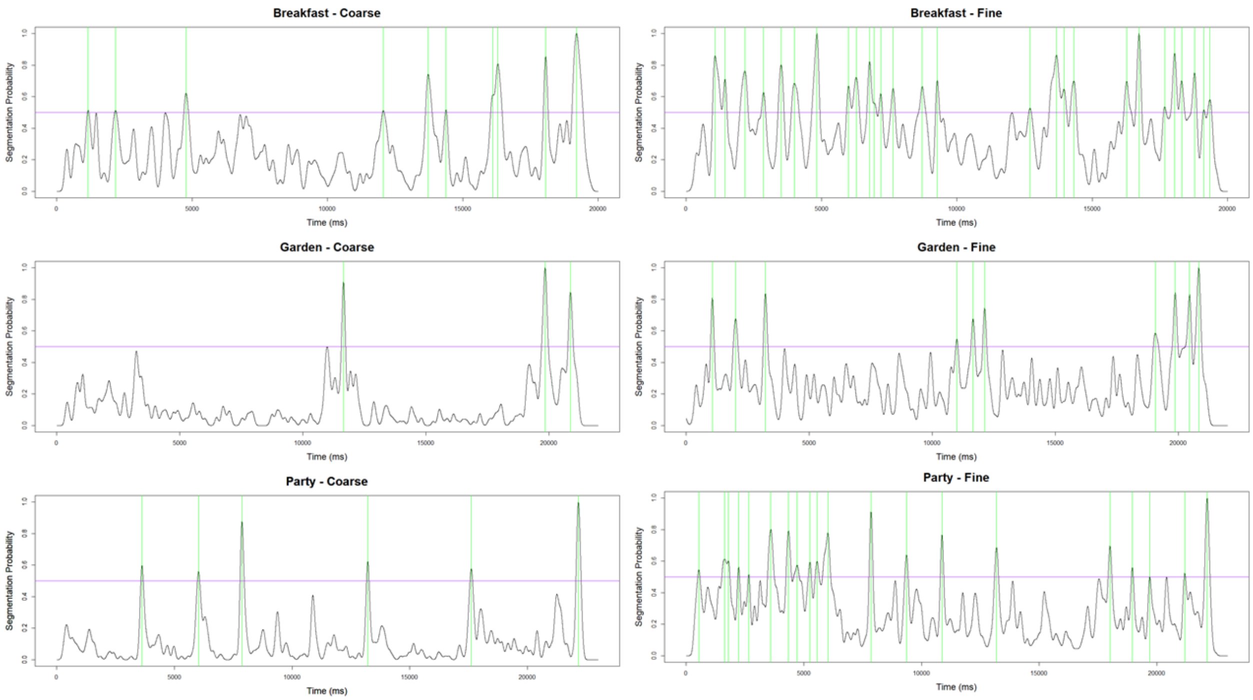

Figure 8. (Left) Probability density plots of segmentation responses for the three videos of everyday events. Green vertical bars indicate normed event boundaries, where the probability of a participant making a segmentation response exceeded .50 within a 3-second temporal bandwidth. Note the high degree of alignment between coarse and fine boundaries.

3.1.2 Segmentation Agreement. Afterward, individual subjects' response probabilities were correlated with the normed segmentation data (i.e. group average) to determine the range of individual variability in segmentation ability in the current study. This provided the basis for determining how closely an individual's segmentation performance matched with the group mean. Then, four general linear mixed models were generated with segmentation agreement being predicted by segmentation grain (i.e., fine versus coarse). The models varied in terms of their random-effects structures. Two models included only subject intercepts for segmentation agreement as the by-subject random effect, and two models included segmentation grain as a by-subject random effect slope. Additionally, half of the models varied by including video as a by-item random effect, while the other two did not include video as a by-item random effect. This test was critical for determining whether the segmentation agreement between fine and coarse segmentation differed with regard to each video, and whether such variability was sufficient to warrant additional flexibility in the overall model. Model fitness indices demonstrated that the more complex model (Grain|Subject and 1|Video) showed better model fitness (BIC = -429.23) than the next best model (1|Subject and 1|Video, BIC = -428.99). Thus, the data suggested that the trends in segmentation agreement did differ by video and that segmentation agreement among participants differed as a function of segmentation grain.

Segmentation agreement among participants was well above chance for coarse-grained segmentation ($B = 0.35$, SEM = 0.031, $t(87) = 11.42$, $p < .001$), suggesting that overall trends in segmentation agreement were more-or-less consistent across participants. However, fine-grained segmentation agreement was slightly (but significantly) lower than coarse grained segmentation ($B = -0.032$, SEM = 0.014, $t(87) = 2.20$, $p = 0.036$). This is counter to previous results which have found better segmentation agreement for fine grained events (Kurby & Zacks, 2011; Sargent et al., 2013), but suggests that there was significantly more variability in the way that participants segmented based upon smaller, stimulus-driven changes in the event.

3.1.3 Segmentation Agreement Across Event Videos. Additional analyses probing differences in coarse versus fine-grained segmentation agreement as a function of video were consistent with trends in object and event order memory. While segmentation agreement was generally lower for the *breakfast* video compared to the *garden* and *party* videos, the most interesting result is that the *breakfast* video was the only one in which coarse segmentation agreement (Mean = 0.26, 95% CI = [0.22, 0.30]) trended lower than fine segmentation agreement (Mean = 0.30, 95% CI = [0.27, 0.34]). In contrast, coarse segmentation tended to be higher than fine segmentation agreement for both the *garden* (coarse: Mean = 0.40, 95% CI[0.36, 0.44]; fine: Mean = 0.31, 95% CI[0.28, 0.36]) and *party* (coarse: Mean = 0.39, 95% CI[0.35, 0.43]; fine: Mean = 0.31, 95% CI[0.30, 0.38]) videos (Fig. 9). Thus, it is likely that the perception of event structure differs substantially between the *breakfast* video and the *garden* and *party* videos.

One source of this difference may reside in amount of motion occurring between the two videos. Motion is perhaps one of the most replicated cues for event boundaries (Kurby & Zacks, 2008; Zacks et al., 2001; Zacks et al., 2006) and facilitates segmentation agreement better at fine grains than coarse grains (Hard, Tversky, & Lang, 2006). In the *garden* and *party* videos, motion is fairly localized, and large movements typically coincided with changes in intention or goal completion. Conversely, the *breakfast* video contained a large range of motion changes within coarse grains as the actor moved from different parts of the kitchen, completing several tasks in parallel. Therefore, effects of multiple points of motion may have been less predictive of coarse boundaries for the breakfast video as they were for the *garden* and *party* videos (Zacks et al., 2010).

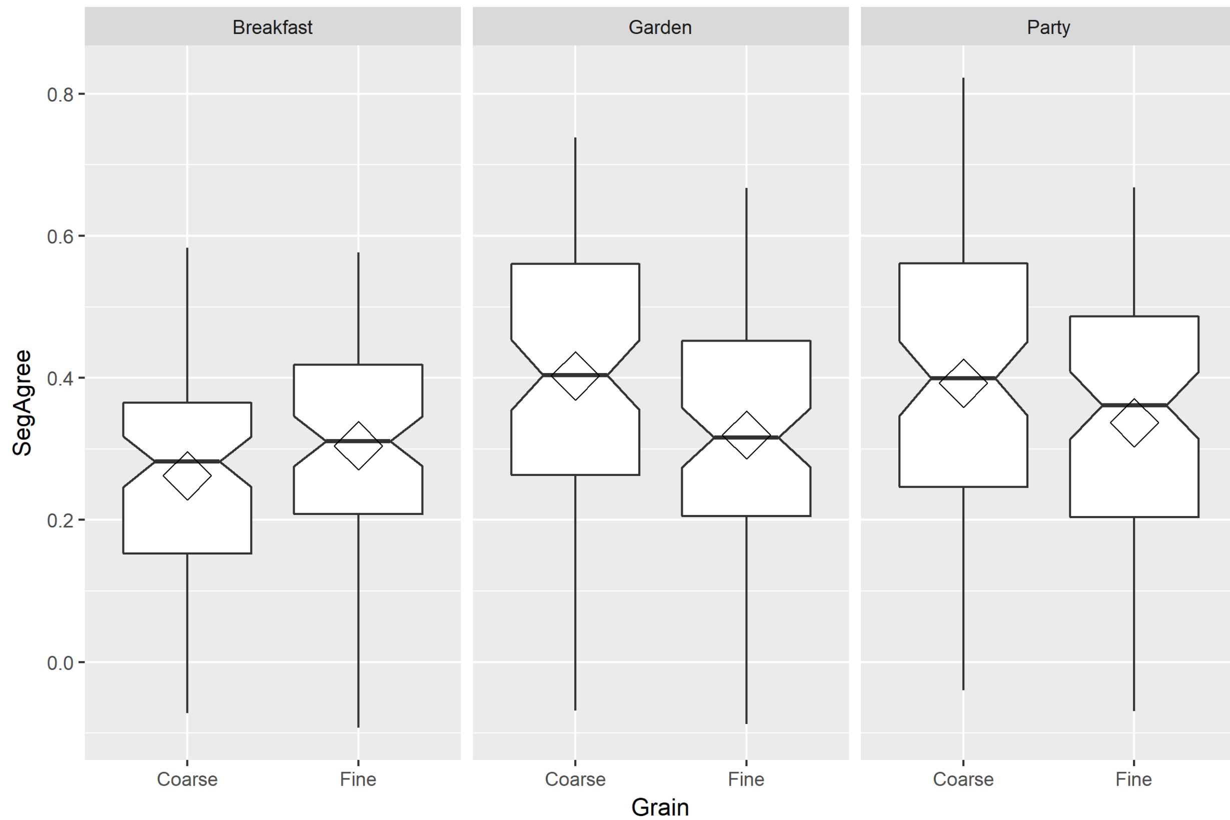

Figure 9. Boxplots of individual subjects' segmentation agreement with the normed segmentation data. Mean diamonds represent a 95% confidence interval around the mean, and the horizontal bars around the midsections represent median segmentation agreement.

3.1.4 Segmentation Agreement Across Studies

In addition to assessing segmentation agreement for participants within the current study, it was also useful to determine whether segmentation performance in this study agreed with segmentation data from Sargent et al (2013). Segmentation performance for participants from Sargent et al.'s (2013) was calculated using the same kernel density smoothing algorithm adapted from Papenmeier (2016). Temporal bandwidths were kept at 3 s, and the probability density distributions over time were correlated between the two studies (Table 2). As can be seen in Figure 10, participants segmentation performance showed a high degree of agreement with the participants from Sargent et al. (2013). However, there did seem to be greater agreement for fine segmentation than for coarse grained segmentation, however there seemed to be a more conservative (i.e., less frequent) response frequency in the current study than in the study by Sargent et al (2013). Nevertheless, the fact that participants between studies appeared to segment similarly would suggest that both groups' perceptions of what was meaningful for each video was interpreted similarly, as well. Additionally, the strong agreement in segmentation data between the two studies suggests that the empirically derived times selected for use in the Gabor discrimination task (Phase 2) were at times in which participants perceived meaningful changes in the event.

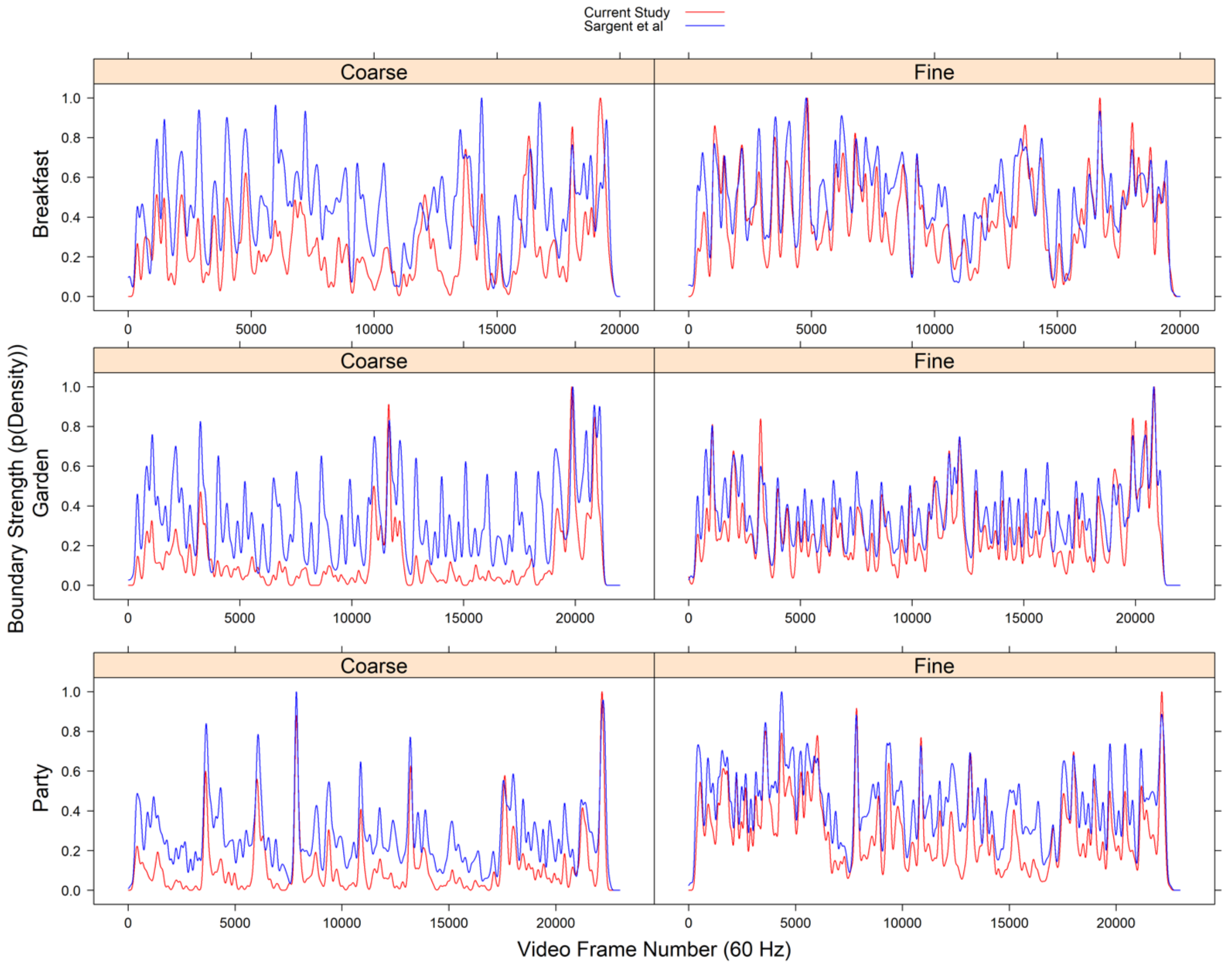

Figure 10. Probability density plots for event segmentation for the current study and from Sargent et al. (2013). Segmentation performance is presented as a function of video and segmentation grain. Note that, although boundaries were identified similarly across the two studies, the data from the current experiment typically showed slightly lower peaks for segmentation probability (especially coarse-grains).

Table 2.

Segmentation Correlations Between the Current Study and Sargent et al. (2013)

Video	Coarse	Fine
Breakfast	0.59	0.82
Garden	0.68	0.88
Party	0.86	0.88

Pearson correlations between the normed segmentation data in Sargent et al. (2013) and the current study. Fine-grained segmentation agreement was consistently high across all three videos, while coarse-grained segmentation agreement was less consistent across the three videos.

3.2 Memory Tasks

Memory for objects and events was also assessed for several important purposes. The first purpose was to ensure that participants were, in fact, attending to the content of the videos and not just devoting all of their attention to the Gabor discrimination task. The second purpose of the memory tasks was to determine if SOA thresholding data served as a useful individual differences predictor for later object memory (Kreitz, Furley, Memmert, & Simons, 2015). A third goal was to determine whether segmentation agreement was a useful predictor for event order memory (DuBrow & Davachi, 2016; Sargent et al., 2013).

3.2.1 Object Memory. Object memory was analyzed by means of a probit mixed-model, evaluating the likelihood of reporting that an object was present. Fixed-effects predictors included object type (e.g, absent, boundary, and non-boundary objects). By-subject random effect structures varied between two models, where one model contained only subject intercepts as a random effects (i.e., 1| Subject), and a second which included object type across subjects (i.e., Object Type | Subject). Finally, another set of models included video as a by-item random effect.

Better model fitness was found when the by-item random effect of video was included ($BIC_{1|Subject+1|Video}$ = 4712.9, $BIC_{Object\ Type|Subject+1|Video}$ = 4754.5) when compared to a model which did not include the by-item random effect ($BIC_{1|Subject}$ = 4860.8, $BIC_{Object\ Type|Subject}$ = 4902.4). However, the model that included object type as a by-subject random effect (i.e., Object Type | Subject + 1|Video) did not demonstrate better fit than the model which contained only subject and video intercepts as random effects. Therefore, the simpler model, which contained only subject and video intercepts as random effects was selected for fixed-effect evaluation.

The fixed effects tests for object memory reveal that participants' memory was sensitive to boundary objects (d' = 1.62, SEM = 0.22) but not non-boundary objects (d' = 0.10, SEM = 0.01), with the two object types differing significantly from each other (z = 18.10, p < .001; Fig. 11). Thus, participants were engaged in attending to and remembering the content related to the event, rather than free-viewing the background information in the scene. This result was also consistent with previous data showing that boundary objects are remembered better than non-boundary objects (Swallow et al., 2009). However, it could be that the simple act of interacting with objects more frequently made these objects more salient than the non-boundary objects, which were interacted with much less frequently.

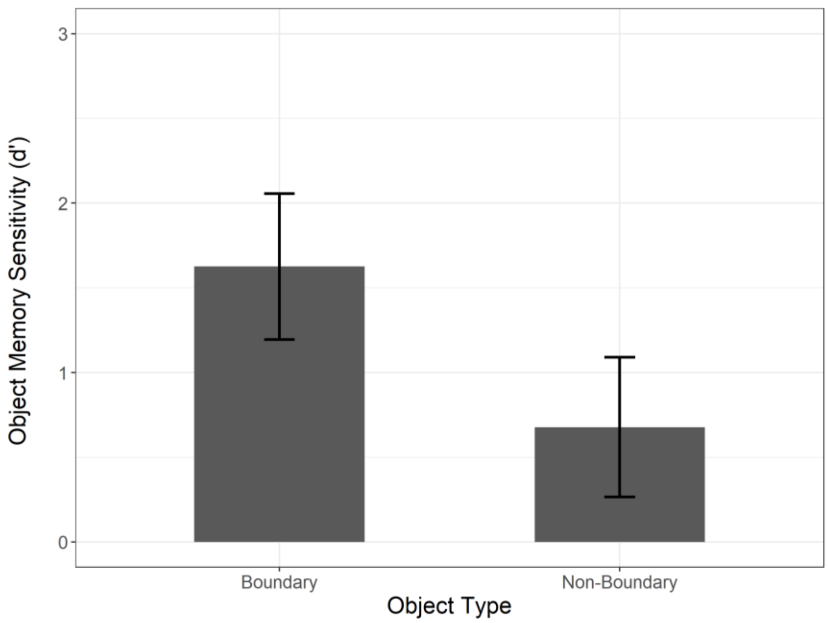

Figure 11. Object memory sensitivity for each video. In all cases, boundary objects were significantly more likely to be correctly recalled compared to non-boundary objects. Error Bars = 95% CI.

3.2.2 Event Order Memory. Like object memory, event order memory was evaluated through a probit mixed-model. The criterion variable was the likelihood of responding that the left-sided event action occurred first, the validity of the left sided action being the only predictor in the regression. The random effects structure varied with four models. Two models included only the subject intercept as a by-subject random effect, and two other models included validity as a by-subject random effect. One of each of the above models also included video type as a by-item random effect. Model fitness was evaluated by BIC values, with the model which included subject intercept as a by-subject random effect and video as a by-item random effect showing the best model fitness (BIC = 3188.5) relative to the next best model, which included target validity|Subject as a by-subject random effect and video as a by item random effect (BIC = 3203.9).

As expected, there was no significant bias to select the right or left action ($B = -0.028$, SEM = 0.09, $z = -0.304$, $p = 0.761$). Furthermore, the results show that participants were minimally but significantly sensitive to event order ($d' = 0.307$, SEM = 0.053, z = 5.80, $p < .001$). Thus, as indicated by the object memory task results, the order memory results show that participants were somewhat actively engaged with encoding the content of the videos during the event perception and Gabor discrimination portion of the experiment. More importantly, there was little chance that the sensitivity to event order could be explained by guessing, since the memory probes were generated from action pairs that discriminated between individuals with high versus low memory performance.

These results agree with order memory results from Sargent et al (2013), though the stimuli in that study were pictures, whereas this study presented descriptions of actions. Given that the videos constituted a meaningful random effect for the results, a subsequent analysis of memory sensitivity for each video found that the event order memory for the *breakfast* video was slightly negatively sensitive ($d' = -.241$, 95% CI[-.49,0.01]), whereas event order memory was positively sensitive for the *garden* video ($d' = .244$, 95% CI[-.01,0.49]) and even more so for the *party* video ($d' = .986$, 95% CI[.72,1.25])(Fig. 12). It is interesting that order memory sensitivity was below chance for the breakfast video, and near chance for the garden video, when segmentation agreement for these two videos was also lower for coarse grains—especially in comparison to fine segmentation agreement for these two videos. Meanwhile, event memory remained highest for the party video, which also showed high (and similar) segmentation agreement for both coarse and fine grains. Thus, it was necessary to compare whether individual subjects segmentation agreement at coarse and fine grains predicted subsequent order memory.

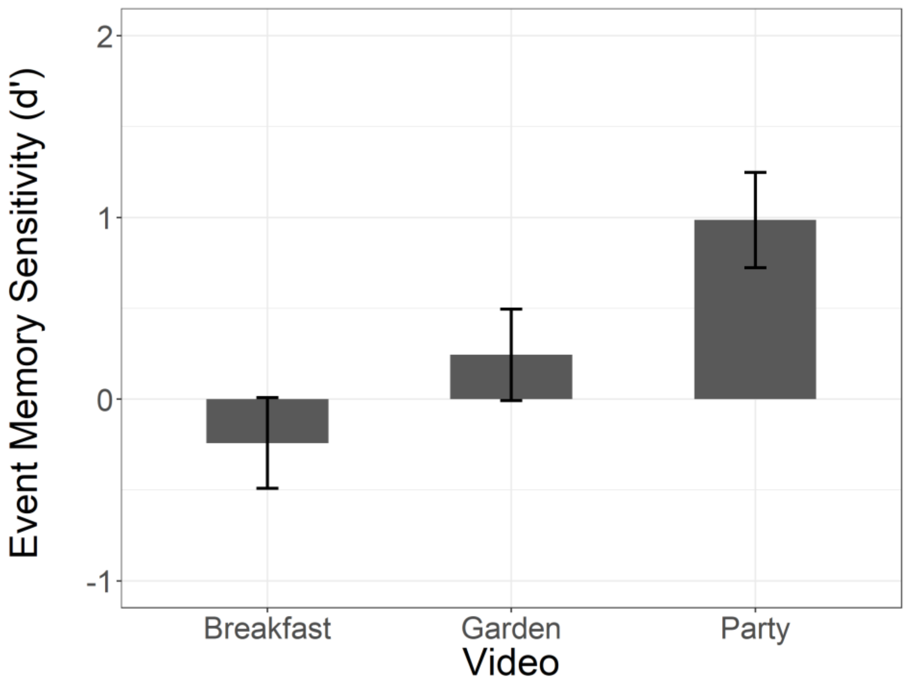

Figure 12. Event memory sensitivity across three videos. Segmentation agreement as a function of video demonstrates clear differences for memory as a function of each video, with the breakfast video showing nearly at-chance memory performance, while the party video shows a relatively high degree of event order memory. Error Bars = 95% Confidence Interval.

Prior research has established that there has been an established relationship between event segmentation agreement and event memory (Sargent et al., 2013, Kurby & Zacks, 2011), and therefore it was important to determine if participants' segmentation agreement predicted later event order memory for the current experiment. For this, three additional models with segmentation agreement as an interacting parameter (Coarse, Fine, Coarse x Fine) were generated to be compared with the above model, which did not include segmentation agreement as a parameter. Model fitness indices for all four models were compared and showed the inclusion of segmentation agreement did improve the goodness of fit in predicting event order memory. Specifically, coarse segmentation agreement positively predicted event order memory ($\Delta d'$ = 0.893, SEM = 0.267, z = 3.34, $p < .001$), significantly improving model fitness over the model which did not include any segmentation agreement (χ^2 (2) = 12.01, p = .002). Thus, recognizing the dramatic shifts in one's event model results in the accurate ordering of event information during long-term memory encoding.

3.3 Gabor SOA Thresholds

In Phase 1 of the experiment, Gabor SOA thresholds on a 1/f noise background were estimated for all participants to determine the baseline speed of processing at different retinal eccentricities (0^o, 4.5^o, and 9^o) and to equalize Gabor accuracy across the visual field prior to manipulating attention in Phase 2. Previous research using the GC-UFOV has only used small (n = 13) sample sizes, and therefore a larger sample would provide a more complete estimate of individual differences in processing times across the visual field. Gabor SOA thresholds were recorded for the 92 participants who completed the entire study, and were analyzed using general linear mixed-modeling, with the log-transformed Gabor SOA being the criterion variable, and Gabor eccentricity being the sole fixed effect. Two models with varying random effects structures were generated; one which included subject intercepts as a random effects variable, and another which included variability across eccentricity by each subject as a random effects variable. The model which included Gabor eccentricity was shown to have a better model fit compared to the model which included only subject intercepts (χ^2 = 6.30, p = .043), and thus it was selected for evaluation of its fixed effects.

Fixed effects tests show that mean SOA was approximately 105 ms (B = 4.65, SEM = 0.05, t = 92.74, p < .001). While there was no fixed effect of eccentricity on processing speeds (B = 0.001, SEM = 0.008, $t(93)$ = 0.15, p = 0.88), as might be expected due to using size-scaled Gabors for each eccentricity, participants did show individual differences in Gabor SOAs from central to peripheral vision. Interestingly, there was a significant positive correlation between subject intercepts and eccentricity slopes (*r(92)* = 0.22, *p* = .03), where participants with lower overall SOAs generally obtained lower SOAs (i.e., faster processing speeds) from central to peripheral vision, and participants with higher overall SOAs obtained higher SOAs (i.e., slower processing speeds) from central to peripheral vision (Fig. 13).

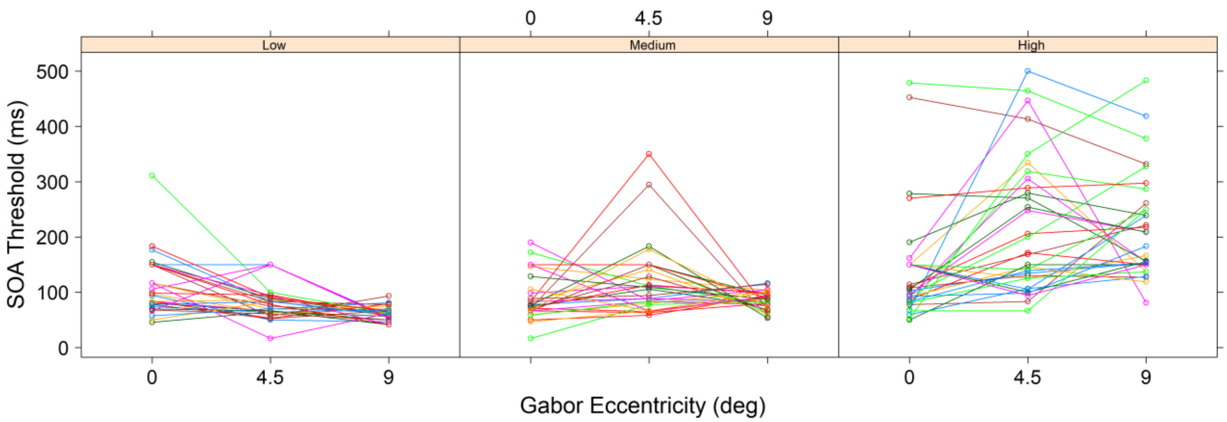

Figure 13. Figure demonstrating the observed Gabor SOAs across retinal eccentricity for each participant. Panels represent three participant groupings, separated by residual scores for Gabor SOA (Low, Medium, and High processing times). Note the slight negative slope for the *low* SOA group compared to the generally positive slope for the *high* SOA group.

An additional factor to consider is whether individual differences in processing speed across the visual field (as per the UFOV task) affected the encoding of objects into long-term memory (Kreitz, Furley, Simons, & Memmert, 2015). The models predicting object memory between boundary and non-boundary objects were augmented to include processing time speeds across the visual field. The factors of mean SOA and SOA eccentricity slope were included as parameters to interact with object type in three different models (mean SOA threshold, eccentricity SOA slope, and mean SOA x eccentricity slope). Model fitness indices were compared, however there was no improvement in model fitness for any of the additional models (mean SOA: χ^2 (3) = 0.89, p = .826; Eccentricity SOA slope: χ^2 (3) = 0.84, p = .839; mean SOA x eccentricity SOA slope: χ^2 (9) = 4.76, p = .85). Therefore, processing time differences across the visual field did not influence object information encoding with regard to event perception.

3.4 Gabor Discrimination During Event Perception

This experiment tested several hypotheses concerning spatial and temporal changes in attention in relation to different event boundaries. To recap, the two competing hypotheses regarding the breadth of attention around event boundaries over time are decomposable into two clusters of outcomes: an *attentional impairment* hypothesis and an *ambient-to-focal shift* hypothesis.

All Gabor analyses were performed on a custom PC with an AMD Ryzen 1800x processor (3.7 GHz), a Mushkin Reactor 1 TB Solid State Drive (SSD), and 16 GB of DDR4 RAM (2400 MHz). The Bayesian mixed models were generated through Hamiltonian Markov Chain Monte Carlo (MCMC) simulation using the brms package in R (version 2.3.1; Burkner, 2017). Posterior distributions were generated by eight sampling chains, distributed over eight processing cores, with 3,000 warm-up iterations and 6,000 iterations, with a total of 24,000 final samples included in the final model[7].

3.4.1 Omnibus Model Structure. Gabor accuracy was used as the criterion variable and coded as 1 or 0 for correct versus incorrect responses, with a Bernoulli distribution being specified. Fixed effects variables included Boundary Type (coarse, fine, non-boundary), boundary offset time (ranging from -4 s to +6 s from the boundary), and Gabor eccentricity (0^{o},

[7] Flat priors for all regression weights were a t-distribution with 3 df, mean of 0, and sd of 10 for the intercept, and a half-t distribution derived from a t-distribution with 3 df, mean of 0, and sd of 10 for all standard deviations of intercepts (across subjects and videos) and within-subject predictor slopes (across subjects).

4.5°, and 9°) as predictors[8]. Video type was included as a by-item random effect in all models. From the available data, two models with differing fixed effects structures were generated.

The model structure of the *attentional impairment* hypothesis (Equation 1) included the interaction of boundary type and boundary offset time and the additive effect of Gabor eccentricity. Thus, if attention is differentially impaired by different boundary types, then Gabor accuracy should be lower at boundaries (fine and/or coarse) compared to non-boundaries. Furthermore, if the time course of attention differs around each boundary, then the boundary offset time (i.e., the time prior-to or after the boundary) should interact with boundary type. However, the attentional impairment hypothesis assumes that attention should be equally affected across the visual field, rather than systematically broadening or narrowing, as is the case in the ambient-to-focal shift hypothesis. Thus, the *ambient to focal* shift hypothesis (Equation 2) differed only in the three-way interaction of event boundary type, boundary offset time, and Gabor eccentricity to measure fluctuations in attentional breadth across time as they differ among each event boundary type.

[8] All continuous predictors were centered to have a mean of 0 when constructing the model.

Equation 1: Fixed-effects structure for the Attentional Impairment hypothesis.

$$\mathrm{Y} = \text{Boundary Type x Time} + \text{Eccentricity}$$

Equation 2: Fixed-effects structure for the Ambient to Focal Shift hypothesis.

$$\mathrm{Y} = \text{Boundary Type x Time x Eccentricity}$$

Random effects structures were both by-subject and by-item to uncorrelate the effects contributed by each subject with each manipulation in the experiment. Nine permutations of each model class (i.e., attentional impairment versus ambient to focal shift) were generated beginning with simple random effect structures with subject intercepts, then singular effects of each fixed effect (e.g., Eccentricity | Subject), additive effects of each fixed effect (i.e., Eccentricity + Boundary Offset Time | Subject), and then the interactive effects of each fixed effect. Model fitness and stability were evaluated by Watanabe-Akaike Information Criterion (WAIC; Vehtari, Gelman, & Gabry, 2015) and additional variance explained between prior and posterior distributions was evaluated by Bayesian R^2 (Gelman, Goodrich, Gabry, & Ali, 2017). Additionally, Bayes factors for the two best fit models by WAIC and Bayesian R^2 were compared to determine the evidence ratio in favor of either of the two hypotheses (Kass & Rafferty, 1995).

3.4.2 Data Filtering. Because participants could respond to the Gabor patches any time prior to the appearance of the next Gabor presentation trigger, the variable which coded accuracy was not reset until the onset of the next Gabor presentation sequence. Thus, the Gabor data were filtered to include only the final data frame prior to the next Gabor trigger. This resulted in a total of 6,210 observations. Additionally, any Gabor presentations in which no mask appeared (i.e., *nil patches*) were removed from the data, resulting in a loss of 1,100 observations, or roughly 17.8% of the overall data. Finally, two subjects were removed for having accuracy at less than chance (< 50%) accuracy, and another was removed for having perfect accuracy (114 observations). Lastly, the distribution of times around the event boundary showed a substantial positive skew, and therefore the data were truncated to include boundary offset times which extended to a maximum of six seconds (44 observations). This resulted in a total of 4,774 observations across 89 participants (See Table 1 for mean Gabor accuracy across eccentricity and boundary type).

The high proportion of nil patches was evaluated for potential causes, the most likely cause being long Gabor presentation durations set during the thresholding task. A hierarchical logistic regression was created to determine whether Gabor SOAs[9] predicted the probability of nil patch presentation. As can be seem in Figure 14, the probability of a nil patch increased significantly with longer SOAs ($\beta = 1.374$, SEM = 0.089, $z = 15.46$, $p < .001$). Gabor eccentricity was included as a factor to determine if this effect varied across eccentricity however there were neither a main effect of eccentricity ($\beta = -0.117$, SEM = 0.096, $z = -1.21$, $p = 0.23$), nor was there an interaction between eccentricity and Gabor SOA ($\beta = 0.020$, SEM = 0.020, $z = 1.01$, $p = .313$). Thus, the nil patch effect seemed to be driven primarily by the amount of time needed for the Gabor to be present before the mask appeared. Therefore, future implementations of the Gabor discrimination task should set lower limits for Gabor SOA times to ensure that the SOA thresholds do not exceed typical fixation durations.

[9] Gabor SOA was log transformed to reduce the degree of positive skew present in the SOA distributions.

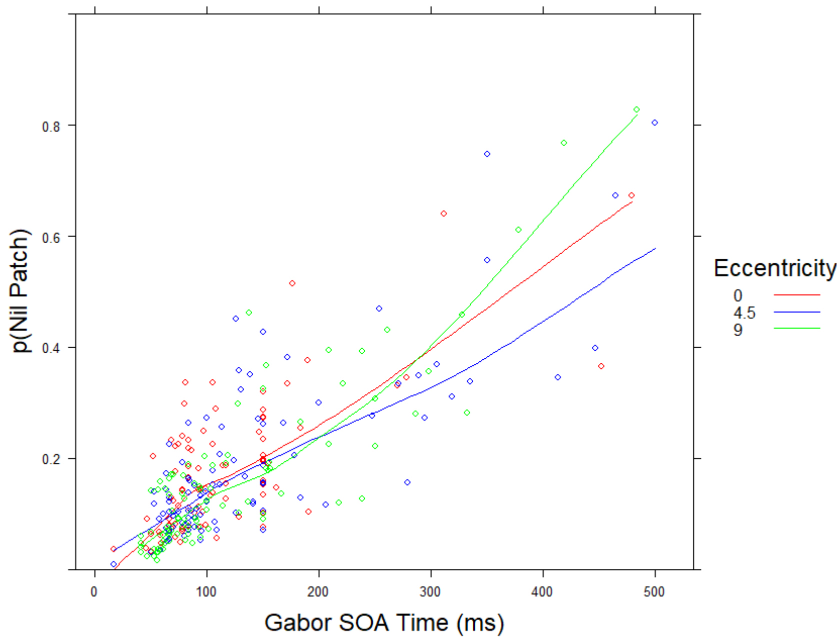

Figure 14. Fitted probabilities of nil patch presentations, or Gabor presentations in which the participant moved their eyes prior to the appearance of the mask. The plot demonstrates that, as the Gabor SOA threshold time increases, the probability of a nil patch also increases. The figure also demonstrates that nil patch presentations were not affected by Gabor eccentricity.

Table 3.

Observed Gabor Discrimination Accuracy Across Eccentricity and Event Boundary Type

	0^o	4.5^o	9^o
Coarse	.78 (.22)	.84 (.21)	.70 (.24)
Fine	.82 (.15)	.80 (.19)	.75 (.16)
Non-Boundary	.89 (.19)	.76 (.29)	.72 (.31)

Mean Gabor accuracy (proportion correct) and standard deviations (in parentheses) across 89 participants. In general, accuracy was within the range of the accuracy target for the SOA thresholding task (Phase 1), however there was a tendency for accuracy to be highest at the fovea (0^o).

A maximum of 76 observations were possible throughout the recorded trials, however data loss due to nil patches (i.e., Gabor presentations where a saccade terminated the gaze-contingent stimuli before a mask presentation) or other possible presentation errors led to differences in the number of observations for each participant. For the 89 participants whose data were included in the Gabor results, roughly 2/3 of the maximum number of observations were recorded (Mean = 54.3, SD = 10). Only four subjects (4.5%) provided less than half of the maximum possible observations.

3.4.3 Linear Omnibus Model Results. All parameter estimates successfully converged (R = 1.00). WAIC's for all models were compared, revealing that Gabor eccentricity was the only by-subject random effect that appeared to be stable between both the attentional impairment and ambient to focal shift models. The WAIC values were lower for the attentional impairment model (WAIC = 4742.98) compared to the ambient to focal shift (WAIC = 4745.06) model. However, the difference in goodness of fit between the two models was roughly equal to the standard error (Δ_{WAIC} = -2.08, SEM = 5.38). Bayesian R^2 values for each of the models were compared, revealing similar results, however the ambient to focal model accounted for a slightly greater degree of variance (R^2 = .092, 95% CI = [.075, .11]) than the attentional impairment model (R^2 = .088, 95% CI = [.072, .10]). Nevertheless, the models did not meaningfully differ from each other. Furthermore, Bayesian R^2 values do not account for the inherent benefit of capturing more explained variance through additional model complexity, as is the case with the additional interaction of eccentricity between boundary type and boundary offset time. Instead, a more appropriate model comparison should be made with Bayes factors (Kass & Raferty, 1995; Kruschke, 2016). When the *attentional impairment* and *ambient to focal* models are compared, an extremely high Bayes factor was found in favor of the *attentional impairment* hypothesis

(Bayes Factor = 931,090). Though at first glance such a result would suggest strong evidence in favor of the *attentional impairment* hypothesis, the more likely explanation for such a large difference between the two models could have to do with poor overall fit of the models by the specification of the parameters of the model, or by improperly specified prior distributions (Burkner, 2017; Kass & Rafferty, 1995)). To identify sources of error in the structure of the model, the population-level fixed effects for both models were evaluated.

Table 4.

Gabor Discrimination Parameter Estimates for Attentional Impairment Model Across Event Grain, Boundary Offset Time, and Eccentricity (Base = Coarse Boundary at Time = 0, Eccentricity = 4.5^{o}*)*

	Event Boundary Grain			Grain x Time Interaction		
	Estimate	SEM	95% CI	Estimate	SEM	95% CI
Coarse (Intercept)	1.328	.530	0.17, 2.39	-0.018	0.030	-0.08, 0.04
Fine (Slope)	0.122	0.085	-0.04, 0.287	0.037	0.037	-0.04, 0.11
Non-Boundary (Slope)	0.211	0.119	-0.02, 0.45	0.037	0.052	-0.07, 0.14

Note. Table showing the regression estimates for the attentional impairment model (Coarse = Base). These parameter estimates model changes in mean Gabor accuracy across time for each boundary type. Slope differences for the fine and non-boundaries assume the Coarse boundary as the intercept point. Thus, the slope values for each condition for overall Gabor accuracy and the interaction between boundary gain and time are relative to the coarse boundary. All parameter estimates converged with Ř < 1.01.

Mean Gabor accuracy (see Table 3) for the coarse boundary (see Fig. 15) was well above chance (*Mean* = 79.1%; β = 1.33, SEM = 0.53, 95% CI = [0.05, 2.63]), whereas accuracy was slightly higher for fine boundaries (*Mean* = 80.9%; β = 0.12, SEM = 0.09, 95% CI = [-0.04, 0.29]). However, non-boundaries showed a higher accuracy relative to coarse boundaries (*Mean*

= 82.3%, β = 0.21, SEM = 0.12, 95% CI = [-0.02, 0.45]). As can be seen in *Figure 16*, attention remained narrowed, with Gabor accuracy decreasing with increasing eccentricity (β = -0.07, SEM = 0.01, 95% CI = [-0.09, -0.04]). This occurred despite the fact that both size-scaling of the Gabors and processing time thresholds were designed to equalize Gabor accuracy across the visual field. Thus, the presence of objects in the scene likely caused a greater degree of foveal attentional bias, compared to the $1/f$ noise backgrounds during the thresholding phase of the experiment. Finally, Gabor accuracy did not change as a function of boundary offset time (Fig. 13) for the coarse boundaries (β = -0.02, SEM = 0.03, 95% CI = [-0.08, 0.04]), fine boundaries (β = 0.04, SEM = 0.04, 95% CI = [-0.04, 0.11), or non-boundaries (β = 0.04, SEM = 0.05, 95% CI = [-0.06, 0.14]). Therefore, the only major finding in the attentional impairment model was that Gabor accuracy decreased slightly among event grains, with the highest accuracy being found for non-boundaries, slightly lower for fine boundaries, and lowest for the coarse boundaries.

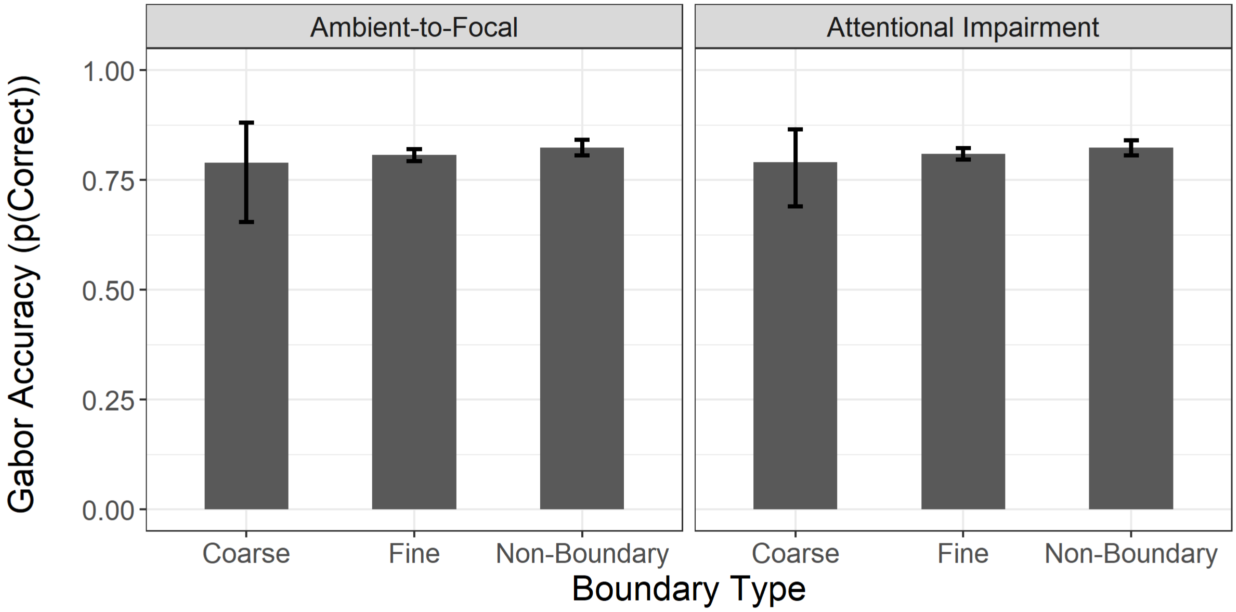

Figure 15. Gabor discrimination overall accuracy means for the coarse, fine, and non-boundaries sampling periods for the *attentional impairment* and *ambient to focal shift* models. Note that each of the models shows the poorest accuracy occurring for the coarse boundaries, however there was a great degree of variability for the coarse boundary, which was also the base category. Error Bars = 1 SEM.

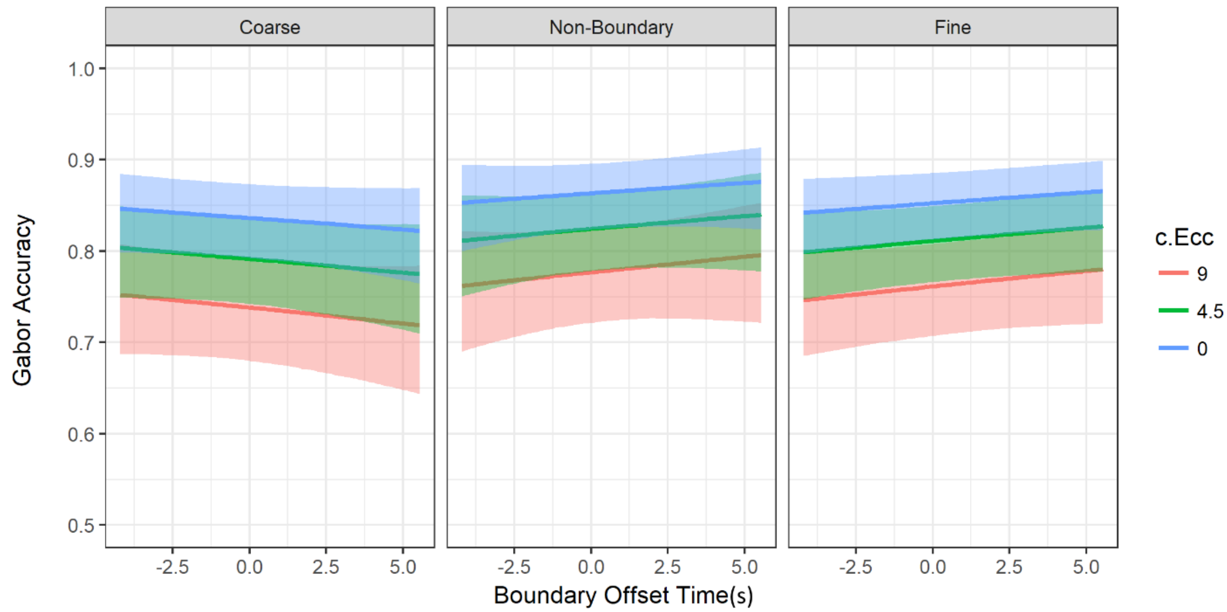

Figure 16. Plotted marginal effects of the attentional impairment model, which fits Gabor accuracy as a function of event boundary grain (including non-boundary times), Gabor eccentricity, and boundary offset time (in seconds). Note that Gabor accuracy is higher for the non-boundary period, whereas accuracy is slightly lower for both boundary types. Error bars = 1 SEM.

In contrast to the *attentional impairment* model (see Table 3), the ambient to focal shift model (Fig. 17) attempted to determine if attention shifts to an ambient state of processing during event boundaries, while non-boundary periods should show narrowing of attention. Like the attentional impairment model, the ambient to focal shift model (Fig. 17) Gabor accuracy was marginally above chance for coarse boundaries (*Mean* = 78.8%; β = 1.318, SEM = 0.682, 95% CI = [-0.17, 2.60]), and accuracy was slightly higher for fine boundaries (Mean = 80.7%; β = 0.112, SEM = 0.085, 95% CI = [-0.01, 0.46]). Also, relative to coarse boundaries, non-boundaries showed relatively higher Gabor discrimination accuracy (non-boundary: *Mean* = 82.4%; β = 0.227, SEM = 0.120, 95% CI = [-0.01, 0.47]). Similar to the attentional impairment model, the ambient to focal shift model showed no main effect of boundary offset time for the coarse boundary (β = -0.01, SEM = 0.03, 95% CI = [-0.07, 0.05]), with no difference from the fine boundaries (β = 0.03, SEM = 0.05, 95% CI = [-0.07, 0.13]) or non-boundaries (β = 0.03, SEM = 0.04, 95% CI = [-0.05, 0.10]) being observed.

One important factor that differs between the *attentional impairment* and *ambient to focal* shift models is the interaction of Gabor eccentricity with boundary grain, where positive slopes depict broadened attention (i.e., ambient processing), while negative slopes depict narrowed attention (i.e., focal processing). The degree of attentional tunneling for the coarse boundaries was similar to the *attentional impairment* model, with a significant negative slope being observed for the coarse boundary (β = -0.07, SEM = 0.02, 95% CI = [-0.11, -0.03]). Also similar to the *attentional impairment* model, attention was broadened for fine boundaries (β = 0.01, SEM = 0.02, 95% CI = [-0.03, 0.06]), and slightly narrowed for non-boundaries (β = -0.04, SEM = 0.03, 95% CI = [-0.10, 0.02]). Furthermore, Figure 17 shows that there was a robust difference in attentional breadth between fine and non-boundaries, with slopes for Gabor accuracy across eccentricity being more positive (and thus more broadened) across retinal eccentricity than non-boundaries (β = 0.12, SEM = 0.023, 95% CI = [0.08, 0.17]). Thus, there is some evidence that non-boundaries remain slightly more tunneled than coarse boundaries and are dramatically more tunneled compared to fine boundaries.

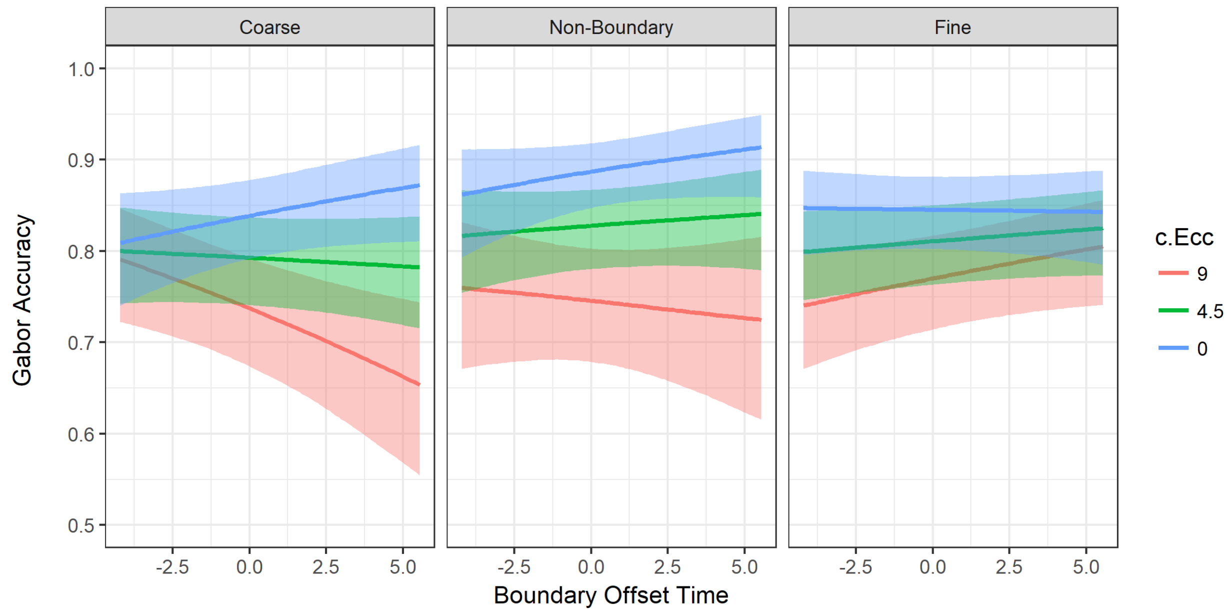

Figure 17. Plotted marginal effects of the ambient to focal shift model, which fits Gabor accuracy as an interactive function of event boundary type (including non-boundaries), Gabor eccentricity, and boundary offset time. This model shows distinct interactions between boundary offset time and eccentricity for each boundary grain. The coarse boundaries show broadened attention prior to event boundaries (with Gabor accuracy being equal for 0°, 4.5°, and 9° eccentricities), with attention narrowing as time moves past the normed event boundary (with Gabor accuracy being greater for 0° than 9°). Fine boundaries show prolonged broadening of attention, whereas non-boundaries show that attention stays narrowed. Error bars = 1 SEM.

Table 5.

Gabor Discrimination Parameter Estimates for Ambient-to-Focal Shift Model Across Event Grain, Boundary Offset Time, and Eccentricity (Base = Coarse Boundary at Time = 0, Eccentricity = 4.5^{o})

A)	Event Boundary Grain			Grain x Time Interaction		
	Estimate	SEM	95% CI	Estimate	SEM	95% CI
Coarse (Intercept)	1.318	.682	-0.17, 2.60	-0.010	0.030	-0.07, 0.05
Fine (Slope)	0.112	0.085	-0.06, 0.28	0.027	0.037	-0.05, 0.10
Non-Boundary (Slope)	0.277	0.120	-0.02, 0.46	0.028	0.053	-0.08, 0.13

B)	Event Boundary Grain x Eccentricity			Event Boundary Grain x Time x Eccentricity		
	Estimate	SEM	95% CI	Estimate	SEM	95% CI
Coarse (Intercept)	-0.069	0.021	-0.11, -0.03	-0.013	0.008	-0.03, 0.00
Fine (Slope)	0.014	0.023	-0.03, 0.06	0.018	0.010	0.00, 0.04
Non-Boundary (Slope)	-0.040	0.032	-0.10, 0.02	0.005	0.014	-0.02, 0.03

A) (Left)This table shows logit mean Gabor accuracy as a function of boundary offset time for each boundary type, as well as (right) the interaction between each boundary type with boundary offset time. All parameter estimates assume the Coarse boundary as the intercept point. Thus, the slope values for each coarse condition are relative to the coarse boundary, and interactions between grain and time assume that the slope differences are in comparison to the coarse boundary slope over time. B) (Left) Table showing the effect of eccentricity on Gabor accuracy with each event boundary grain. Each comparison is being made to the two-way interaction of the event boundary grain and eccentricity at the coarse boundary. Negative slopes indicate better Gabor accuracy with smaller eccentricities (i.e., *tunnel vision*) and positive slopes indicate broadened attention. (Right) Slopes of the 3-way interaction between event boundary grain, boundary offset time, and Gabor eccentricity on Gabor accuracy. Each comparison is being made to the three-way interaction event boundary grain, boundary offset time, and eccentricity at the coarse boundary.

However, another crucial difference between the model structures of the two hypotheses is that the ambient to focal shift model contains a three-way interaction between boundary type, boundary offset time, and Gabor eccentricity (see Table 5B, right panel). As seen in *Figure 17*, Coarse boundaries showed a minimally negative interaction between time and eccentricity (β = -0.013, SEM = 0.008, 95% CI = [-0.03, 0.00]), suggesting that attention may start broadened prior to coarse boundaries, and narrows as time passes the event boundary. However, it must be emphasized how slight the effect seems to be, especially when considering the strong degree of error for the parameter estimate. Meanwhile, the fine boundary differed from coarse boundaries substantially, with a positive interaction between boundary time and eccentricity (β = 0.018, SEM = 0.010, 95% CI = [0.00, 0.04]), suggesting that attention became broader as the boundary passed in time. Finally, the non-boundary condition showed no difference in its interaction between time and eccentricity compared to the coarse boundaries (β = 0.005, SEM = 0.014, 95% CI = [-0.02, 0.03]), though *Figure 17* gives the appearance of less of a difference between Gabor eccentricities over time. Therefore, the three-way interaction in the *ambient-to-focal shift* model suggests that the coarse boundaries produce the strongest change in eccentricity over time, whereas attentional breadth during non-boundary and fine boundary periods is relatively stable over time.

Given that the differences between the two competing models (e.g., attentional impairment vs. ambient to focal shift) were not discriminable using WAIC values and showed only marginal gains in explained variance, an additional exploratory analysis of the data was conducted to determine where sources of error variance might exist, and how they might be modeled. Furthermore, all models showed difficulties in generating posterior distributions, with large numbers of divergent chains after the warmup period, and other models showing that all samples drawn after the warmup period exceeded the maximum tree depth for the Markov chain. This occurred despite increasing the number of pre- and post-warmup samples, as well as increasing the acceptability of extreme samples in the chain (delta = 0.99) from the default settings (Burkner, 2017).

Additional consideration was given to the data itself to determine the sources of extreme noise. Therefore, several exploratory analyses were conducted. The first set of exploratory analyses assessed curvilinearity that may exist for Gabor accuracy over time. The second set of exploratory analyses determined whether the addition of the non-boundary condition introduced excessive noise and evaluate model fitness in its absence. The third set of exploratory analyses determined if the introduction of the by-item random effects structure created an overly complex model. Models with and without the by-item random effects structures were created, while also specifying weakly informative prior distributions on the basis of the target accuracy thresholds.

3.4.4 Examination of Curvilinearity of Gabor Accuracy Using Categorical Time Bins. Previous data from both Crundall, Underwood, and Chapman (2002) and Huff, Papenmeier, and Zacks (2012) suggests that attentional shifts in time do not occur monotonically, and instead show a shift around the boundary time and resume to a baseline level. Therefore, the first step will be to determine *if* Gabor accuracy demonstrates such a shift over time. Raw natural spline fits of the data (Fig. 18) shows that Gabor accuracy does seem to show very different interactions between Gabor eccentricity and boundary offset time for each boundary type. While there seems to be only minimal nonlinearity among both the coarse and fine-grained boundaries, there were noticeable bends in the shape of the non-boundary (control) periods, where the 0^{o} and 4.5^{o} eccentricities were relatively flat, whereas Gabor accuracy at the 9^{o} eccentricity showed a positive bend after the event middle (0 s), or the point in time where the normed segmentation data indicated that a boundary was unlikely to be detected. This is in contrast to the interactive model fit which shows very little difference between time and eccentricity for the coarse and non-boundary grains. Nevertheless, the LOESS spline fits do not provide specification for when the bends in Gabor accuracy occur, nor do they provide estimates for the degree of magnitude for the changes in Gabor accuracy. Furthermore, the Loess natural splines may also be susceptible to over-regularization, or forced linearity despite sharp, curves that would be indicative of a sudden shift in an otherwise linear function (Knowles & Renka, 2012; Rice & Rosenblatt, 1983). Therefore, it was important to specify the knots necessary to measure the data.

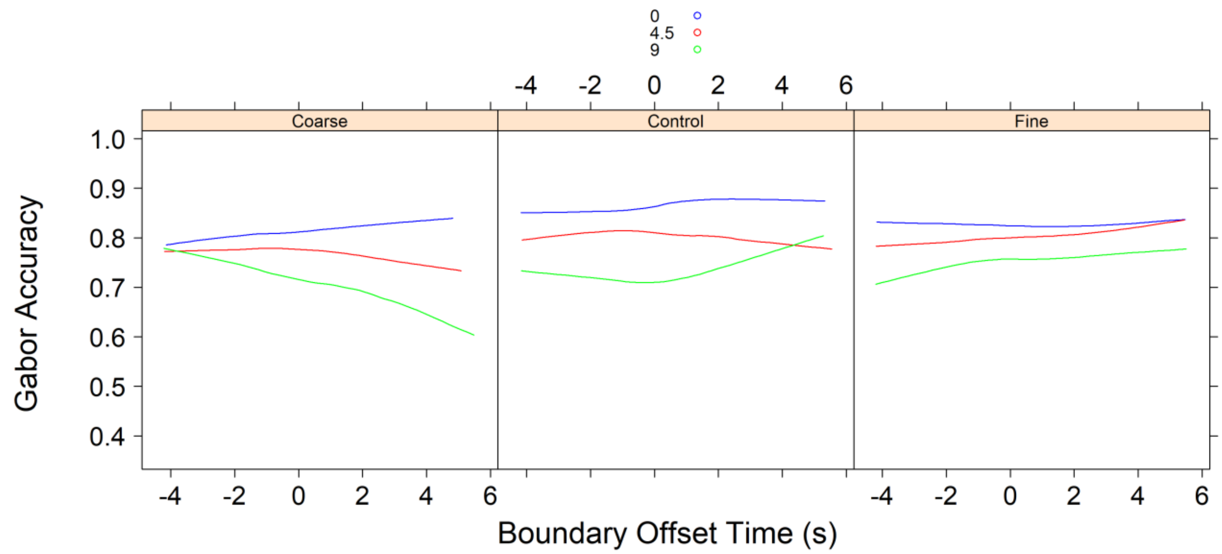

Figure 18. Raw fits of Gabor accuracy as a function of boundary offset time and eccentricity, for each segmentation grain. The raw fit of the data strongly resembles the ambient-to-focal shift model. Notice that for the coarse-grained boundaries (left) attention is broadened prior to the event boundary (0 s) and narrows with increasing time, whereas attention at the fine-grained boundary remains broadened over time. The primary difference in the raw and fitted data is that the control condition seems to show some curvilinearity following the event middle (0 s), though this could be due to some outlying observations.

To determine where the bends in Gabor accuracy occur, a two-step process was applied. In the first step, boundary offset times were binned into six categorical levels, and analyzed using the Bayesian hierarchical regression. In the second step, the resulting number of bends in Gabor accuracy over time were used to inform the number of knots to be included in a fitted natural spline, where time was once again coded as a continuous variable.

For the first step in modeling the bends in Gabor accuracy over time, a phase-shift time series analysis was conducted. For this analysis, boundary offset time was divided into 6 separate bins (-4 s – -2.5s, -2.5 s – -1 s, -1s – 0.5 s, 0.5 s – 2s, 2 s – 3.5s, 3.5 s – 6s), which corresponded to the intervals for the Gabor presentation trigger times. Like the original Bayesian hierarchical regression, two different models were generated using the additive and interactive structures which corresponded to the attentional impairment and ambient-to-focal shift hypotheses, respectively. The number of warmup and post-warm up samples taken were identical to the original omnibus analysis.

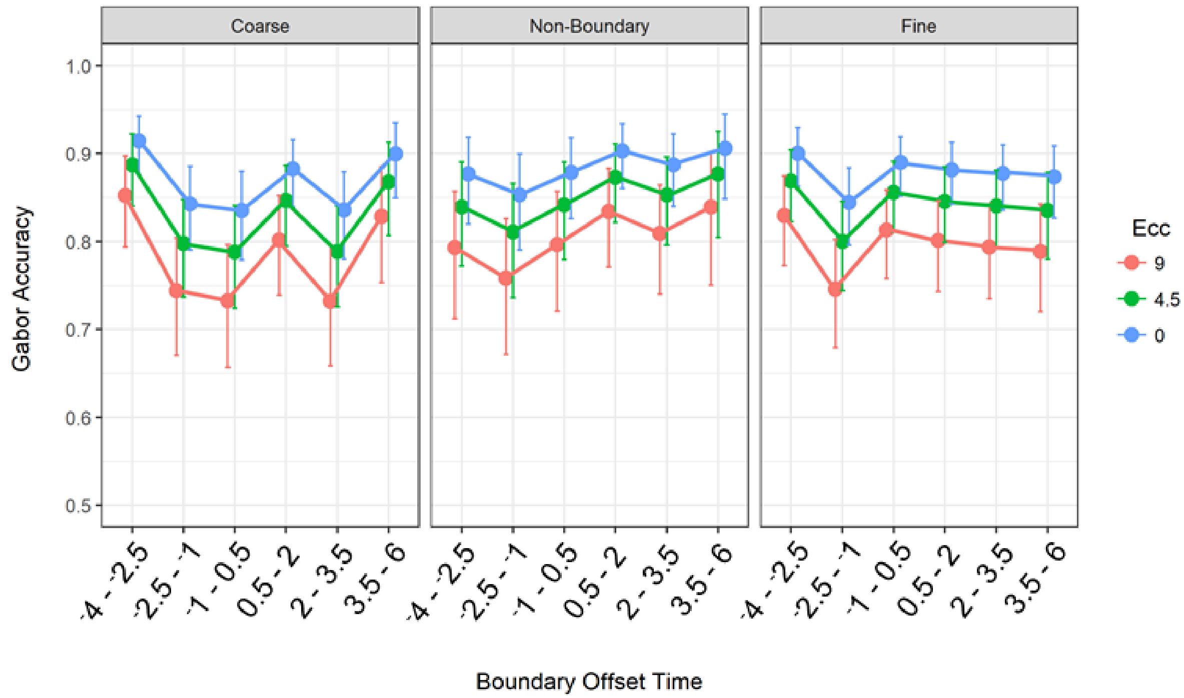

Figure 19. Fitted marginal effects of the attentional impairment model where boundary offset time has been binned into six levels: level 1 (-4 s – -2.5 s), level 2 (-2.5 s – -1 s), level 3 (-1 s – 0.5 s), level 4 (0.5 s – 2 s), level 5 (2 s – 3.5 s), and level 6 (3.5 s – 6 s). Note that both the coarse and non-boundary times both show the maximum number of knots in Gabor accuracy over time (3), which suggests that this level of flexibility will be necessary to capture the shape of attentional changes over time. Error Bars = 1 SEM.

Table 6.

Gabor Discrimination Parameter Estimates For Attentional Impairment Model Using Categorical Time Bins (Base = Coarse Boundary at Bin 1)

	Bin 1 (-4s – -2.5s)	Bin 2 (-2.5s – -1s)	Bin 3 (-1s – 0.5s)	Bin 4 (0.5s – 2s)	Bin 5 (2s – 3.5s)	Bin 6 (3.5s – 6s)
Coarse (Intercept)	2.054 (0.619)	-0.693 (0.286)	-0.750 (0.292)	-0.361 (0.291)	-0.751 (0.284)	-0.178 (0.383)
Fine	-0.174 (0.292)	0.183 (0.355)	0.637 (0.363)	0.169 (0.359)	0.516 (0.356)	-0.090 (0.455)
Non-Boundary	-0.407 (0.365)	0.485 (0.476)	0.767 (0.476)	0.628 (0.466)	0.849 (0.461)	0.486 (0.624)

Note. Table showing the changes in mean Gabor accuracy across six time bins for each boundary type. All differences assume the Coarse boundary as the intercept point. Thus, the slope values for each coarse condition are relative to the coarse boundary at the first time bin and each comparison for the fine and non-boundary conditions are being made with the coarse boundary for its respective time bin.

For the attentional impairment model (Table 6), Gabor accuracy for the coarse boundary (Fig. 19, left panel) was well above chance in the 1st time bin (-4s – -2.5 s; β = 2.05, SEM = 0.70, 95% CI = [0.57, 3.44]). Following the 1st time bin, accuracy dropped substantially for bins 2 (-2.5 s – -1 s; β = -0.69, SEM = 0.29, 95% CI = [-1.27, -0.13]) and 3 (-1 s – .5 s; β = -0.75, SEM = 0.30, 95% CI = [-1.34, -0.17]). Gabor accuracy briefly returned to near baseline accuracy in the 4th time bin (0.5 s – 2 s; β = -0.36, SEM = 0.30, 95% CI = [-0.96, 0.22]), before it dropped again in the 5th time bin (2 s – 3.5 s; β = -0.75, SEM = 0.30, 95% CI = [-1.33, -0.19]), and then returned to baseline in the final (6^{th}) time bin (3.5s – 6 s; β = -0.18, SEM = 0.39, 95% CI = [-0.93, 0.59]). In total, three bends were observed for the coarse boundaries. The first bend occurred prior to the event boundary, keeping accuracy relatively low until the 4^{th} time bin, where accuracy momentarily increased, and fluctuated again between the 5^{th} and 6^{th} time bins.

Relative to the coarse boundary, the fine boundary (Fig. 19, right panel) did not differ substantially from the coarse boundary at the 1^{st} time bin (-4 s – 2.5 s; β = -0.17, SEM = 0.30, 95% CI = [-0.77, 0.41]), nor did it differ from the coarse boundary at the 2^{nd} time bin (-2.5 s – -1 s; β = 0.18, SEM = 0.36, 95% CI = [-0.52, 0.88]). In the 3^{rd} time bin, Gabor accuracy increased slightly, relative to the coarse boundary (-1 s – 0.5 s; β = 0.633, SEM = 0.37, 95% CI = [-0.09, 1.37]), but did not differ significantly from the coarse boundary relative to 4^{th} bin (0.5 s -- 2 s; β = .17, SEM = 0.36, 95% CI = [-0.55, 0.88]), 5^{th} bin (2 s – 3.5 s; β = 0.51, SEM = 0.36, 95% CI = [-0.93, 1.23]), or 6^{th} bin (3.5s – 6 s; β = -0.09, SEM = 0.46, 95% CI = [-0.99, 0.80]). Thus, like the coarse grain boundaries, Gabor accuracy dropped prior to the event boundary. However, unlike the coarse-grained boundary, Gabor accuracy around fine-grained boundaries remained relatively high and flat across time after the boundary had passed, requiring only one knot to fit nonlinearity in Gabor accuracy over time.

The non-boundary condition (Fig. 16, center panel) produced Gabor accuracy that was, again, quite similar to the coarse boundaries, with accuracy being slightly (though not reliably) lower in the 1^{st} time bin (-4 s – -2.5 s; β = -.398, SEM = 0.37, 95% CI = [-1.12, 0.32]). Again, the 2^{nd} time bin showed a drop in Gabor accuracy similar to the coarse boundary, though muted in comparison (-2.5 s – -1 s; β = 0.758, SEM = 0.48, 95% CI = [-0.47, 1.41]). This refractory period was also short-lived in comparison to the coarse boundary, with Gabor accuracy being noticeably higher in the 3^{rd} (-1 s – 0.5 s; β = 0.76, SEM = 0.49, 95% CI = [-0.21, 1.71]) and 4^{th}

(0.5 s – 2 s; β = .62, SEM = 0.47, 95% CI = [-0.31, 1.54]) time bins. Gabor accuracy was similar to the coarse boundary in the 5th (2 s – 3.5 s; β = 0.83, SEM = 0.46, 95% CI = [-0.08, 1.75]) and 6th time bin (3.5s – 6 s; β = -0.48, SEM = 0.46, 95% CI = [-0.51, 0.61]). Given the similarities between coarse and non-boundaries regarding Gabor accuracy over time, it appears that the non-boundary period would also need three knots to appropriately fit the data.

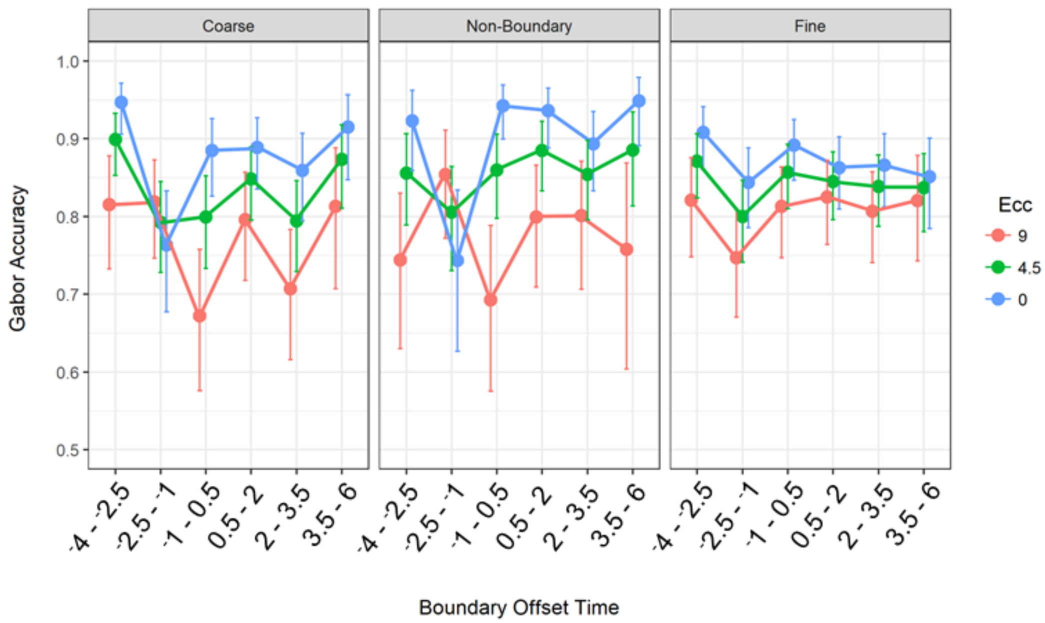

Figure 20. Fitted marginal effects of the *ambient-to-focal* model where boundary offset time has been binned into six levels beginning with bin 1 (-4 s – -2.5 s), bin 2 (-2.5 s – -1 s), bin 3 (-1 s – 0.5 s), bin 4 (0.5 s – 2 s), bin 5 (2 s – 3.5 s), and bin 6 (3.5 s – 6 s). Note that both the coarse and non-boundaries remain tunneled across time, except for brief moments in bin 2 (-2.5s – -1s) where attention broadens. Conversely, attention remains broadened for the majority of the fine boundary time. Error bars = 1 SEM.

Table 7.

Gabor Discrimination Parameter Estimates for Ambient-to-Focal Shift Model Using Categorical Time Bins (Base = Coarse Boundary at Bin 1)

A) Boundary Type x Time Bin (Base = Coarse Grain at Bin 1)

	Bin 1 (-4s – -2.5s)	Bin 2 (-2.5s – -1s)	Bin 3 (-1s – .5s)	Bin 4 (0.5s – 2s)	Bin 5 (2s – 3.5s)	Bin 6 (3.5s – 6s)
Coarse (Intercept)	2.184 (0.711)	-0.851 (0.314)	-0.811 (0.326)	-0.470 (0.319)	-0.845 (0.316)	-0.253 (0.415)
Fine	-0.281 (0.322)	0.323 (0.379)	0.834 (0.536)	0.716 (0.520)	0.818 (0.504)	0.522 (0.686)
Non-Boundary	-0.399 (0.416)	0.480 (0.516)	0.688 (0.393)	0.252 (0.385)	0.582 (0.384)	-0.024 (0.483)

B) Grain x Eccentricity x Time Bin (Base = Coarse Grain at Bin 1, Eccentricity = 4.5^{o})

	Bin 1 (-4s – -2.5s)	Bin 2 (-2.5s – -1s)	Bin 3 (-1s – .5s)	Bin 4 (.5s – 2s)	Bin 5 (2s – 3.5s)	Bin 6 (3.5s – 6s)
Coarse x Eccentricity	-0.156 (.071)	0.193 (.082)	0.009 (0.085)	0.076 (0.319)	0.052 (0.083)	0.054 (0.110)
Fine x Eccentricity	-0.070 (0.085)	-0.174 (0.100)	0.006 (0.103)	-0.022 (0.102)	-0.015 (0.102)	0.007 (0.686)
Non-Boundary x Eccentricity	-0.005 (0.108)	0.047 (0.134)	-0.070 (0.137)	-0.060 (0.137)	0.027 (0.133)	-0.093 (0.179)

Note. A) Mean Gabor accuracy as a function of boundary offset time for each boundary type. All differences assume the Coarse boundary at time bin 1 as the intercept point. Thus, the slope values for each coarse condition are relative to the coarse boundary at the 1^{st} time bin and each comparison for the fine and non-boundary conditions are being made with the coarse boundary for its respective time bin. B) This table demonstrates the interaction with boundary offset time and Gabor eccentricity for each boundary type, with each comparison for being made to the Coarse boundary at the first time bin. Negative slopes indicating better Gabor accuracy with smaller eccentricities (i.e., *tunnel vision*) and positive slopes indicating broadened attention.

For the ambient-to-focal shift model (Fig. 20, Table 7), the same six time bins were generated, however the model structure utilized the three-way interaction between boundary type, boundary offset time, and Gabor eccentricity. Mean Gabor accuracy for the Coarse boundaries in the 1^{st} time bin was also relatively high, and well-above chance (-4 s – -2.5 s; β = 2.18, SEM = 0.71, 95% CI = [0.72, 3.49]) before dropping substantially in the 2^{nd} (-2.5 s – -1 s; β = -0.811, SEM = 0.31, 95% CI = [-1.49, -0.26]) and 3^{rd} (-1 s – 0.5 s; β = -0.811, SEM = 0.33, 95% CI = [-1.47, -0.19]) time bins. In the 4^{th} time bin, attentional impairment from the baseline level was slight (0.5 s – 2 s; β = -.470, SEM = 0.32, 95% CI = [-1.11, 0.14]), but less than for the 2^{nd} and 3^{rd} time bins. Gabor accuracy decreased again in the 5^{th} time bin (2 s – 3.5 s; β = -0.845, SEM = 0.32, 95% CI = [-1.49, -0.25]), before reaching the same level of accuracy as the first pre-boundary time bin (3.5s – 6 s; β = -0.253, SEM = 0.41, 95% CI = [-1.06, 0.57]). Thus, there were three critical bends in overall Gabor accuracy, with one longer decrease in overall attention occurring prior and at the event boundary, before increasing slightly immediately following the boundary then dropping briefly dipping at roughly 2 – 3.5 s after the event boundary. Nevertheless, additional bends were needed to fit the interaction between Gabor offset time and eccentricity for each boundary type.

In terms of the interaction between time and eccentricity, attention was tunneled in the 1st bin of the coarse boundary condition (-4 s – -2.5 s; β = -0.156, SEM = 0.07, 95% CI = [-0.30, -.02]), before the trend completely reversed, showing a bias for peripheral information in the 2nd bin (-2.5 s – -1 s; β = 0.193, SEM = 0.08, 95% CI = [.04, 0.36]). In the 3rd bin, attention shifted to become tunneled (-1 s – 0.5 s; β = 0.009, SEM = 0.09, 95% CI = [-0.21, 1.71], to a similar degree as the 1st time bin, which was the case for all other times (see Table PhaseShiftAmb2Foc for full regression reports). Differences in the time by eccentricity interaction were minimal between coarse and non-boundary conditions, with trends being in the same direction over the entire duration of the sampling period (i.e., bins 1-6). Conversely, the fine boundary remained substantially broadened for the entirety of the sampling period and failed to show the *vista vision* effect, where attention is better in the periphery than the fovea, at the 2nd time bin (-2.5 s – -1 s; β = -0.174, SEM = 0.10, 95% CI = [.38, 0.17]). Thus, approximately five knots were necessary to capture the complex interaction between eccentricity and time for each boundary type.

The fact that Gabor accuracy was more stable for the fine boundaries that it was at non-boundaries is a surprising result, considering that event models in WM should be stable at this time, as well (Kurby & Zacks, 2008; Reynolds, Zacks, & Braver, 2007). However, behavioral and EEG data for central and peripheral cueing tasks demonstrates that, while sustained attention can be used to improve target recognition at a validly cued location, spatial and object-based attention seem to cycle in antiphase with each other, routinely switching between cued and uncued locations. Although the cycling rate between object and space-based attention (3-5 Hz) occur at a faster rate than the cycles for the attentional shifts in this study (3-5 seconds), it is possible that the shift from broad to narrow attention in dynamic environments is a normal state of affairs for attentional processing, whereas the prolonged broadened attention is indicative of a disruption of this system. However, further electrophysiological data would be necessary to confirm or reject this hypothesis.

Based upon the results from the above analyses, it is possible that a range of knots will be needed to capture the curvilinearity of Gabor accuracy across time and space. However, there are several caveats inherent to changing continuous variables into nominal variables that must be considered when interpreting the above analyses, which emphasizes why the above results are better interpreted as a guide for generating future models, and not being used as a true model. First, the binning of time reduces power by creating discrete cut-off points in which the continuous extremes of one bin may be more closely related to the mean of their neighbor than their own, thereby increasing the amount of residual error in each parameter estimate (Irwin & McClelland, 2003). To combat such residual error, the range of each time bin was determined by the Gabor trigger times, which did seem to correspond with the modes in which the presentation times occurred. However, creating a fixed number of time bins may have also hidden further bends in the timeseries, which could not be quantified by either model. Therefore, the results from the categorical timeseries analysis was used to inform the structure of four models in which boundary offset time was again coded as a continuous variable, while curvilinearity in the data was fit using natural splines with prespecified degrees of freedom.

3.4.5 Examination of Curvilinearity of Gabor Accuracy Using Natural Splines. Four additional models were generated to determine whether model fitness for the linear models from *Section 3.4.3* could be improved upon with added flexibility in the form of a natural spline fit. The results from section 3.4.4 suggested that there were potentially three or five shifts in Gabor accuracy around event boundaries, for the attentional impairment and ambient-to-focal shift hypotheses, respectively. For each of the attentional impairment and ambient-to-focal shift models, time was fitted by two different spline fits (Bates & Venables, 2014)(*splines* library, v3.5.0; Venables & Bates, 2018): one in which time was fitted with three degrees of freedom and another in which time was fitted with 5 degrees of freedom (based on the results from section 3.4.4). By including equal numbers of knots in the linear function for both model types, each hypothesis could be evaluated with the same number of bends model fitness. These models were then compared with the two models in which no curvilinearity was assumed.[10] After models were generated, comparisons of model fitness were made on the basis of WAIC and Bayesian R^2 values.

The best-fit model was still the linear attentional impairment model (WAIC = 4743.5), the simplest model of the six tested (for full spline fit results, see Table 8). The next best model fits were all roughly equal and included the linear ambient-to-focal shift model (WAIC = 4746.1, Δ_{WAIC} = -2.63, SEM = 5.38), and the attentional impairment models fitted with three degrees of freedom (WAIC = 4748.8, Δ_{WAIC} = -5.52, SEM = 5.25). Thus, on the basis of WAIC values, the addition of splines did not outweigh the benefit that the more simple, linear models had in predicting the data. Conversely, the Bayesian R^2 values show that, within their respective curvilinear structure, the ambient-to-focal models generally outperform the attentional impairment models (Table 9). However, with spline degrees of freedom being equal between each hypothesis, the increase in explained variance was still smaller than the standard error. Thus, the question of which hypothesis provides a truly better fit has not yet been resolved, however given two models with equivalent fitness, the simpler (attentional impairment) model should be reserved.

[10] Model convergence for spline fits failed with R-hat exceeding 1.1. Post burn-in samples were increased from 24,000 to 32,000. This resulted in stable model fits with R-hat being near 1.00 for all parameters.

Table 8.

WAIC Values for Natural Spline Fits Across Attentional Impairment and Ambient-to-Focal Shift Models.

	Attentional Impairment	Ambient-to-Focal Shift
Linear	4743.47 (74.54)	4746.10 (74.72)
D.F. = 3	4748.98 (74.79)	4760.14 (75.29)
D.F. = 5	4756.68 (75.09)	4765.17 75.98

Note. WAIC values for the linear and natural spline fits for Gabor offset time. Each row represents different numbers of degrees of freedom and each column represents its effect on the models representing each hypothesis. Parenthesis contain SEM of the WAIC value. Note that model complexity increases monotonically between both the increasingly complex interactions in the fixed effects structure (i.e., attentional impairment versus ambient-to-focal shift) and with increasing degrees of freedom, which provides added flexibility for changes in Gabor accuracy over time.

Table 9.

Bayesian R^2 Values for Natural Spline Fits Across Attentional Impairment and Ambient-to-Focal Shift Models.

	Attentional Impairment	Ambient-to-Focal Shift
Linear	0.089 (.072, 0.107)	0.092 (0.075, 0.110)
D.F. = 3	0.092 (0.075, 0.109)	0.096 (0.079, 0.114)
D.F. = 5	0.094 (0.077, 0.112)	0.103 (0.086, 0.121)

Note. Bayesian R^2 values (and 95% Confidence Intervals) for the six models, which include the original linear models for each hypothesis, as well as the spline-fitted models. Though the proportion of variance explained by the posterior distributions for each parameter tended to increase with model complexity, the confidence intervals for all models tended to overlap.

The previous two attempts to improve model fitness have been in terms of the fixed effects structure. Thus, the problem may lie in the random effects structure or from the use of flat, uninformed prior distributions in generating the posterior distribution, as WAICs (Gelman, 2006; Vehtari et al., 2015) and Bayes factors (Kass & Rafferty, 1995) sensitive to both of these issues. Regarding the random-effects structure, one issue in sampling efficiency is from a low number of observations within each condition, with sparse numbers of observations in each condition causing model fitness indices to underestimate how accurately a model can predict data from a posterior distribution (Vehtari et al., 2015). This problem is exacerbated when models become more complex, where parameters (and their interactions) are less likely to have sufficient data from each participant at each level of the model's hierarchy.

As mentioned in section 3.4.3, there were several subjects with less than half of the maximum possible number of observations. With the additional complexity of a by-subject and by-item random-effects structure, it becomes increasingly less likely for a sufficient quantity of data to be modeled for each condition. One solution to this problem is to remove these participants, however this option would leave open the possibility that participants with a high number of nil patches may be indicative of poor inhibitory control (Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001) and working memory capacity (Unsworth, Schrock, & Engle, 2004). Removing these participants would therefore underestimate the degree of variability in attentional changes that are inherent in the population, and any subsequent models may not generalize to future populations. Instead, it may be more beneficial to reduce the complexity of the random effects structure to remove any random effects which demonstrate a poor fit. While by-subject random effects structures showed good fits for the intercept and eccentricity in both the attentional impairment (intercept: $\sigma = 0.707$, SEM = 0.074, 95% CI = [0.58, 0.86]; eccentricity: $\sigma = 0.075$, SEM = 0.018, 95% CI = [0.038, 0.11]) and ambient-to-focal shift (intercept: $\sigma = 0.710$, SEM = 0.074, 95% CI = [0.58, 0.87]; eccentricity: $\sigma = 0.077$, SEM = 0.018, 95% CI = [0.04, 0.11]) models, this was not the case for video type. The parameter estimates for the standard deviation demonstrated a poor fit for the linear attentional impairment ($\sigma = 0.678$, SEM = 0.918, 95% CI = [0.10, 3.33]) and ambient-to-focal shift ($\sigma = 0.791$, SEM = 1.185, 95% CI = [0.11, 4.15]) models. Thus, the random effect for video failed to effectively capture a reliable degree of variance, and it would be of greater benefit to remove it from the model. An additional analysis was conducted, with the by-item random effect omitted, while weakly informative prior distributions were used to limit the sampling range for the posterior distribution.

3.4.6 Model Optimization for Gabor Discrimination During Event Perception Through Informed Prior Distributions and Simplified Random Effects

To reduce the computational demands inherent in the previous analyses, six models were constructed, with two versions being created for each of the attentional impairment and ambient-to-focal shift hypotheses: a set of models in which boundary offset time was fitted to a linear function, a set of models fit to a natural spline with three degrees of freedom, and a set of models fit to a natural spline with five degrees of freedom. While the by-subject random effect of eccentricity was maintained from the previous models, the by-item random effect of video was omitted to reduce the impact of participants with low numbers of observations for a given video. Furthermore, weakly informative priors for the population intercept were included to constrain the sampling range of the prior distribution. The intercept mean of the prior distribution was set to the target Gabor threshold (logit(0.75) = 1.1) with a standard deviation of .7, and prior distributions for slopes were set to a mean of 0 (SD = 1), as suggested by (Gelman, Jakulin, Pittau, & Su, 2008).

Contrary to all other Gabor discrimination task models, no transitions in the MCMC simulation diverged beyond the acceptable distance from the predicted path. However, the WAIC values for all models did not appear to change meaningfully compared to the previous models (see Table 10), and in fact increased slightly (i.e., the new models showed a slightly poorer fit). Nevertheless, the trend from previous models showing a better (though not significant) fit for the attentional impairment model versus the ambient-to-focal shift model was also true for the models which did not include video as a by-item random effect but did include weakly informative priors. But again, the WAICs for two best models, the linear attentional impairment model (WAIC = 4758.13, SEM = 74.30) and the linear attentional impairment model

(WAIC = 4761.56, SEM = 74.50), were less than one standard error of measurement from each other (Δ_{WAIC} = -3.42, SEM = 5.08). Thus, in terms of overall model fitness, the two models seem to show similar goodness of fit, but the overall standard error for population level parameter estimates was slightly smaller compared to all previous models. Bayes factors between the two linear models were generated, and found an abnormally large value in favor the *attentional impairment* model (Bayes factor = 81,197,357), which, like the results in section 3.4.3, is should not be considered reliable.

Table 10.

WAIC Values for the Linear and Natural Spline Fits for Gabor Offset Time.

	Attentional Impairment	Ambient-to-Focal Shift
Linear	4758.14 (74.30)	4761.56 (74.50)
D.F. = 3	4764.13 (74.58)	4776.33 (75.07)
D.F. = 5	4771.23 (74.87)	4781.66 (75.82)

Note. For these models, the by-item random effect of video had been removed and a weakly informative prior had been specified to be equal to the target Gabor accuracy mean during SOA thresholding in Phase 1. Each row represents different numbers of degrees of freedom and each column represents its effect on the models representing each hypothesis. Parenthesis contain SEM of the WAIC value. Note that model complexity increases monotonically between both the increasingly complex interactions in the fixed effects structure (i.e., *attentional impairment* versus *ambient-to-focal shift*) and with increasing degrees of freedom, which provides added flexibility for changes in Gabor accuracy over time.

Regarding the proportion of variance explained between the prior and posterior distributions, the Bayesian R^2 again showed that the ambient-to-focal shift models predicted a marginally higher proportion of variance compared to the attentional impairment models (see Table 11). However, the most noticeable difference between in variance explained occurred with the ambient-to-focal shift model with five degrees of freedom (Bayes R^2 = 0.099, SEM = 0.009), though its attentional impairment counterpart was still roughly 1% lower and within the 95% confidence interval of the ambient-to-focal shift model.

Table 11.

Bayesian R^2 Values for the Linear and Natural Spline Fits for Gabor Offset Time.

	Attentional Impairment	Ambient-to-focal shift
Linear	0.085 (0.068, 0.102)	0.088 (0.071, 0.105)
D.F. = 3	0.088 (0.071, 0.105)	0.092 (0.075, 0.109)
D.F. = 5	0.090 (0.073, 0.107)	0.099 (0.082, 0.116)

Note. Bayesian R^2 values (and 95% Confidence Intervals) for the six models, which include the original linear models for each hypothesis, as well as the spline-fitted models. Though the proportion of variance explained by the posterior distributions for each parameter tended to increase with model complexity, the confidence intervals for all models tended to overlap.

3.4.7 Gabor Discrimination At Event Boundaries As a Function of Boundary Strength. The fact that segmentation agreement differed as a function of video also suggests attention was differentially affected by the videos, despite the computational difficulties brought by the by-item random effects structure in section 3.4. Furthermore, the clear differences across the videos as a whole and in terms of the boundary strength at each moment in time would suggest that a finer grained assessment of the attentional analyses is necessary. In other words, the degree to which attention is affected at a particular boundary may be dependent upon the salience of that boundary. Finally, the high degree of segmentation alignment between coarse and fine boundaries suggests that both grains play a role in affecting attention, despite boundaries being previously labeled as fine *or* coarse. Therefore, an additional analysis of the Gabor discrimination data was conducted to determine the unique effects of Gabor boundary salience on attention.

To again test the competing *attentional impairment* and *ambient-to-focal shift* hypotheses, five Bayesian hierarchical mixed models were generated for each hypothesis. Weakly informative priors for accuracy (intercept) were set at a mean of 1.1 (SD = .7), the inverse logit of which matched the target threshold for Gabor accuracy), and slopes were set to a mean of 0 (SD = 1). All models had a fixed effects structure where boundary type interacted with boundary offset time. For the *attentional impairment* model, Gabor eccentricity was included as an additive effect whereas the *ambient-to-focal shift* model included Gabor eccentricity as an interaction. To test the effects of boundary strength, the probability density value for coarse and/or fine boundaries was included for four models (e.g., coarse, fine, coarse + fine, coarse * fine)[11]. These models were compared to a simpler model, which did not include the probability density values for coarse or fine-grained segmentation. In keeping with the trends from the previous Gabor discrimination task results in section 3.4, eccentricity was included as the only by-subject random effect. Given that the by-item random effect of video caused substantial computational problems, and that the segmentation data provides information unique to each

[11] The non-negative nature of the event boundaries required a root transformation to satisfy assumptions of normality.

video, no by-item random effects were included. Model selections were made on the basis of WAIC and Bayesian R^2 values.

Comparisons of WAIC values again demonstrated that the models supporting the attentional impairment hypothesis provided a more parsimonious fit of the data, though not significantly so. Among the *attentional impairment* models, the best fit came from the model which included the interaction between boundary type, boundary offset time, and the probability density for fine boundaries (WAIC = 4745.72), however this did not differ substantially from the next best attentional impairment model, which included additive effects of coarse and fine boundaries (WAIC = 4753.31, Δ_{WAIC} = -7.59, SEM = 8.15). However, the model which included fine boundary probability density did fit marginally better than the model which included no boundary information at all (WAIC = 4761.12, Δ_{WAIC} = -15.40, SEM = 10.87). Additionally, model fitness indices for the best *ambient to focal* model (which included the additive effects of coarse and fine boundaries (WAIC = 4754.46, Δ_{WAIC} = -8.75, SEM = 8.28). Bayes factors for the two models were generated, and again found an exceedingly large Bayes factor (Bayes factor = 1.8×10^{10}). Therefore, again, neither model could be reliably selected.

With regard to Bayesian R^2 values, the ambient-to-focal shift model which included the interaction between coarse and fine boundary density accounted for the greatest proportion of variance explained (Bayesian R^2 = 0.105, SEM = 0.009, 95% CI = [.089, .122]), however showed virtually no difference from the best attentional impairment model, which also included the interaction between coarse and fine boundary probability density (Bayesian R^2 = 0.097, SEM = 0.009, 95% CI = [.088, .114]). Therefore, the inclusion of boundary strength among the fine and coarse boundaries did not show sufficient evidence to tease apart the differences between the two competing hypotheses.

3.4.8 Gabor Discrimination During Event Perception as a Function of Individual Segmentation Agreement. Individual segmentation agreement has been shown to affect later memory for event information in both young and aging individuals. Given the intimate connection between attention and memory, it was critical to determine if individual segmentation agreement would be useful in providing improved model fitness in predicting attentional changes around event boundaries. Specifically, stronger attentional changes might be observed by participants who were attending normatively to the event structure, whereas weaker attentional changes would be predicted for participants who segmented atypically. Furthermore, differences in segmentation agreement as a function of grain might provide an additional degree of precision in predicting event order memory. In particular, previous research has suggested that fine event segmentation ability is better at predicting later event memory (Sargent et al., 2013; Kurby & Zacks, 2012). However, the current data has found that coarse segmentation agreement seemed to be better in predicting event order memory, and therefore the models generated from this data may predict that coarse segmentation would be more efficient in predicting attentional changes at the event boundary.

To test these hypotheses, individual participants' segmentation agreement for each video and segmentation grain was included as a predictor for the Gabor discrimination task, with five Bayesian generalized linear mixed models being constructed for each of the attentional impairment and ambient to focal shift hypotheses, which maintained the same fixed-effects structure of the previous models. These models included one in which there were no effects of segmentation agreement, a model with additive effects of fine event segmentation, a model with additive effects of coarse segmentation agreement, a model with additive effects of fine and coarse segmentation agreement, and a model with interactive effects of fine and coarse segmentation. No by-item random effects were included to maintain a simple model structure, and because the segmentation agreement values were specific to each video.

The results again found that the best model fitness by WAIC was found for the models that did not contain segmentation agreement information (see Table 12). This was the case for both the *attentional impairment* (WAIC = 4186.03) and *ambient-to-focal shift* (4189.80) hypotheses, which again, did not differ from each other (Δ_{WAIC} = -3.78, SEM = 5.07). However, on the basis of Bayesian R^2 values (Table 13), the *ambient-to-focal shift* model which included the interaction of coarse and fine segmentation agreement was marginally better (Bayesian R^2 = 0.101, SEM = 0.009, 95% CI = [0.08, 0.12]) than the attentional impairment counterpart (Bayesian R^2 = 0.091, SEM = 0.009, 95% CI = [0.07, 0.11]). Nevertheless, because the Bayes factors for the two models were extremely large (Bayes factor = 8.6 x 10^7), the results again show very little reliable support for a difference between the *attentional impairment* and *ambient-to-focal* shift hypotheses.

Table 12.

WAIC Values for Gabor Discrimination Ability for the Attentional Impairment and Ambient-to-Focal Models with the Inclusion of Segmentation Agreement for Fine and Coarse Boundaries.

Segmentation Agreement	Attentional Impairment	Ambient-to-Focal
Absent	4186.03 (69.90)	4189.80 (70.15)
Fine	4196.96 (70.17)	4208.63 (70.71)
Coarse	4195.70 (70.23)	4209.93 (70.75)
Fine + Coarse	4197.22 (70.20)	4227.64 (71.32)
Fine x Coarse	4213.76 (70.93)	4243.46 (72.02)

Note. WAIC values and their standard errors (in parenthesis) in relation their inclusion with the Attentional Impairment and Ambient to Focal Shift Models each model. Fixed effects of segmentation agreement are in addition to the fixed effects of the models used in previous linear equations for the *attentional impairment* and *ambient-to-focal shift* models.

Table 13.

Bayesian R^2 Values for Gabor Discrimination Ability for the Attentional Impairment and Ambient-to-Focal Models with the Inclusion of Segmentation Agreement for Fine and Coarse Boundaries.

Segmentation Agreement	Attentional Impairment	Ambient-to-Focal
Absent	0.085 (.009)	0.088 (.009)
Fine	0.087 (.009)	0.092 (.009)
Coarse	0.087 (.009)	0.091 (.009)
Fine + Coarse	0.087 (.009)	0.096 (.009)
Fine x Coarse	0.091 (.009)	0.101 (.009)

Note. Bayesian R^2 values and their standard errors (in parenthesis) in relation their inclusion with the Attentional Impairment and Ambient-to-Focal Shift Models each model. Fixed effects of segmentation agreement are in addition to the fixed effects of the models used in previous linear equations for the *attentional impairment* and *ambient-to-focal shift* models.

3.4.9 Gabor Discrimination Task Modeling Summary

Despite the difficulty in establishing robust effects of event boundaries on attention, there are several consistent effects that were observed across the majority of models that were run. The first clear finding is that there is a consistent tradeoff in model complexity and parsimony between the ambient to focal and attentional impairment models, respectively. Models testing both hypotheses tend to show that the coarse boundaries display a stronger degree of attentional impairment, replicating previous results from Crundall, Underwood, and Chapman (2002) and Huff, Papenmeir and Zacks (2012). In contrast, the attentional impairment effect for the fine boundaries was substantially weaker, with the difference in Gabor accuracy between fine and non-boundaries only being half of the difference of the coarse versus non-boundary comparison. Thus, it seems that coarse boundaries represent a greater degree of cognitive load compared to fine boundaries.

The second important contribution of this study is that it is the first to systematically compare differences between attention across the visual field during real-world event perception. Perhaps the most robust effect from all the models was the main effect of eccentricity, which showed a negative slope for Gabor accuracy (i.e., attention was biased to central vision). This is important to note because both Gabor size-scaling and Gabor processing time thresholds were devised to hold Gabor accuracy equal across the visual field when attention was not biased to any particular location. By thresholding processing time in the 1/f noise backgrounds and testing attention in real-world scenes, attention was able to narrow to focus on objects present within the scene. Therefore, the main effect of eccentricity is not at all surprising. The critical question, then, is whether the interactions between eccentricity and boundary offset time provide meaningful and reliable effects that generalize to event perception as a whole, since this is the crux of the ambient-to-focal shift hypothesis.

Attentional tunneling results from the presence of a foveal perceptual load (Williams, 1988; 1989), and in the context of the ambient to focal processing model (Unema et al., 2005) attentional tunneling represents a period in which scene context determines top-down salience of an object within that scene, and thus the need for fine-detailed processing of that object arises. While the model fitness indices of the WAICs show that both the simple attentional impairment models and more complex ambient-to-focal shift models provide equal goodness of fit, the ambient-to-focal shift models tend to provide marginally improved explained variance from the prior to posterior distributions. Furthermore, there is a dramatically different effect between coarse and fine boundaries, where coarse boundaries are in a brief state of broadened attention before narrowing into focal mode after the event boundary. Conversely, the fine boundaries remain in ambient mode throughout the sampling period (-4s – 6 s). If there is a difference between coarse and fine boundaries beyond slightly greater attentional impairment at coarse boundaries, it could be that fine boundaries exhibit a less efficient transition between one event model to the next, whereas coarse boundaries facilitate a more rapid shift in time from one model to the next. However, these interpretations must be tempered with the fact that model fitness indices failed to discriminate between the *ambient to focal* and *attentional impairment* models. And although the *ambient to focal* model shows that Gabor accuracy dropped 15% from 4 seconds before to 6 seconds after the event boundary, the residual error was quite high, making this finding unreliable. Therefore, further investigation will be needed to confirm such if such an effect is real.

Chapter 4 - Discussion

4.1 Overview

The purpose of this experiment was to determine the spatiotemporal dynamics of visual attention during real-world event perception. This was the first experiment of its kind to estimate attentional dynamics of real world events in a way that disentangled effects of visual attention from sensory limitations, by utilizing the Gaze-Contingent Useful Field of View methodology (Ringer et al., 2014). Furthermore, this study utilized the framework of Event Segmentation Theory (Zacks, Tversky, & Iyer, 2001; Zacks & Swallow, 2007) to operationalize the varied and dynamic structure of real-world events to predict when critical changes in attention might occur. Nevertheless, the current study was unable to find sufficient support for either the *attentional interference* (Crundall et al., 2002; Huff et al., 2012) or the *ambient-to-focal shift* (Unema, et al., 2005; Velichovsky, et al., 2005) explanations of attention at event boundaries. This study also utilized cutting edge techniques in hypothesis testing through the use of Bayesian hierarchical modeling to determine the viability of one hypothesis over another, rather than the null hypothesis testing that is common with traditional statistical approaches. This is important to note given that productive scientific research should be systematic and incremental, and the ability of Bayesian hierarchical modeling allows for the use of prior distributions to more efficiently estimate population level effects that will generalize to future populations.

4.2 Hypothesis Evaluation Summary: Attentional Impairment versus Ambient-to-focal shifts

Although the model clusters generated to test the two competing hypotheses were not completely discriminable, there were parameter estimates consistent with both hypotheses. For both the attentional impairment and ambient-to-focal shift hypotheses, Gabor accuracy for coarse boundaries was generally lower than non-boundaries. Both models demonstrated that accuracy at the coarse boundary (≈ 79% correct) was slightly, but consistently lower in comparison to non-boundaries (≈ 83% correct). Meanwhile, Gabor accuracy at the fine boundaries (≈ 81%) was between coarse and non-boundaries, but did not differ from either. Therefore, there seems to be reliable evidence for attentional impairment (Crundall et al., 2002) at coarse boundaries. However, what has not been ruled out is the presence of changes in the size of the useful field of view at event boundaries.

The critical difference between the two hypotheses is in terms of the interaction of Gabor eccentricity with boundary offset time among the three boundary types (i.e., coarse, fine, and non-boundary). In many ways, the attentional impairment versus ambient-to-focal shift hypothesis distinction is similar to comparisons between general interference of attention (Crundall et al., 2001; 2002; Huff, Papenmeier, & Zacks, 2012) and attentional tunneling (i.e., tunnel vision; Williams, 1988; 1989). In a previous version of the Gabor discrimination task, these two attentional states were not mutually exclusive (Ringer et al., 2016). While an auditory working memory dual-task load was found to only produce general interference, a foveal dual-task load resulted in tunnel vision *and* general interference (Ringer et al., 2016). Additionally, the foveal load produced stronger general interference ($\Delta d'$ = 3.23 between single and dual-task sensitivity) compared to the auditory load ($\Delta d'$ = 0.733 between single and dual-task sensitivity). In terms of multiple resource theory (Wickens, 2002), this difference between the two types of loads (perceptual versus cognitive processing) suggests that interfering with information at earlier stages of processing produces a stronger degree of interference as the same perceptual channels are occupied. Conversely, the working memory load occurs at a later stage, during the cognitive processing stage, which results in more information being preserved from the earlier perceptual channels (Wickens, 2002). Since, the degree of general interference of attention that occurs at event boundaries is relatively weak, this would support the idea that event boundaries represent a working memory load and not a perceptual load. Furthermore, the perceptual load hypothesis predicted that attention would have narrowed at event boundaries, since prediction error would have made it more difficult to extract meaning from the event. In actuality, if event boundaries do show fluctuations in attentional breadth, they are in the opposite direction, where the event boundary results in broadened attention and the event middles result in narrowed attention. Thus, it is likely that event boundaries represent a *change* in attentional control and not overexertion of attention. Thus, the effects of event boundaries on attention are more likely to be at a higher cognitive level (e.g., WM), rather than at a perceptual level.

If the ambient-to-focal shift model is accurate in describing effects of event boundaries on attention, a working-memory explanation for changes in attention would still be able to explain changes in attentional breadth in response to event boundaries, despite the shift from ambient to focal mode. In the case of the current data, coarse boundaries showed evidence of an interaction between boundary offset time and eccentricity, such that attention was broadly distributed prior to the boundary and narrowed after the boundary, while fine boundaries remained broadly distributed over the course of the ten-second duration around the boundary. Meanwhile, event middles showed prolonged narrowed attention. Interestingly, the effect of coarse boundaries on attention was more discrete, Gabor accuracy being nearly identical for all eccentricities prior to the boundary, and quickly narrowing in the seconds that followed. Thus, it seems that attention would be more strongly affected by coarse boundaries, which would support multiple neuroimaging studies evaluating brain activity at event boundaries (Kurby & Zacks, 2018; Speer et al., 2003; Zacks et al., 2001; Zacks et al., 2006). This would also make a unique prediction for SPECT (Loschky et al., 2018) where global shifts occur at a faster rate than the mapping stage, where new information is gradually incorporated into the event model.

In terms of the temporal aspects of attention, if the effect of event boundaries is to cause momentary broadening of attentional breadth to aid in the updating of the event model then the effect of fine boundaries is stronger than coarse boundaries because the attentional effect is sustained over a longer period of time. Thus, if the ambient-to-focal shift account of attention at event boundaries is true, then coarse boundaries would represent a stronger, faster attentional shift, whereas fine boundaries do not broaden as much as coarse boundaries and require more time to transition from one event grain to the next. Because attention remained in ambient mode longer around fine boundaries than at coarse boundaries, front-end perceptual input provided scene-level information to WM over a longer period of time (Loschky et al., 2018). At the same time, less inhibitory control was exerted over attentional selection in the absence of an appropriate event framework (Gernsbacher et al., 1995) and the information provided to update the working memory model was at a coarser level than it would be in the event middle.

4.3 Event Perception and the Dual-Task Framework

An assumption motivating this study was the notion that changes in attention could be described with a dual-task framework. In the case of dual-task interference, two unrelated tasks require one to divide mental resources at perceptual, cognitive processing, and behavioral/motor output levels (Wickens, 2002). Thus, the strain on cognition comes from the discontinuity between the two tasks, which requires one to serially switch from one task to a second task when the needs of the second task arise (Pashler, 1994). In the case of dynamic event perception, SPECT proposes that during an event model shift, event information is recalled from episodic memory, while event information in working memory is encoded into episodic memory. Thus, the strain on cognitive resources may come from the need to simultaneously encode and consolidate new information to LTM while also retrieving accurate event information from LTM to update one's event model (Loschky et al., 2018). However, these three tasks are not completely independent from each other—in fact, they depend very much on one another. Therefore, it is perhaps the case that updates in event models are not analogous to dual-task interference in the traditional sense, where cognitive resources are allocated between multiple opponent tasks. Nevertheless, the statistical models for both the attentional impairment and ambient-to-focal models show support for subtle changes in attentional gain across the visual field, as do several other studies that have measured covert attention around critical changes in dynamic events (Crundall, Underwood, & Chapman, 2002; Huff, Papenmeier, & Zacks, 2012). However, the reasons for predicting changes in the *shape* of attentional breadth (i.e., tunneling or broadening) must be reconciled with the theoretical underpinnings of event perception in addition to needing further empirical evidence before it can be validated. The traditional dual task framework assumes that tunneling is a result of strained attentional demands (Williams, 1988), however these methods make the importance of the central (e.g., foveal) visual task an explicit priority. During naturalistic event perception, the predicted importance of specific object or location is *implicit* and controlled by the accuracy of the event model (Huff, Meitz, & Papenmeier, 2014).

The evidence for ambient-to-focal shifting of attention around the event boundary is currently tenuous, though the available data does reject one hypothesis regarding the shape of attention around the event boundary. It was previously suggested that attention might begin to tunnel around event boundaries. This was because, as the event becomes more uncertain, information regarding objects or agents within the event becomes more difficult to extract and requires a greater degree of attentional focus to shift from one event grain to the next (Nuthmann, Smith, Engbert, & Henderson, 2010). However, if changes in the size of attentional breadth were at all present, it seems that the opposite effect occurred: attention broadened at event boundaries and narrowed toward the event middle. This effect is consistent with SPECT (Loschky et al., 2018) which suggests that, since event boundaries represent moments of uncertainty, it is advantageous to expand attention over the entire scene, which will help to facilitate the transition from one event to the next. The simultaneous drop in attention suggests that attention was not only expanding, but that there is a lack of control from central executive resources in constraining the focus of attention as event information in WM is encoded and consolidated into LTM (Rayner, Sereno, Morris, Schmauder, & Clifton, 1989).

Furthermore, the reduction in attention at the event boundary puts into question how EST's gating mechanism works. Event Segmentation Theory suggests that the gating mechanism is opens at the onset of the event boundary and directly connects sensory input to event model information in WM (Zacks et al., 2007; Zacks & Sargent, 2010; Kurby & Zacks, 2008). This argument predicts that attention should be *better* at event boundaries since the circuit between early perception and working memory has been completed. However, the empirical evidence by this study and the results from Crundall et al. (2002) and Huff et al. (2012) consistently find impairment of attention, rather than facilitation. Huff, Papenmeier, & Zacks (2012) further argue that the decrease in attention is due to a complete reorienting of top-down attentional resources, given that their multiple object tracking paradigm did not correlate with low-level perceptual bottle-necks (e.g., crowding of the soccer players), and thus later cognitive processing constraints are necessary to explain attentional impairment. Instead, their explanation of attentional impairment is akin to the attentional blink where attentional processing for an earlier stimulus interferes with the processing of a later (but temporally close) second stimulus (Raymond, Shapiro, & Arnell, 1992). Thus, event boundaries could be a real-world example of the attentional blink where the recognition of event information immediately following the boundary is not detected because event information prior to the boundary is still being encoded (Crundall et al., 2002; Huff et al., 2012). Even more interesting is that the attentional blink peaks at approximately 200-300 ms (Chun & Potter, 1995) or roughly the average duration of a single eye-fixation (Rayner, 1998). According to SPECT, would mean that an event model update results in the loss of front-end information for one fixation. However, the brief suppression of new perceptual input following the event boundary could help facilitate the encoding of information at the event boundary.

Neuroimaging research has also been able to demonstrate some of the physiological underpinnings of the enhancement of event information at the beginning of the event model, by suggesting that cognitive processing is temporarily alleviated from interference between new sensory input and the storage of recent episodic information to LTM. In one study, hippocampal activation was monitored during the free-viewing of short film clips, which were followed by either a film clip, a pixel-scrambled film clip, or nothing (Ben-Yakov, Eshel, & Dudai, 2013). Anterior hippocampal activation that immediately followed viewing the videos was strongest when only one video was displayed, and with slightly less strong activation for the noise videos. The weakest anterior hippocampal activation occurred when participants were required to watch two normal film clips. The results from Ben-Yakov et al. (2013) also found that there were distinctive spikes in anterior hippocampal activation that occurred at the end of each video, regardless of whether it was followed by another video. Thus, even during event perception in which the event stimuli contain multiple continuous events, it is likely that the hippocampus undergoes periodic offloading of event information into long-term storage. However, this ability to encode new information can be impaired through interference from new, incoming sensory information. Therefore, the suppression of new sensory information at the event boundary via attention would be useful in reducing the degree of interference during memory encoding.

Ben-Yakov et al. (2013) also found that stronger anterior hippocampal activation that occurred when videos were followed by a blank screen also produced stronger posterior hippocampal activation during memory recall. The increase in posterior hippocampal activation during memory recall resulted in better subsequent memory. Therefore, it could be possible that *stronger* attentional suppression at coarse boundaries results in *faster* event model update because there is less interference by sensory input during retrieval from LTM. Meanwhile, fine event boundaries show weaker attentional impairment in *space* and slower shifting in *time*. Therefore, the spatial effects and the extended temporal effects of fine event boundaries could be explained by less suppression of sensory input, which could reduce the speed and accuracy at which the new event model can be updated.

Although a dual-task paradigm framework explanation of event model updating is quite distinct from the traditional context of an individual attempting to complete two distinct tasks, its basic principles are still relevant. Consciousness is supported by processes with limited resources—both in perception and memory—however we have adapted strategies that streamline the process and optimize performance.

4.4 Attention, Event Models, and Global versus Incremental Updating

In addition to the concurrent encoding and retrieval of information necessary during dynamic event perception, another question remains for how, when, and what information is updated between working and long-term memory. When a global event model shift occurs when an entirely event model is generated from sensory information (Kurby & Zacks, 2012; Zacks & Swallow, 2007). However, according to event indexing theory, event models can be updated incrementally when only a small portion of the event model is updated (Huff et al., 2014; Zwaan, Langston, & Graesser, 1995). Whether one updates incrementally or globally may depend on the number of event index changes (Huff, Meitz, & Papenmeier, 2014), or the types of index changes that occur (Zacks et al., 2010).

The current results have some trouble reconciling with the traditional assumptions behind the sensory gating mechanisms proposed by Event Segmentation Theory, which argues that the event model affects the selective filtering of information at later sensory processing stages, but sensory information does not affect the event model *until* the event boundary (Reynolds, Zacks, & Braver, 2007; Kurby & Zacks, 2011). The results from this study show that Gabor accuracy at coarse boundaries (79%) and fine boundaries (81%) is only slightly lower compared to non-boundary times (82%), so if a global sensory gate does exist, it is very weak. However, it is possible that the strength of the attentional manipulation (namely the event videos) were not sufficient to precipitate a strong enough change in attention. Thus, more arousing or salient video stimuli may be more effective in enhancing the effect of the events on attention. Likewise, manipulating the Gabor patches to appear for less than the window of the attentional blink might also provide greater sensitivity to attentional changes during event perception, especially when determining the difference in attentional impairment between coarse and fine boundaries.

If the sensory gating mechanism relatively weak, an alternative to EST’s sensory-event model gate would be for a gating mechanism to exist between working memory and long-term memory, with constant communication occurring between the event model in WM and new sensory input. Only when new sensory input sufficiently deviates from the event model is the connection between sensory information and WM (i.e., attention) broken, and the connection between LTM and WM reestablished so that novel event information can be incorporated with prior knowledge in LTM. Some models do suggest that this relationship exists. The Structure Building Theory (Gernsbacher, 1990) and SPECT (Loschky et al, 2018) both argue that after the foundation for the event model has been laid, new information is mapped on to the global structure. Similarly, the Event Indexing Model (Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998) argues that event information is decomposed into multiple dimensions which are gradually updated if new information is similar, and the likelihood of updating one's event model increases as more of these indices change. All of these models argue that new, incoming information is mapped onto existing event schemas in working memory during event perception. In particular, SPECT can help to explain how the differential perceptual processing of early, coarse information (e.g., gist) and later object information in the front-end affect back-end cognitive processing while laying the foundation for the event, and during subsequent mapping. Furthermore, the movement of broad to narrow processing in the ambient to focal shifting of attention provide an additional level of importance for constant monitoring of information at different levels of specificity (i.e., coarse to fine processing, respectively). Therefore, the changes in covert attention provide a window into the mechanics of event model shifting at event boundaries.

Because attentional impairment around event boundaries was relatively weak for both the attentional impairment and ambient-to-focal models, it is more likely that the more critical gating mechanism is in terms of attentional breadth (e.g., tunnel vision) and not overall attentional gain (e.g., general interference of attention). Compared to the few percentage point differences in overall Gabor accuracy among boundary and non-boundary periods, the evidence for Gabor accuracy as a function of eccentricity and time was much more distinctive for different boundary and non-boundary periods (see Fig. 13). To reiterate, the ambient-to-focal shift model found some evidence of attention remaining tunneled during non-boundaries, but quickly moved from broad-to-narrow during coarse boundaries. Specifically, the difference between central and peripheral was only 3-percentage point difference prior to the boundary and increased to an approximately 20 percentage point difference in the five seconds after the boundary. Fine boundaries showed slightly less broadened attention for longer periods of time. Thus, sensory gating during event perception may be in terms of attentional breadth. However, sizeable error in this model do not conclude with absolute certainty that the ambient-to-focal account of attention at event boundaries. Therefore, it must be emphasized that the merits of this model require further investigation before it can be validated.

A key component of SPECT argues that broad extraction of information (i.e., through the gist of the scene) is an important "front-end" process to facilitate the transition from one event to the next (Loschky et al., in press). The ambient-to-focal shift model proposes that the broadening of attention during event boundaries exists because the absence of a valid event model means that there is no context to guide inhibitory processes with constraining the focus of attention. It is also important to note that the methods used in this experiment have factored out low-level perceptual limitations of the visual system (e.g., cortical magnification, photoreceptor density, etc.), and thus the results reflect a change in the sensory gate that is specific to attention.

4.5 Structural and Physiological Models of Working Memory and Attention During Event Perception

There are many structural and physiological models that attempt to explain the relationship between attention, perception, and memory, though the primary division in the current literature has been between the continuum-based perspective of "long-term working memory" (Ericsson & Kintsch, 1995) and the more categorical multicomponent model (A. Baddeley, 2003; A. D. Baddeley & Hitch, 1974). However, developing a parsimonious, physiological representation of these memory hierarchies can become quite challenging when attempting to understand the multiple perceptual and conceptual dimensions in which information can be represented, and how these dimensions are nested within one another.

Hierarchical Process Memory (HPM; Hasson, Chen, & Honey, 2015) is a neurophysiological model which embodies the long-term working memory theory of Ericsson and Kintsch (1995), where WM and LTM are not dissociated from each other, and instead WM represents an active memory trace in LTM. An additional feature of HPM is that it specifies a relationship between the *temporal receptive window* (TRW, i.e., the length of time in which information can be accumulated) and the level of information processing. Specifically, early sensory areas accumulate information over short periods of time, and are thus updated frequently, whereas higher-order processes that are responsible for global (e.g., narrative) comprehension accumulate information over longer periods of time (Hasson, Chen, & Honey, 2015). Furthermore, the receptive field size for the early and late stages a direct relationship between receptive field size and TRW, with smaller receptive field sizes occurring for shorter TRWs and larger receptive field sizes occurring with longer TRWs (Himberger, Chien, & Honey, in Press). Thus, broader scene space is more resistant to small, sudden changes as it relates to higher-order processing and provides stable modulation of low-level perceptual input.

For instance, HPM predicts particularly shallow temporal normalization (i.e., long TRWs) for the temporal-occipital junction and the intraparietal sulcus (Himberger et al., in Press; Zhou, Benson, Kay, & Winawer, 2017). During free viewing of *The Red Balloon,* changes in goal orientation, a high-level event characteristic, resulted in strong reductions in intraparietal sulcus activation following coarse and fine boundaries (though the effect was stronger for coarse boundaries) (Zacks, Speer, Swallow, & Maley, 2010). Furthermore, these reductions in activation for IPS coincided with strong error signals from the anterior cingulate cortex, which is a crucial component for event model updating (Reynolds, Zacks, & Braver, 2007)[12]. Therefore, if (a) perception activates traces in LTM to form WM, and (b) higher level brain regions are less amenable to new information, a closed sensory gate would slowly starve global contextual hierarchies of relevant information, thus reducing activation of dorsal stream structures until a global update. . The error signal from the anterior cingulate would then be responsible for reopening the sensory gates to flood coarse-level systems with new information from global visual space. Thus, the event model can be reset to reflect a more accurate state of the visual environment.

For instance, although event boundaries are triggered by several important dimensions, like characters, location and goals (Zwaan & Radvansky, 1998; Zacks et al., 2007), these dimensions do not seem to equally affect the likelihood and type of event updating. Object-based changes are more likely to result in fine boundary segmentation, whereas changes in space and causality are more likely to result in coarse boundary segmentation (Kurby & Zacks, 2012; Speer, Zacks, & Reynolds, 2007). Thus, coarse scene information changes results in coarse

---

[12] However, error responses from subcortical dopamine systems are also critical for detecting prediction error. (Zacks, Kurby, Eisenberg, & Haroutunian, 2011).

boundary updates, whereas fine object information changes result smaller, fine boundary updates.

Furthermore, when attention is limited to only the coarse event information, it can impair the ability to attend to information at finer levels of specificity (but not necessarily boundary information). In one study, participants were told to read for specific types of content, where one group was instructed to attend to character shifts, while another group was instructed to attend to spatial shifts (Bailey, Kurby, Sargent, & Zacks, 2017). The *character change* group showed very little difference in accessibility of spatial and character information when compared to a control group, in which no specific instructions were given. However, the *spatial change* group showed substantial difficulty at accessing information at spatial changes and character-based non-changes in the narrative. These results can be explained in terms of the stages of processing involved with spatial properties of attention and working memory (Bailey, Kurby, Sargent, & Zacks, 2017; Kurby & Zacks, 2012). Spatial processing occurs at a coarser level of perceptual processing, compared to faces and objects, and by limiting attention to coarse spatial information, fine character information was not effectively maintained across the narrative. Thus, to the extent that spatial systems can help to contextualize information from a top-down perspective, it is also inherent in the fine information from the bottom-up but lacks the specificity for the finer distinctions that are necessary for detailed recognition.

If event memory is hierarchical (Hard, Tversky, & Lang, 2006), with conceptual information helping to organize the interpretation of current information and the expectations for future event information (Zacks, Tversky, & Iyer, 2001), then the predictions made by the ambient-to-focal model would suggest that the narrowing of attention is analogous to the narrowing of information activated in LTM. In this sense, the breadth of attention represents the breadth of information in working memory. when in the middle of an event, the event model can make specific predictions about what is about to occur next and attention is maximally narrowed. When those predictions produce errors over time, and when those errors are at a deep level of the memory hierarchy (e.g., space, goals, intentions, etc.), the event model must be updated globally (Reynolds, Zacks, & Braver, 2007; Zacks, Kurby, Eisenberg, & Haroutaunian, 2011). However, if there are only a few changes to the event model (Huff, Meitz, & Papenmeier, 2014), or those changes to the event are at a superficial level of the hierarchy, event updates are mapped onto the existing, long-term event structure (Gernsbacher, 1995; Loschky et al., 2018). Likewise, when larger increases in prediction error necessitate a larger (e.g., global) event model update, the event model shift would require updating the stored model at a deeper level of the LTM hierarchy. This would be similar to laying the foundation of a new event model (Gernsbacher, 1990; Loschky et al., 2018). When laying a new foundation for an event, the event model would be relatively unspecified, and attention would be minimally constrained until the appropriate model is selected. In this state, it would be advantageous for attention to cover a large area of visual space, or (with respect to Event Segmentation Theory) the sensory gate would be maximally opened to provide the broadest sample of visual information. Upon recognizing the gist of the environment, perceptual cues can help to narrow an appropriate event that is contextually and perceptually consistent with the coarse-level information (Larson, Hendry, & Loschky, 2012).

4.6 Event Memory and Perceptual Modalities

Event segmentation theory posits that event models constrain attention while also allowing viewers to make predictions about what should occur next in an event (Kurby & Zacks, 2008). Thus, the event model in working memory provides a buffer between prior knowledge in

LTM and new information from the senses. When predictions are valid, the new information is redundant with information in LTM and therefore there is no need for LTM and WM to interface. However, when new event information deviates from the event model in working memory, error signals increase in the anterior cingulate cortex (Zacks et al., 2010), and the event model in WM must be updated to improve prediction rates for new event information (Huff, Meitz, & Papenmeier, 2014; Reynolds, Zacks, & Braver, 2007; Zacks, Kurby, Eisenberg, & Haroutunian, 2011). Event model updating is also said to be a period in which sensory gating mechanisms between perception and LTM are opened to encode new event information, and to recall accurate event information. The novel event information that is consolidated into LTM is often constrained (and biased) by prior experience and expectations (Huff et al., 2017). However, there have been inconsistent findings regarding the importance of the types of information that cue coarse versus fine event model updating.

A common theme between the previous research that used the videos presented in the current study (Sargent et al., 2013; Kurby & Zacks, 2011) is that they all used the same memory probe stimuli, namely in the form of visual recognition memory. Generally, these studies found that fine event boundary agreement predicted better overall event recognition memory when participants were asked to discriminate between still images of videos that were either present or absent from the video they had seen (Kurby & Zacks, 2011; Zacks, Speer, Vettel, & Jacoby, 2006; Sargent et al., 2013). Event order memory (through ordering still images of the event video) has been slightly less consistent, however. In one study, fine segmentation agreement predicted picture order memory for older adults only (Zacks et al., 2006), whereas in two other studies, no relationship between fine segmentation agreement and event order recognition memory was found, regardless of age (Kurby & Zacks, 2011). In another study that used the same videos as the current study, segmentation agreement was not found to correlate with event order memory (Sargent et al., 2013). These results are inconsistent with the present study, which found event order memory to be predicted by coarse segmentation agreement, but not fine segmentation agreement.

The current experiment did replicate previous results for object recognition during event perception. Specifically, it demonstrated an advantage for boundary objects compared to non-boundary objects using conceptual recognition cues. In previous studies, participants showed very little difference in recognizing pictures of objects at boundaries versus non-boundaries, while a large memory advantage for boundary objects (versus non-boundary objects) was observed when the memory cues were conceptual (i.e., words; Swallow, Zacks, & Abrams, 2009). However, they did not find an effect for either fine or coarse segmentation agreement in predicting the advantage for boundary versus non-boundary objects.

4.7 Limitations

While the current study was able to show trends in favor of two non-mutually exclusive hypotheses, the best-fit models in the study were not able to differentiate between the two. Even more troubling was the differences between the two models' critical effects. For the attentional impairment hypothesis, the overall attentional effect between coarse, fine and non-boundary was small (≈3%), however the standard error among participants was quite small, and thus the effect was relatively consistent across participants. Conversely, the ambient to focal model showed strong differences in Gabor accuracy as a function of eccentricity and boundary offset time among the three event grains. However, the error among participants was quite large, and thus not as consistent of an effect. Nevertheless, two alternative hypotheses have become substantially less likely. First, it is very unlikely that improvements in memory around the event boundary can be explained by *increases* in attention, since both models show a slight drop in attention, and these results replicate those by Crundall et al. (2002) and Huff et al. (2012). Furthermore, it is very unlikely that attentional tunneling occurs at event boundaries since the current evidence shows that, if the size of the UFOV changes during dynamic event perception, it is most likely to broaden at the event boundary. Nevertheless, the question of which of the two hypotheses is more likely than the other is still unanswered and will require further research.

Several limitations occurred in this experiment, which must be addressed by future studies. The first limitation of this study was that despite a large sample size (n = 89), there were still a large proportion of participants with missing data in several conditions (e.g., video type, eccentricity, and boundary type). The low number of responses across participants was such that it made the complex random effects structures (both by-item and by-subject) computationally intensive (Betancourt & Girolami, 2013), especially with the inclusion of a three-way interaction in the fixed effects structure. Therefore, while it is important that future research include more videos, it is equally important that more attentional probes are included *within* each video so that a more stable estimate of variability across participants and videos can be obtained.

A second limitation of this study was that the individual segmentation data was not conducive to remapping the Gabor presentations to individualized event boundaries. Approximately 52% of participants segmented five times or less, which provides a very limited range of samples around the event boundary. During the experiment, some participants were unsure of when to segment, despite being given a practice segmentation video and a feedback counter to provide a hint to the experimenter and the participant as to whether they were completing the task correctly. This was to protect the data from contamination by the experimenter in guiding their responses and imposing unnatural uniformity in segmentation data.

In the future, additional shaping procedures will be necessary to ensure that participants clearly understand the instructions for segmentation in a way that is uniform and free from undue influence (Sargent et al., 2013; Zacks, 2004).

An additional limitation of the GC-UFOV method is that only one Gabor presentation per eccentricity is possible in time, which makes the presence of an unmasked Gabor even more detrimental for sampling attention around a limited number of event boundaries. Furthermore, the range of boundary offset times (i.e., the time around the event boundary) may have been too limited to capture a complete image of the changes in anticipation of an event boundary. For instance, the fine boundary showed sustained broadened attention around the event boundary, but there was no indication of when that change occurred, or whether it was a discrete shift. Likewise, the sampling period around coarse boundaries may have also not been early enough to estimate when attention broadens in anticipation of a coarse boundary. However, the current data were able to estimate when attention begins to narrow in response to the event boundary, which appears to begin approximately four seconds prior to the boundary time.

4.8 Future Research

This study was the first of its kind to estimate changes in attention—without conflating task performance with low-level visual sensitivity—during real-world event perception. Despite the lack of conclusive results for the GC-UFOV during event perception, there are several modifications to the current experimental paradigm that can be made to improve upon it. However, to ensure that the changes from one experiment to the next can be attributed to a single methodological change, these modifications must be implemented slowly and systematically (Platt, 1964).

The first follow-up study should make minimal changes to the experimental paradigm and should instead make use of informed prior distributions to determine if the trends found in the current study are stable and replicable. For complex hierarchical models, model stability indices like the WAIC and LOO can have difficulty showing any difference between the weak, flat prior distribution and posterior sampling, and thus the difference between the two similarly structured models will can be quite minimal (Vehtari et al., 2015). Importantly, Bayes factors for Bayesian hierarchical models can be very misleading when flat priors and (at times) when weakly informative prior distributions are specified (Vehtari et al., 2015). Instead, specifying informative priors for intercepts and slopes can greatly improve error estimations for the intercepts and slopes of future data (Vehtari et al., 2015). Furthermore, model stability can be improved with more observations within each level of fixed *and* random effects (Betancourt & Girolami, 2013). Therefore, this study would include more observations by flanking the event boundary with two observations. However, the temporal inter-presentation interval between Gabor patches would need to be randomized so that one Gabor patch does not serve as a pre-cue for another, while also providing an appropriate (minimum) buffer time between presentations so as to not detract from the event video. The increase in Gabor presentations would also serve as a fail-safe against the unmasked Gabor presentations, which cannot be used to accurately measure attention within this paradigm.

If it is the case that a second sample can provide more definitive evidence of either the *attentional impairment* or *ambient-to-focal* hypotheses, the next logical step would be to determine if the changes around coarse and fine boundaries for one set of videos will generalize to the next. A second study could determine if the posterior sampling distributions from this study can be used to generate prior distributions that generalize to other samples of participants and videos, however such a study would require multiple days of testing and would require additional (financial) incentives to reduce attrition. More engaging and surprising stimuli could also be used, however primary obstacle for the GC-UFOV method only allows for one observation per Gabor eccentricity, orientation (right vs. left), and boundary. So, it requires many trials to estimate attentional breadth, and participant engagement may remain an important consideration. Continually showing “surprising” stimuli could bias participants to become less sensitive to the surprises, however film could be a way to reduce variability in the perception of events while maintaining a high degree of engagement on the part of the participants (Loschky et al., 2015).Finally, a third study—independent of the GC-UFOV—could be implemented to determine the degree of global versus local perceptual processes that change during an event boundary, and *when* they occur. To date, virtually all neurophysiological research in event segmentation has been conducted using fMRI. Although fMRI has tremendous spatial resolution for functional brain activity, its temporal resolution is limited to approximately 1 s intervals (Lin et al., 2018). Alternatively, techniques like electroencephalography (EEG) can have a temporal resolution as fine as 1 ms, with spatial resolution being its primary tradeoff. One important limitation of EEG too consider is the propensity for overt motor activity to interfere with scalp electrical readings of brain activity, and this includes oculomotor behavior. Given that eye-movements are a necessary component to real-world perception and produce electrical activity in relatively low frequency channels (> 65 Hz; Yuval-Greenberg et al., 2008) measuring individual ERP components would yield quite noisy data. However, using steady-state topography (SST) EEG, one can measure changes in the brain at low-frequency bandwidths that are outside of the frequencies expressed by electrical activity from eye movements (Muthukumaraswamy, 2013), specifically in the α and β bands, which are observed at 8-12 Hz and 13-35 Hz, respectively

(Pfurtscheller & Aranibar, 1977). With regard to event perception, several experiments have been able to connect global inhibitory control of sustained attention to the α (Kelly, Lalor, Reilly, & Foxe, 2006) and β (Engel & Fries, 2010) bandwidths.

The Information via Desynchronization Hypothesis (Hanslmayr, Staudigl, & Fellner, 2012) has used research with Steady State Electroencephalography (SST EEG) to argue that reduced power in the α and β bandwidths impairs inhibitory control exerted by higher-level processes that would otherwise prevent new information from being encoded (Hanslmayer, Leipold, Pastotter, & Bauml, 2009; Klimesch, 1996). These low frequency bandwidths have been previously associated with improved memory for a given stimulus when they follow the presentation of that stimulus, thus α and β bandwidths may be reflective of a period of a global event update. Desynchronization of the α and β bands also produces higher γ bandwidth power, which is associated with greater integration of sensory information (Nyhus & Curran, 2010), and therefore may reflect perceptual processing activity at the sensory gate. An additional finding relevant to event perception is that the α power correlates negatively with activity in the cingulate cortex (Jann et al., 2009), an area of the brain responsible for error monitoring and that has shown strong activation in response to coarse event boundaries (Zacks et al, 2010). By gaining more sensitive temporal insights into event perception, it would then be possible to more accurately estimate *when* event boundaries are detected by viewers. If this sensitive, continuous measure of attention were to be paired with the GC-UFOV task, then a clearer relationship between event segmentation and changes in attentional breadth may be established.

Chapter 5 - Conclusions

Attention is the process that selects and enhances portions of our environment for in-depth analysis by cognitive systems. Real-world perception requires seamless transitions of attention from one moment to the next, when information from long-term memory is used to guide those shifts in attention. These shifts in attention occur at *event boundaries*, where meaningful changes in event information violate our expectations of what should occur next. Event Segmentation Theory (Zacks & Swallow, 2007) argues that when our expectations are violated, the windows of perception open to incorporate new sensory information with prior event information.

The results from this study do not support the notion that attention improves at these boundaries, however. If anything, overall attentional gain slightly *decreases* at event boundaries when compared to non-boundaries, which replicates previous findings from Crundall et al. (2002) and Huff et al. (2012). Nevertheless, there is evidence that the size of the attentional window increases at event boundaries, much like the aperture of a camera opens to expose film to light. Even more important was that evidence for change in the size of the attentional window was found despite the fact that low-level perceptual limitations (i.e., photoreceptor density, cortical magnification, etc.) had been neutralized by thresholding attentional stimuli to be equally discriminable from central to peripheral vision in the absence of any cognitive or perceptual load (Ringer et al., 2014; Ringer et al., 2016). Though the nature of systematic broadening and narrowing of attention still requires further experimentation to be effectively modeled, this study was among the first to find evidence of covert attentional tunneling and narrowing that can be attributed to real-world event structure.

References

Baayen, R. H., Davidson, D. J., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. . *Journal of Memory and Language, 59*(4), 390-412.

Bach, M. (2006). The Freiburg Visual Acuity Test-Variability unchanged by post-hoc re-analysis. *Graefe's Archive for Clinical and Experimental Ophthalmology, 245*(7), 965-971. doi:10.1007/s00417-006-0474-4

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829-839.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

Bailey, H., Kurby, C., Sargent, J. Q., & Zacks, J. M. (2017). Attentional focus affects how events are segmented and updated in narrative reading. *Memory and Cognition, 45*, 940-955.

Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: Expanding the useful field of view. *Journal of the Optical Society of America, 5*(12), 2210-2219. doi:10.1364/josaa.5.002210

Ball, K. K., Edwards, J. D., & Ross, L. A. (2007). The Impact of Speed of Processing Training on Cognitive and Everyday Functions. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 62*(Special Issue 1), 19-31.

Ball, K. K., Owsley, C., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1993). Visual attention problems as a predictor of vehicle crashes in older drivers. *Investigative Ophthalmology & Visual Science, 34*(11), 3110-3123.

Bates, D., & Venables, B. (2014). Regression spline functions and classes. *R-Package Version 2, 2*(1).

Ben-Yakov, A., Eshel, N., & Dudai, Y. (2013). Hippocampal immediate poststimulus activity in the encoding of consecutive naturalistic episodes. *Journal of Experimental Psychology: General, 142*(4), 1255.

Betancourt, M., & Girolami, M. (2013). Hamiltonian monte carlo for hierarchical models. *arXiv*, 1-11.

Blasing, B. E. (2015). Segmentation of dance movement effects of expertise, visual familiarity, motor experience and music. . *Frontiers in psychology, 5*, 1500.

Burkner, P. C. (2017). brms: An R package for Bayesian multilevel models using STAN. *Journal of Statistical Software, 80*(1), 1-28.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research, 51*(13), 1484-1525. doi:10.1016/j.visres.2011.04.012

Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. P. (2000). Spatial covert attention increases contrast sensitivity across the CSF: Support for signal enhancement. *Vision Research, 40*, 1203-1215. doi:http://dx.doi.org/10.1016/S0042-6989(00)00024-9

Chun, M. M., & Potter, M. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception & Performance, 21*(1), 109.

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

Crundall, D. E., Underwood, G., & Chapman, P. R. (1999). Driving experience and the functional field of view. *Perception, 28*, 1075-1087. doi:10.1068/p2894

Crundall, D. E., Underwood, G., & Chapman, P. R. (2002). Attending to the peripheral world while driving. *Applied Cognitive Psychology, 16*(4), 459-475. doi:10.1002/acp.806

Daniel, P. M., & Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *The Journal of physiology, 159*(2), 1961.

de Haas, B., Schwarzkopf, D. S., Anderson, E. J., & Rees, G. (2014). Perceptual load affects spatial tuning of neuronal populations in human early visual cortex. *Current Biology, 24*(2), R66-R67. doi:http://dx.doi.org/10.1016/j.cub.2013.11.061

DuBrow, S., & Davachi, L. (2016). Temporal binding within and across events. *Neurobiology of Learning and Memory, 134*, 107-114.

Eckstein, M. P., Drescher, B. A., & Shimozaki, S. S. (2006). Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychological Science, 17*(11), 973-980. doi:10.1111/j.1467-9280.2006.01815.x

Eisenberg, M. L., & Zacks, J. M. (2016). Ambient and focal visual processing of naturalistic activity. *Journal of Vision, 16*(2), 5: 1-12. doi:10.1167/16.2.5

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777-813.

Engel, A. K., & Fries, P. (2010). Beta-band oscillations: signaling the status quo? *Current opinion in neurobiology, 20,* 156-165.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), Apr 1995, 1211-1245.

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics, 40*(4), 225-240.

Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception & Performance, 11*(5), 583-597. doi:10.1037/0096-1523.11.5.583

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, 4*(12), 2379-2394.

Gaspar, J. G., Ward, N., Neider, M. B., Crowell, J., Carbonari, R., Kaczmarski, H., . . . Loschky, L. C. (2016). Measuring the useful field of view during simulated driving with gaze-contingent displays. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 58*(4), 630-641. doi:10.1177/0018720816642092

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis, 1*(3), 515-534.

Gelman, A., Goodrich, B., Gabry, J., & Ali, A. (2017). R-Squared for Bayesian regression models. .

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics, 2*(4), 1360-1383.

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception and Psychophysics, 67*(7), 1252-1268.

Gernsbacher, M. A. (1990). *Language comprehension as structure building* (Vol. xi). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(2), 245-262.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*(1), 20-25. doi:10.1016/0166-2236(92)90344-8

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371-395.

Hanslmayer, S., Leipold, P., Pastotter, B., & Bauml, K. H. (2009). Anticipatory signatures of voluntary memory suppression. *Journal of Neuroscience, 29*(9), 2742-2747.

Hanslmayr, S., Staudigl, T., & Fellner, M.-C. (2012). Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Frontiers in Human Neuroscience, 6*(74), 1-12.

Hard, B. M., Tversky, B., & Lang, D. S. (2006). Making sense of abstract events: Building event schemas. *Memory & Cognition, 34*(6), 1221-1235. doi:10.3758/bf03193267

Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: memory as an integral component of information processing. *Trends in Cognitive Science, 19*(6), 304-313.

Hegdé, J. (2008). Time course of visual perception: coarse-to-fine processing and beyond. *Progress in Neurobiology, 84*(8), 405-439.

Henderson, J. M. (1992). Object identification in context: The visual processing of natural scenes. *Canadian Journal of Psychology, 46*, 319-341.

Himberger, K. D., Chien, H.-Y., & Honey, C. J. (in Press). Principles of temporal processing across the cortical hierarchy. *Neuroscience*, 1-14.

Huff, M., Meitz, T. G., & Papenmeier, F. (2014). Changes in situation models modulate processes of event perception in audiovisual narratives. *40*(5), 1377–1388.

Huff, M., Papenmeier, F., Maurer, A. E., Meitz, T. G. K., Garsoffky, B., & Schwan, S. (2017). Fandom Biases Retrospective Judgments Not Perception. *7*(43083), 1-8. doi:10.1038/srep43083

Huff, M., Papenmeier, F., & Zacks, J. M. (2012). Visual target detection is impaired at event boundaries. *Visual Cognition, 20*(7), 848-864. doi:10.1080/13506285.2012.705359

Hutson, J. P., Smith, T. J., Magliano, J. P., & Loschky, L. C. (2017). What is the role of the film viewer? The effects of narrative comprehension and viewing task on gaze control in film. *Cognitive Research: Principles and Implications, 2*(1), 46. doi:10.1186/s41235-017-0080-5

Irwin, J. R., & McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research, 40*(3), 366-371.

Ito, J., Yukako, Y., Suzuki, M., Maldonado, P., Fujita, I., Hiroshi, T., & Grun, S. (2017). Switch from ambient to focal processing mode explains the dynamics of free viewing eye movements. *Scientific Reports, 7*(1082), 1-14.

Jahn, G., Oehme, A., Krems, J. F., & Gelau, C. (2005). Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. *Transportation research. Part E, Logistics and transportation review, 8*(3), 255-275.

Kaernbach, C. (1990). A single-interval adjustment-matrix (SIAM) procedure for unbiased adaptive testing. *The Journal of the Acoustical Society of America, 88*(6), 2645-2655.

Kane, M. J., Conway, A. R., Miura, T. K., & Colflesh, G. J. (2007). Working memory, attention control, and the N-back task: a question of construct validity. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 33(3), 615-622. doi:10.1037/0278-7393.33.3.615

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience, 17*(11), 4302-4311.

Kass, R. E., & Rafferty, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 89*(430), 773-795.

Kelly, S. P., Lalor, E. C., Reilly, R. B., & Foxe, J. J. (2006). Increases in alpha oscillatory power reflect an active retinotopic mechanism for distractor suppression during sustained visuospatial attention. . *Journal of Neurophysiology, 95*, 3844-3851.

Klimesch, W. (1996). Memory processes, brain oscillations, and EEG synchronization. . *International Journal Of Psychophysiology, 24*(1-2), 61-100.

Knowles, I., & Renka, R. J. (2012). Methods for numerical differentiation of noisy data. *Electronic Journal of Differential Equations, 21*, 235-246.

Kreitz, C., Furley, P., Memmert, D., & Simons, D. J. (2015). Working-memory performance is related to spatial breadth of attention. *Psychological Research*, *79*, 1034-1041.

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences, 12*(2), 72.

Kurby, C. A., & Zacks, J. M. (2011). Age differences in the perception of hierarchical structure in events. *Memory and Cognition, 39*, 75-91.

Kurby, C. A., & Zacks, J. M. (2012). Starting from scratch and building brick by brick in comprehension. *Memory & Cognition, 40*(5), 812-826. doi:10.3758/s13421-011-0179-8

Larson, A. M., Hendry, J., & Loschky, L. C. (2012). Scene gist meets event perception: The time course of scene gist and event recognition. *Journal of Vision, 12*(9), 1077. doi:10.1167/12.9.1077

Leibowitz, H. W., & Post, R. B. (1982). The two modes of processing concept and some implications. In J. Beck (Ed.), *Organization and representation in perception* (pp. 343-363). Hillsdale, NJ: Lawrence Erlbaum.

Lin, F. H., Polimeni, J. R., Lin, J. F. L., Tsai, K. W. K., Chu, Y. H., Wu, P. Y., & Kuo, W. J. (2018). Relative latency and temporal variability of hemodynamic responses at the human primary visual cortex. *Neuroimage, 164*, 194-201.

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science, 240*(4853), 740-749. doi:10.1126/science.3283936

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108*(2), 393-434.

Loschky, L. C., Hutson, J. P., Smith, M. E., Smith, T. J., & Magliano, J. P. (2018). Viewing Static Visual Narratives Through the Lens of the Scene Perception and Event Comprehension Theory (SPECT). In J. Laubrock, J. Wildfeuer, & A. Dunst (Eds.), *Empirical Comics Research: Digital, Multimodal, and Cognitive Methods* (pp. 217-238): Routledge.

Loschky, L. C., & Larson, A. M. (2010). The natural/man-made distinction is made prior to basic-level distinctions in scene gist processing. *Visual Cognition, 18*(4), 513-536.

Loschky, L. C., Larson, A. M., Magliano, J. P., & Smith, T. J. (2015). What Would Jaws Do? The Tyranny of Film and the Relationship between Gaze and Higher-Level Narrative Film Comprehension. *PLoS ONE, 10*(11), 1-23. doi:10.1371/journal.pone.0142474

Loschky, L. C., & McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multi-resolutional display. *Journal of Experimental Psychology: Applied, 8*(2), 99-117. doi:10.1037/1076-898X.8.2.99

Loschky, L. C., McConkie, G. W., Yang, J., & Miller, M. E. (2005). The limits of visual resolution in natural scene viewing. *Visual Cognition, 12*(6), 1057-1092. doi:10.1080/13506280444000652

Loschky, L. C., & Wolverton, G. S. (2007). How late can you update Gaze-contingent Multiresolutional Displays without detection? *Transactions on Multimedia Computing, Communications, and Applications, 3*(4), 1-10. doi:10.1145/1314303.1314310

Magliano, J. P., Kopp, K., Higgs, K., & Rapp, D. N. (2017). Filling in the Gaps: Memory Implications for Inferring Missing Content in Graphic Narratives. *Discourse Processes*, 1-14. doi:10.1080/0163853X.2015.1136870

Magliano, J. P., Kopp, K., McNerney, M. W., Radvansky, G. A., & Zacks, J. M. (2012). Aging and perceived event structure as a function of modality. *Aging, Neuropsychology, and Cognition, 19*(1-2), 264-282. doi:10.1080/13825585.2011.633159

Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin, 81*(12), 899-917.

McMains, S. A., & Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron, 42*(4), 677-686. doi:10.1016/S0896-6273(04)00263-6

Müller, M. M., & Hübner, R. (2002). Can the spotlight of attention be shaped like a donut? Evidence from steady-state visual evoked potentials. *Psychological Science, 13*(2), 119-124.

Müller, N. G., Bartelt, O. A., Donner, T. H., Villringer, A., & Brandt, S. A. (2003). A physiological correlate of the "zoom lens" of visual attention. *The Journal of Neuroscience, 23*(9), 3561-3565.

Muthukumaraswamy, S. (2013). High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations. *Frontiers in Human Neuroscience, 7*, 138.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29*(11), 1631-1647.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal Of Personality And Social Psychology, 28*(1), 28.

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology, 38*(5), 752-760. doi:10.1111/1469-8986.3850752

Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model of fixation durations in scene viewing. *Psychological Review, 117*(2), 382-405. doi:10.1037/a0018924

Nyhus, E., & Curran, T. (2010). Functional role of gamma and theta oscillations in episodic memory. *Neuroscience And Biobehavioral Reviews, 34*(7), 1023-1035.

Pannasch, S., Dornhoefer, S. M., Unema, P. J. A., & Velichkovsky, B. M. (2001). The omnipresent prolongation of visual fixations: saccades are inhibited by changes in situation and in subject's activity. *Vision Res, 41*(25-26), 3345-3351. doi:S0042-6989(01)00207-3 [pii]

Pannasch, S., Helmert, J. R., Roth, K., Herbold, A. K., & Walter, H. (2008). Visual fixation durations and saccade amplitudes: Shifting relationship in a variety of conditions. *Journal of Eye Movement Research, 2(2):4,* 1-19. doi:10.16910/jemr.2.2.4

Papenmeier, F. (2016). Determine Event Boundaries in Event Segmentation Experiments (Version 1.2.4): CRAN.

Pashler, H. (1994). Divided attention: Storing and classifying briefly presented objects. *Psychonomic Bulletin & Review, 1*(1), 115-118.

Patten, C. J. D., Kircher, A., Ostlund, J., & Nilsson, L. (2004). Using mobile telephones: cognitive workload and attention resource allocation. *Accident Analysis & Prevention, 36,* 341-350. doi:10.1016/S0001-4575(03)00014-9

Pfurtscheller, G., & Aranibar, A. (1977). Event-related cortical desynchronization detected by power measurments of scalp EEG. . *Electrencephalography and Clinical Neurophysiology, 42*(6), 817-826.

Piras, A., Lobeietti, R., & Squatrito, S. (2010). A study of saccadic eye movement dynamics in volleyball: comparison between athletes and non-athletes. *Journal of Sports Medicine and Physical Fitness, 50*(1), 99-108.

Platt, J. R. (1964). Strong inference. *Science, 146*(3642), 347-353.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3-25.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18*(3), 849.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422. doi:10.1037//0033-2909.124.3.372

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movements and on-line language comprehension processes. *Language & Cognitive Processes, 4*, 21-50.

Rayner, K., Smith, T. J., Malcolm, G. L., & Henderson, J. M. (2009). Eye movements and visual encoding during scene perception. *Psychological Science, 20*(1), 6-10. doi:10.1111/j.1467-9280.2008.02243.x

Reimer, B. (2009). Impact of cognitive task complexity on drivers' visual tunneling. *Transportation Research Record: Journal of the Transportation Research Board, 2138*(1), 13-19.

Reynolds, J. R., Zacks, J. L., & Braver, T. S. (2007). A computational model of event segmentation from perceptual prediction. *Cognitive Science, 31*, 613-643.

Rice, J., & Rosenblatt, M. (1983). Smoothing splines: Regression, derivatives and deconvolution. *The Annals of Statistics, 11*(1), 141-156.

Ringer, R. V., Johnson, A. P., Gaspar, J. G., Neider, M. B., Crowell, J., Kramer, A. F., & Loschky, L. C. (2014). Creating a new dynamic measure of the useful field of view using gaze-contingent displays. In P. Qvarfordt & D. W. Hansen (Eds.), *Proceedings of the Symposium on Eye Tracking Research and Applications* (pp. 59-66): ACM.

Ringer, R. V., Throneburg, Z., Johnson, A. P., Kramer, A. F., & Loschky, L. C. (2016). Impairing the useful field of view in natural scenes: Tunnel vision versus general interference. *Journal of Vision, 16*(2), 7:1-25. doi:10.1167/16.2.7

Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research, 37*(3), 495-510. doi:10.1007/bf00236819

Rovamo, J., Virsu, V., & Naesaenen, R. (1978). Cortical magnification factor predicts the photopic contrast sensitivity of peripheral vision. *Nature, 271*(5640), 54-56. doi:10.1038/271054a0

Sargent, J. Q., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Kurby, C. A., Bailey, H. R., . . . Beck, T. M. (2013). Event segmentation ability uniquely predicts event memory. *Cognition, 129*(2), 241-255. doi:http://dx.doi.org/10.1016/j.cognition.2013.07.002

Schulman, G. L., Tansy, A. P., Kincade, M., Petersen, S. E., McAvoy, M. P., & Corbetta, M. (2002). Reactivation of networks involved in preparatory states. *Cerebral Cortex, 12*(6), 590-600.

Schwartz, S., Vuilleumier, P., Hutton, C., Maravita, A., Dolan, R. J., & Driver, J. (2005). Attentional load and sensory competition in human vision: Modulation of fMRI responses by load at fixation during task-irrelevant stimulation in the peripheral visual field. *Cerebral Cortex, 15*(6), 770-786. doi:10.1093/cercor/bhh178

Shioiri, S. (1993). Postsaccadic processing of the retinal image during picture scanning. *Perception & Psychophysics, 53*(3), 305-314.

Smith, T. J., Whitwell, M., & Lee, L. (2006). *Eye movements and pupil dilation during event perception.* Paper presented at the 2006 Symposium on Eye Tracking Research & Applications.

Somers, D. C., & Sheremata, S. L. (2013). Attention maps in the brain. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(4), 327-340. doi:10.1002/wcs.1230

Speer, N. K., Swallow, K. M., & Zacks, J. M. (2003). Activation of human motion processing areas during event perception. *Cognitive and Affective Behavioral Neuroscience, 3*(4), 335-345.

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision, 11(5)*(13), 1-82. doi:10.1167/11.5.13

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General, 138*(2), 236-257. doi:2009-05547-006 [pii]

10.1037/a0015631

Theeuwes, J., Kramer, A. F., Hahn, S., Irwin, D. E., & Zelinsky, G. J. (1999). Influence of attentional capture on oculomotor control. *Journal of Experimental Psychology: Human Perception & Performance, 25*(6), 1595-1608.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*(4), 766-786.

Treisman, A. M. (2006). How the deployment of attention determines what we see. *Visual Cognition, 14*(4-8), 411-443. doi:10.1080/13506280500195250

Trevarthen, C. B. (1968). Two mechanisms of vision in primates. *Psychologische Forschung, 31*(4), 299-337. doi:10.1007/bf00422717

Unema, P. J. A., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition, 12*(3), 473-494. doi:10.1080/13506280444000409

Unsworth, N., Schrock, J. C., & Engle, R. (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.

Vehtari, A., Gelman, A., & Gabry, J. (2015). Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. *arXiv, 1507.04544v5*, 1-28.

Virsu, V., & Rovamo, J. (1979). Visual resolution, contrast sensitivity, and the cortical magnification factor. *Experimental Brain Research*, *37*(3), 475-494. doi:10.1007/BF00236818

Ward, N., Gaspar, J. G., Neider, M., Crowell, J. A., Carbonari, R., Kaczmarski, H., . . . Kramer, A. F. (2018). Older adult multitasking performance using a gaze-contingent useful field of view. *Human Factors, 60*(2), 236-247.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159-177. doi:10.1080/14639220210123806

Williams, L. J. (1988). Tunnel vision or general interference? Cognitive load and attentional bias are both important. *American Journal of Psychology*, *101*, 171-191.

Williams, L. J. (1989). Foveal load affects the functional field of view. *Human Performance, 2*, 1-28.

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science, 28*(6), 979-1008.

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., . . . Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience, 4*(6), 651-655. doi:10.1038/88486

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction error associated with the perceptual segmentation of naturalistic events. *Journal of Cognitive Neuroscience, 23*(12), 4057-4066. doi:10.1162/jocn_a_00078

Zacks, J. M., Speer, N. K., Swallow, K. M., & Maley, C. J. (2010). The brain's cutting-room floor: Segmentation of narrative cinema. *Frontiers in Human Neuroscience, 4*. doi:10.3389/fnhum.2010.00168

Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science, 16*(2), 80-84. doi:10.1111/j.1467-8721.2007.00480.x

Zacks, J. M., Swallow, K. M., Vettel, J. M., & McAvoy, M. P. (2006). Visual motion and the neural correlates of event perception. *Brain Research, 1076*(1), 150-162. doi:S0006-8993(05)01970-0 [pii]

10.1016/j.brainres.2005.12.122

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin, 127*(1), 3-21.

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology-General, 130*(1), 29-58.

Zelinsky, G. J., & Schmidt, J. (2009). An effect of referential scene constraint on search implies scene segmentation. *Visual Cognition, 17*(6-7), 1004-1028. doi:10.1080/13506280902764315

Zhou, J., Benson, N. C., Kay, K., & Winawer, J. (2017). Compressive temporal summation in human visual cortex. *Journal of Neuroscience*, 1724-1817.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 6*(5), 292-297. doi:10.1111/j.1467-9280.1995.tb00513.x

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*(2), 162-185.

1.