THE ROLE OF GIST PROCESSING IN BOUNDARY EXTENSION

by

AISHA PATRICE SIDDIQUI

(Under the Direction of James M. Brown)

ABSTRACT

Boundary extension is a ubiquitous phenomenon in which participants remember seeing more of an image than was previously shown. According to the Multisource Model of Scene Perception, viewers create a representation of their environment within a single fixation that integrates presented scene information with their expectations of what should exist there. The purpose of the following experiments was to determine the extent to which boundary extension relies on gist processing, and particularly, whether access to certain scene gist properties could be used to explain previous findings in the boundary extension literature. The gist of the scene is readily available at short timespans and was hypothesized to facilitate the existence of boundary extension by contributing schematic information about the scene into the representation. By shortening the encoding duration, access to the gist was gradually reduced to threshold levels. Experiment 1 was used to determine whether boundary extension could be elicited with natural complex scenes. The results from Experiment 1 showed that boundary extension could be elicited from complex natural, landscape scenes; however, it heavily depended on the participants' perceived depth rating and their perceived navigability of the scene. Experiment 2 manipulated access to the mean depth gist property and showed that boundary extension decreased as access to the gist was reduced, whereas Experiment 3 manipulated access to the navigability gist property and found that boundary extension increased as access to the gist was reduced. Experiment 2 provides evidence that the gist property depth is essential to the existence of boundary extension, making available schematic information to our scene representation which leads to increases in the source monitoring error. Experiment 3 shows evidence that navigability also provides schematic information to our scene representation however due to the adaptive function of the property it serves to provide veridical memory for the scene.

INDEX WORDS: Boundary Extension; Gist processing; Multisource Model of Scene Perception

THE ROLE OF GIST PROCESSING IN BOUNDARY EXTENSION

by

AISHA PATRICE SIDDIQUI

BS, University of Florida, 2007

MS, University of Georgia, 2009

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2013

THE ROLE OF GIST PROCESSING IN BOUNDARY EXTENSION

by

AISHA PATRICE SIDDIQUI

Major Professor:	James M. Brown
Committee:	B. Randy Hammond
	Dean Sabatinelli

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2013

ACKNOWLEDGEMENTS

Thank you to all my friends and family who have helped me along the way. A special thanks to Dr. James Brown for his support.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

As humans, we rely on our visual system to anticipate events or objects in our environment. This anticipation can be adaptive and advantageous in identifying possible predators and potential prey; however, it can also lead to errors. One example of such an error is boundary extension. Boundary extension is a ubiquitous phenomenon in which participants remember seeing more of an image than was previously shown (Intraub & Richardson, 1989). According to the Multisource Model of Scene Perception, viewers create a representation of their environment within a single fixation (e.g., 250 ms) that integrates presented scene information with their expectations of what should exist there (Intraub, 2010, 2012). Top-down processes add information to our scene representation from an egocentric perspective, suggesting our personal experiences (Intraub, 2010) and specific viewpoints (Gagnier, Intraub, Oliva, & Wolfe, 2011) are crucial to how we will extend (i.e., viewers will not add information that would hypothetically exist behind their view perspective). When the original image is removed from view, a source monitoring error occurs, in which subjects mistakenly report the constructed representation of their environment as the image they actually saw. Essentially, the Multisource Model of Scene Perception claims boundary extension is a result of routine scene perception.

If boundary extension is a result of scene perception, then at short time spans our constructed representation must rely on the integration of gist properties. The gist is an abstract, spatial representation of a scene clear enough to elicit and activate categorical information about the scene from memory (Friedman, 1979; Oliva & Torralba, 2001). Researchers have found viewers are capable of activating the scene schema (the gist) from an image in about 250 – 300 ms (Schyns & Oliva, 1994) and have recently shown the threshold for some of these global properties can be accessed in as little as 19 ms (e.g., "natural scene" vs. "man-made scene"; Greene & Oliva, 2009a). Consequently, activation of schema related information should increase the discrepancy between the actually presented image and the constructed representation, according to the multisource model. Access to gist related information could therefore facilitate the existence of boundary extension and perhaps could be the predominant mechanism underlying the effect. If so, eliminating or minimizing access to the gist should influence the magnitude of boundary extension. The purpose of the following experiments was to determine the extent to which boundary extension relies on gist processing, and particularly, whether access to certain scene gist properties can explain previous results in the literature. Additionally, the current experiments addressed the following questions: (1) can we elicit boundary extension using complex natural scenes, (2) can we elicit boundary extension at very short study timespans and (3) does the magnitude of boundary extension depend on access to a gist property?

Boundary Extension

Boundary extension was first described by Intraub and Richardson (1989). They found a consistent error in scene perception, in which participants reported their memory of a recently-viewed image contained more information than was shown, appearing as a zoomed-out version of the original. Participants viewed a series of 20 images for 15 s each and were instructed to remember as many details as possible. Half of the images were presented as close-angle versions, and the other half were presented as wide-angle versions. A close-angle version in this study mandated that one of the four picture boundaries cropped the main object; however, subsequent studies have shown that cropping of the main object is not necessary (Bertamini, Jones, Spooner,

& Hecht, 2005; Gagnier, Dickinson, & Intraub, 2013; Intraub, Bender, & Mangels, 1992; Intraub & Bodamer, 1993). When participants recreated the scene using a drawing task 48 hours later, boundary extension was found in 96% of the closer-up images and 87% of the wider-angle images (Intraub & Richardson, 1989). These results have been replicated numerous times, with retention intervals between study and test as short as 42 ms (Dickinson & Intraub, 2008), suggesting that access to a schematic spatial representation occurs very rapidly. Participants tend to include a proportionately greater amount of the background, and scale down the size of the objects and background to appear as if they hit “zoom-out” on their mental camera lens (Intraub, 1992; Intraub & Bodamer, 1993; Intraub & Richardson, 1989; Legault & Standing, 1992; Intraub, Gottesman, Wiley, Zuk, 1996). Even when a viewer is capable of physically moving the horizontal and vertical borders closer or farther away they show consistent boundary extension (borders farther out), rather than boundary restriction (closer in) or veridical memory for the scene (Intraub, Hoffman, Wetherhold, & Stoehs, 2006).

In addition to recall tests and psychophysical adjustment, boundary extension can be elicited using a recognition test, as well. The study phase for a recognition task is similar to the study phase in recall tasks, except half of the images presented at study are presented in the same view at test, and the other half are presented in the opposite view. In other words, four conditions are created: a close angle image shown at study with a close-angle image at test (CC); a wide-angle image at study with a wide-angle image at test (WW); a close-angle image at study with a wide-angle image at test (CW); and a wide-angle image at study with a close-angle image at test (WC). Subjects give their response on a 5-pt rating scale indicating whether the test image was *the same as the presentation image* (3), or whether the camera was a *little farther away* (4)*, a lot further away* (5)*, a little closer-up* (2)*, or a lot closer up* (1), followed by a confidence rating, *sure, pretty sure, not sure,* and *don't remember picture* (Intraub & Richardson, 1989). Some subsequent studies (e.g., Gottesman & Intraub, 2002; Intraub, Gottesman, Bills, 1998) slightly modified the wording for each value on the scale to include a reference to the main object: (1) *much closer up; object looks much bigger*; (2) *slightly closer-up; object looks slightly bigger*; (3) *the same*; (4) *slightly wider angle; object looks slightly smaller;* (5) *much more wide-angle; object looks much smaller*. The numbers are then converted to a -2 to +2 value for computation, in which images appearing to be the same are valued at 0.

When participants were presented with an identical picture at test and study (CC and WW), they tended to indicate the test image appeared closer up than before, despite being the exact same image (Intraub, 1992; Intraub et al, 1992; Intraub & Bodamer, 1993; Intraub & Richardson, 1989). For wider-angle images (WW), extension occurs to a lesser extent or magnitude, (Intraub et al., 1992; Intraub et al., 1998). When a distractor image was shown at test (the opposite view from the study image), participants tended to report a wider angle image at test (CW) was more similar to the study image, and a closer angle image at test (WC) as being too close. The defining quality of this result is that the absolute values of these responses were not equivalent; they were asymmetric despite the original images showing an equivalent change in view. Thus, the criteria for exhibiting boundary extension when both close and wide images are shown, can be thought of as (1) CC significantly less than 0 (when converted to a -2 to +2 scale), (2) WW significantly less than zero or not different from 0, and (3) an asymmetry between CW and WC.

Interestingly, boundary extension was still elicited when identical views were presented between study and test throughout the entire experiment (i.e., always showing the identical image at study and test). Gagnier and colleagues (2011, Exp. 2) used identical views at study and test and found an attenuated boundary extension effect. In this case, boundary extension was determined by analyzing how much of a change was reported at test from the identical image shown at study. Given that the presentation view never changed, a participant should indicate a boundary extension rating of zero (the image is the same) if their memory for the images were veridical; however, if a participant reported a value less than zero (the test image looks closer), then it would indicate that the participant boundary extended, and if they reported a value greater than zero, then it would indicate that the participant boundary restricted.

Multisource Model of Scene Perception

The multisource model of scene perception was proposed as an alternative to traditional visuo-cognitive models of visual scene representation. Traditional models suggest scene perception relies on visual scene representations that are created in a linear process, which gradually integrates visual input and semantic knowledge. Traditional models depend on various memory stores to maintain the representation (transsaccadic memory, visual short term memory, conceptual short term memory, and visual working memory), while semantic information elicits another representation that will be integrated at a later stage (Intraub, 2012). However, these models were incapable of explaining how boundary extension could occur so rapidly, at retention intervals of 42 ms (Dickinson & Intraub, 2008; Intraub, 2012; Intraub & Dickinson, 2008). In contrast, the multisource model simultaneously integrates multiple sources of information in an egocentric fashion. Instead of relying solely on visual input, the multisource model can utilize other sensory systems, such as the haptic system, to construct an egocentric spatial representation. In this way, we allow for a view of the world as we would typically encounter it.

The creation of an egocentric spatial representation is dependent on various inputs of information. Figure 1 shows a graphical representation of the multisource model (McDunn, Siddiqui, & Brown, 2013; Siddiqui, McDunn, & Brown, 2012). Sensory visual input provides information from the scene. The fidelity of this input is graded due to our visual system's optical limitations; information located foveally has the greatest acuity, whereas portions of the scene located peripherally have lower acuity. For example, a viewer might be presented with an image of a school. During the initial fixation (~250 ms), the clearest visual information will be located wherever the viewer has fixated (e.g., the door of the building). The clarity or resolution of the representation tapers off gradually towards the borders of the image, due to the resolution of our visual system. To "fill in the blank", or rather to account for the graded fidelity of the representation, we utilize amodal perceptual processes (i.e., completion or continuation). These processes complete possible partially occluded objects and allow for a perceived continuity beyond the borders of the image. For example, the viewer might complete the edge of the school house wall even though the visual input is not clear, or may understand that the wall continues around the corner despite not being able to see the wall.

In addition to amodal processes, schematic and conceptual knowledge of the scene are integrated into the image to provide further context as well as information about objects and their relations (Intraub, 2012). Beyond the edges of the image, the viewer can utilize stored schematic information about a school to enhance the representation further. These pieces of information tend to be congruent to the categorical schema; for example, the viewer might add a flagpole, or children playing on a playground, rather than surrounding the school with barbed wire, or disco lights. The activation of additional schematic and contextual information during this process is also integrated into our representation of the scene. Upon initial fixation, these additional inputs, or sources of information, become integrated into our scene representation, updating and refining it as subsequent fixations are made. At test, the viewer mistakenly chooses this representation with additional top-down information rather than the actual image, making an error that leads to boundary extension rather than restriction.

When participants mistakenly report seeing more of an image than was present (i.e., they exhibit boundary extension), it has been attributed to making an error in source monitoring (Intraub, 2010, 2012; Johnson, Hashtroudi & Lindsay, 1993), or more specifically an error in reality monitoring (Johnson & Raye, 1981). Source monitoring is a cognitive process involved in attributing the origin of memories, knowledge, and/or beliefs to a source. Disruption of source monitoring can be increased using divided attention tasks, or anything that hinders encoding or limits decision processes at retrieval (e.g., time pressure, severe stress, distraction, alcohol). Sources can have a spatial, temporal, and/or social context, as well as indicate quality of the sensory detail at encoding. Source monitoring is based on characteristics of memory that were established when the initial memory trace was formed. The characteristics of a memory include perceptual sources (sound/color), contextual sources (spatial/temporal), semantic detail, affective information (emotional reactions), and the cognitive operations involved (record of organizing, elaborating, retrieving, identifying).

A source monitoring error occurs when a subject misattributes one of these characteristics to the wrong source. For instance, in the classic “Becoming Famous Overnight” study (Jacoby, Kelley, Brown, & Jasechko, 1989) participants attributed fame to previously seen images of non-famous people. Subjects were shown a series of non-famous and famous names and were asked to indicate whether they belonged to a famous person or not. Most people did very well on this task. Twenty-four hours later, they were shown some of the same non-famous names that they correctly identified the day before. However, now the subjects indicated the non-famous names were famous. The subjects could remember having seen the name but not the context of where they had seen the name previously and assumed it was because they were famous. The subjects misattributed familiarity to fame.

Source monitoring attempts to distinguish between internally generated sources and externally generated sources (Johnson et al., 1993). In external source monitoring, participants must discriminate between two externally generated pieces of information. For example, when a person is trying to determine whether they heard some important piece of news from CNN or FOX news, they are discriminating between two external sources. Internal source monitoring occurs when a person must discriminate between two internally generated pieces of information, such as determining whether a memory actually occurred versus the person only dreamed about it (*cf*. Johnson et al., 1993, footnote #1). Reality monitoring describes a situation in which externally generated sources and internally generated sources are evaluated; it is responsible for discriminating between memories for thoughts and imagination versus memory for perceived events (Johnson et al., 1993; Johnson & Raye, 1981).

A source monitoring error, or more specifically a reality monitoring error, occurs when the two sources become confused. There becomes a discrepancy between the actual source of information (external) and what one might have thought they perceived (internal). This scenario is most closely related to boundary extension. The initial image shown serves as the externally generated source, and the representation containing additional top down schematic information of the scene becomes the internally generated source. When there are discrepancies between these sources of information a reality monitoring error occurs, and thus boundary extension occurs. Upon removal of the visual image, a person is unable to separate the original sensory input from the internally generated detail with which it was combined and the subject misattributes their memory of the scene as containing more information (Intraub, 2012).

Consequently, the predominant explanation for boundary extension is that it is a result of participants consistently relying on the constructed representation rather than a veridical representation.

Memory and Gist Processing in Scene Perception

During scene perception, access to the gist of a scene is readily available and contributes to our representation by providing schematic and categorical information. In simple terms, the "gist" of a scene contains enough information to give viewers a basic understanding of their environment. The gist is elicited at very short timespans and the information obtained within this glance can provide the viewer with both conceptual and perceptual sources of information (Oliva, 2005). The extraction of perceptual information encompasses structural features of the scene, including both low and intermediate visual information (e.g., color, contours, shapes, textures, etc.), whereas the extraction of conceptual information consists of inferred semantic and schematic knowledge about a scene.

Schyns and Oliva (1994) showed participants could utilize coarse perceptual information, available via lower spatial frequencies, to activate scene schemas and obtain the gist of a scene at very short timespans. The researchers showed participants a series of images for either 30 ms (short) or 150 ms (long) from four semantic categories (e.g., highway, city, living room, or valley). These images were low-pass filtered, high-pass filtered, normal (unfiltered), or hybrid combinations of two semantic categories (one of which was high-pass, and the other of which was low-pass filtered). Following image presentation, participants viewed a target image (always normal) and made a yes/no judgment as to whether the target matched the category of the previous image set. During the short time span, participants viewed the hybrid image as matching the category label of the low-pass images, whereas during the long time span they perceived the same hybrid image as matching the category of the high-pass images. Thus, when participants view a scene they initially rely on coarse perceptual information to obtain the gist of a scene, followed by a fine-grained inspection of the scene.

Viewers also use conceptual gist information to facilitate scene processing. Potter and Levy (1969) found participants were very good at identifying a target image when it was presented alone, but when presented along with a series of images rapidly presented (Rapid Serial Visual Presentation, RSVP) they performed poorly (e.g., only about 11% of images were remembered correctly after 125 ms). On the other hand, when participants were shown an image followed by a visual noise mask, Potter (1976) observed that participants could remember a single picture presented for 50 ms about half of the time, but now the probability of remembering the item increased to about 80% of the time when the picture was presented for 120 ms. Potter and colleagues argued that processing of the image in a rapid stream was conceptually masked by the other images, and the gist could not be consolidated before the presentation of a new image, whereas a visual noise mask allows for the conceptual gist to still be accessed. Interestingly, participants can use the conceptual gist of a scene to focus attention on the target item. Performance on an RSVP task can be improved if viewers are given the picture target or a verbal description of the target prior to viewing the stream (Potter, 1976). Ultimately, viewers have access to the perceptual and the conceptual gist of a scene very rapidly, and can use the varying pieces of information to identify and remember the scene.

In fact, Greene and Oliva (2009a) found threshold differences in presentation time for discriminating between different types of gist information, specifically global perceptual properties. Their experiment tested different global properties identified by Oliva and Torralba (2001) as important and crucial to scene understanding: concealment, mean depth, naturalness, navigability, openness, temperature, and transience (see Appendix A for distinctions). Each of the global properties have distinct statistical regularities (e.g., spatial frequency orientation and scale, color, texture density, affordances, etc.) in their scene content that enables automatic semantic activation and contributes to reliable scene categorization (Oliva & Torralba, 2001; Oliva, 2005). The images tested were piloted to exhibit either a high or low degree of the property's characteristics by selecting images ranked in the upper or lower quartiles of their respective global property (for ranking procedures, see Greene & Oliva, 2009b). High degree images (i.e., targets) contained perceptual consistencies prototypical of a given global property, whereas low degree images (i.e., distractors) represented images in the opposite extreme of a given property, containing a different set of perceptual consistencies (Oliva & Torralba, 2001), indicating images with a high degree of depth characteristics would elicit different energy spectra than images with a low degree of depth.

Each test image was followed by four texture synthesized color masks shown for 20 ms each (80 ms total) and participants gave a yes/no rating as quickly and accurately as possible to indicate whether the presented image exhibited a high degree of the given global property (Greene & Oliva, 2009a). The researchers used a 3-up/1-down staircase procedure to measure threshold accuracy. The first image was always shown for 50 ms, increasing 10 ms if the response was incorrect (200 ms ceiling) and decreasing 30 ms if correct (floor of 10 ms). Threshold was determined when accuracy reached 75%. Participants had access to global properties, such as determining the mean depth of an image (i.e., whether the scene contained a great amount of depth, depicting kilometers of space, as opposed to a few meters of space), in as little as 26 ms (sd = 2.8). They achieved 100% accuracy at a maximum exposure duration of 75 ms (sd = 4.9). In contrast, participants needed at least 36 ms (sd = 4.5) to achieve threshold in determining the navigability of a scene (i.e., whether a scene appeared to contain a clear path, compared with an obstructed one). They achieved maximum exposure at a duration of 120 ms (*sd* = 9.2). In totality, their study showed that we have access to gist properties at early stages in scene processing and it varies depending on the type of gist property. Given our ready access to gist properties and their respective contribution of schematic information to the scene representation, it is likely that specific gist properties play an essential role in the existence of boundary extension.

Boundary Extension as a Gist Processing Error

Boundary extension is so ubiquitous that only a few instances exist showing an elimination or a reduction in magnitude. For instance, Intraub and Bodamer (1993) explicitly warned their subjects to pay close attention to details at the object and border, and in some conditions demonstrated what boundary extension was to the subjects. They found boundary extension could be reduced in the demonstration condition but never eliminated. In fact, only two limitations to the effect have been described: the degree of extension is reduced as an image gets wider angle, and the proxy image shown must represent an abbreviated view of a continuous world (Gottesman & Intraub, 2003).

The first limitation suggests the degree of extension decreases as an image gets more wide-angle. When comparing the magnitude of boundary extension for close-angle and wide-angle images, wide-angle images typically elicit less boundary extension (Intraub et al., 1992). One explanation given as to why wide angle images elicit less extension suggests very wide angle images already contain enough information about the attended object and its surroundings. In turn, the viewer does not have any further expectations for continuity about the scene.

Perhaps, it has more to do with the type of gist information available to the viewer as an image becomes wider angle, which could be changing and activating different schematic information.

How one determines the distinction between a close and wide-angle image, however, is arbitrarily determined by the author; there is no established metric for measuring close and wide angles in a scene. Siddiqui et al. (2012) attempted to standardize the change in view angle using abstract shapes on random dot backgrounds, which controlled for changing semantic content as view angle changed. They found that as the difference in view angle increased (i.e., as the difference between close and wide angle images increased), less boundary extension was elicited, specifically images with a 40% relative size change elicited a regression to the mean size difference rather than extension, indicating close images boundary extend and wide image boundary restrict. Therefore, when close and wide angles were used in the following experiments, a standard percent change (16%) in relative size was utilized. Possibly, *mean depth* of an image might serve as a better metric for calculating the difference between close and wide angle images.

The suggestion that wider angle images contain enough information about the attended object and its surrounding is made with the assumption that an image has a singular main object on some background (Gottesman & Intraub, 2003), rather than a complex meaningful scene. Actually, most images used in boundary extension studies utilize simple images with a single salient foreground object on a large background surface (Gagnier et al., 2011). One exception was a study looking at the effect of clutter on boundary extension (Gottesman, 2011). Participants were presented with cluttered (e.g. rugged rocky shore, busy city street) and uncluttered (e.g., sandy beach, open field) scenes. Cluttered images elicited more boundary extension than uncluttered scenes suggesting the spatial scene layout of an image is pertinent to extension. These results also suggested that utilizing complex natural scenes may not elicit boundary extension if they were perceived to be clutter free or easy to navigate through. Unfortunately, in this paradigm the construct “clutter” was poorly defined. Clutter was defined “as the amount of perceptual detail present in a scene”, but does not reference anything more about the concept. Given the examples cited (e.g., sandy beach or rocky shore), it is possible the author was describing the *navigability* of a scene, however the images probably varied in a range of global gist properties presented such as *mean depth* (Greene & Oliva, 2009a, 2009b). Greene and Oliva (2009b) define *navigability* as corresponding “to the ease of self-propelled movement through the scene” (p.140) and ranges from a complete inability to pass through a scene due to obstacles, clutter, or treacherous conditions, to complete free movement in any direction throughout the scene. They define *mean depth* as corresponding “to the scale or size of space” (p. 140), ranging from close-up views to more panoramic scenes. It is possible Gottesman (2011) confounded uncluttered scenes as being more wide angle or panoramic, which would also decrease boundary extension. Together, Exps. 2 and 3 served to disentangle these results by manipulating images based on the gist properties of mean depth and navigability, respectively.

The second limitation for producing boundary extension suggests the presented image must depict a truncated view of a scene, in that it represents a proxy view of a continuous world (Gottesman & Intraub, 2003). In other words, in order to extend the boundaries of a scene the participant must expect the world continues beyond the edge of the borders. Intraub et al. (1998) removed the background of a scene and presented outline drawings of the main objects alone on a white background. They eliminated the expectation that the outlined object existed in a real, continuous scene and almost eliminated boundary extension. In this case, CC trials appeared closer, but WW appeared farther away (i.e., the same wide angle image appears wider angle), but the distractors (CW and WC) exhibited an asymmetric response (see criteria for boundary extension discussed earlier). Interestingly, they were able to elicit boundary extension with the same images if they asked participants to imagine the outlined object existed in a real scene. Hypothetically, imagining of a background provided additional internally generated inputs and caused a greater discrepancy in the original view and the constructed representation. This greater discrepancy leads to a larger source monitoring error, and hence greater boundary extension.

Recently, Intraub (2010, 2012) suggested that study designs (or stimuli) that lead to poorer memory of a scene should lead to greater boundary extension. Similar to tasks that lead to the disruption of source monitoring, divided attention tasks or anything that hinders encoding or limits decision processes at retrieval (e.g., time pressure, severe stress, distraction, alcohol) should therefore lead to increases in boundary extension. Using a divided attention task, Intraub, Daniels, Horowitz, and Wolfe (2008) showed that boundary extension was greater when attention was divided than when it was not. According to the authors, focused attention should lead to better encoding of and more veridical memory for source material and thereby less boundary extension (i.e., less of a source monitoring error). However, when attention becomes divided, the source information becomes fuzzier, or less clear, leading to a greater source monitoring error and boundary extension.

If poorer memory for scenes leads to greater boundary extension, then more boundary extension should also be elicited when encoding time decreases. Previous research has shown that subjects can recognize about 80% of masked images presented for 120 ms, but only about half of those images when presented for 50 ms (Potter, 1976, 2012). Thus, according to the multisource model, reducing this stimulus parameter – temporal access – should lead to greater boundary extension. As study presentation time decreases, source information becomes degraded, or fuzzy, causing a decrease in memory, and boundary extension should increase. Intraub et al. (1996) presented color images of single objects on simple natural backgrounds. They presented the study images for 4s or 250 ms, and measured boundary extension using both a drawing task, and the 5-pt recognition scale. In both tasks, they found the shorter encoding time (250 ms) tended to elicit a greater degree of boundary extension than the longer encoding time (4s). Interestingly, the shortest presentation time reported in a boundary extension study has been 250 ms[1]. It is possible that boundary extension increases as encoding time decreases because of an increased reliance on gist processing at shorter timespans (e.g., Brainerd & Reyna, 2002). As presentation time decreases, subjects tend to rely more on course perceptual information (i.e., gist information; Schyns & Oliva, 1994). The additional reliance on gist information to the constructed scene representation would further increase the discrepancy between the original sensory image and the internally generated representation, thus leading to greater boundary extension.

Current Study

The purpose of the following experiments was to determine the extent to which boundary extension relies on gist processing, and particularly, whether access to certain scene gist properties explains previous results in the boundary extension literature. Specifically, the experiments explored how the mean depth of an image could explain previous results showing wider angle images elicit less boundary extension, and how navigability of a scene might be used to further clarify the effects of clutter. Additionally, the experiments attempted to answer the following questions: (1) does boundary extension occur when viewing complex natural scenes? (2) can boundary extension occur at shorter presentation times, and is there a temporal limit to how quickly boundary extension happens? And (3) does the magnitude of boundary extension depend on access to a gist property?

If boundary extension can be elicited at shortened time frames, which have never been tested before, it is possible the mechanism that allows boundary extension to occur is facilitated by the rapid automatic extraction of gist features. To determine this, we varied the inherent global properties of a scene and biased subjects to encode only one of the gist properties using an incidental encoding task (e.g., a perceptual judgment task). Then, in order to manipulate access to the gist property, the images were presented at various durations, including the average threshold for accessing that specific property. If boundary extension relies on the rapid automatic extraction of gist properties, when access to the gist is eliminated or reduced, there should be a noticeable decrease in the amount of boundary extension. Given previous research in boundary extension concerning the proxy view angle of an image and clutter, the gist properties of *mean depth* and *navigability* were chosen, respectively. Recall that Greene and Oliva (2009a) found threshold differences in presentation time for discriminating between different global perceptual properties. Participants needed 26 ms (*sd* = 2.8) to achieve 75% accuracy in determining the *mean depth* of an image (i.e., whether the scene takes up kilometers of space or less than a few meters of space), with a maximum exposure duration of 75 ms (*sd* = 4.9). In determining the *navigability* of a scene (i.e., whether a scene appears to contain a clear path or an obstructed one), participants needed 36 ms (*sd* = 4.5) to achieve 75% accuracy, with maximum exposure at a duration of 120 ms (*sd* = 9.2). However, in order to vary global properties of a scene, one must be able to utilize more complex images than previously used in boundary extension studies. Therefore, the purpose of Exp. 1 was to evaluate whether boundary extension could be elicited using more complex scenes that were more similar to the images used in Greene and Oliva (2009a).

CHAPTER 2

EXPERIMENT 1

Method

Participants

Participants came from the University of Georgia research participant pool. There were 42 (29 female) participants. All participants signed informed consents, had normal or corrected-to-normal vision (self-described and verified using a visual acuity test), and indicated no history of an attention deficit disorder (self-described).

Stimuli and Apparatus

All images were adjusted and presented at a 256 x 256 pixel resolution, similar to Greene and Oliva (2009a) and shown in 24-bit color. Images were presented on an 18-in CRT monitor (85 Hz) using E-Prime v. 2. Participants sat 92.5 cm from the monitor, and images were centrally located, subtending a visual angle of 7° x 7°. A mixture of 260 images were taken and presented from the SUN database (Scene UNderstanding; Xiao, Hays, Ehringer, Oliva, & Torralba, 2010).

Procedure

Participants were situated in a darkened room with a chin rest on a desk, located 92.5 cm from a computer monitor. The experimenter remained in the room while the instructions were read by the participant, ensuring the participant understood the task.

Experiment 1a. The purpose of this experiment was to indicate whether, and to what extent, boundary extension could occur with complex scenes, not necessarily containing a singular, central main object. Each trial began with a centrally located fixation mark in order to orient the subject to the center of the image. Similar to the masking procedure used in Intraub and Dickinson (2008), participants initiated the trial with a press of the spacebar, while the fixation mark was on the screen. Subsequently, the study image was presented for 257 ms, followed by a dynamic, colored mosaic mask. The first mask appeared for 151 ms, then switched to another mask for 105 ms (total of 256 ms), similar to Intraub and Dickinson (2008). Finally the test image was presented on the screen until a response was given (see Fig. 2). Half of the images were presented at study in a close-angle view and the other half at a wide-angle view (see Fig. 2 & 3).

At test, the images were shown in either the same view, or in the opposite view, creating four presentation states (CC- close at study, close at test; WW- wide at study, wide at test; CW- close at study, wide at test; WC- wide at study, close at test). Participants gave a boundary extension rating of the second image on the following 5-pt scale: 1 = much closer up; object looks much bigger; 2 = slightly closer-up; object looks slightly bigger; 3 = the same; 4 = slightly wider angle; object looks slightly smaller; 5 = much more wide-angle; object looks much smaller. After the boundary extension rating was given, participants made a confidence judgment in their response, in which 3 = sure, 2 = pretty sure, 1 = not sure, and 4 = don’t remember picture. Trials given a rating of 4 (don’t remember picture) by the participant were excluded from analyses. This procedure was comparable to previous studies investigating boundary extension (Siddiqui et al., 2012, 2013).

Experiment 1b. After all boundary extension ratings were given, participants viewed the same images again and gave two global perceptual ratings about the image. They rated the perceived depth of the scene and the navigability of the scene on a 9-pt scale. Scores close to 1 indicated the scene showed a low degree of that property and scores close to 9 indicated the scene showed a high degree of that property. Images high in *mean depth* would be described as a scene that "takes up kilometers of space" (9), as compared with a scene taking up "less than a few meters of space" (1). Images high in *navigability* would be described as a scene containing "an obvious path that is free of obstacles" (9), as compared with a scene containing "many obstacles, or difficult terrain" (1). These values were used to determine if an individual's perceived rating of an image influenced boundary extension differentially.

Results and Discussion

Experiment 1a

As previously stated, if boundary extension exists, the following criteria have been suggested to occur: (1) CC should be significantly less than zero (i.e., participants should indicate the test image looked closer than it did before, rather than indicating they were viewing the identical image, or a wider angle image); (2) WW should be significantly less than or equal to zero (i.e., participants should indicate the test image looked closer than it did before, or indicate they were viewing the same image, rather than viewing a wider angle image); and (3) the absolute values of CW and WC should be significantly different from each other. Given that the amount of change between close angle images and wide angle images are the same (16% size change), participants making veridical judgments about CW and WC should have identical perceived change magnitudes, therefore a significant difference between the absolute values of CW and WC should indicate an error in memory and hence boundary extension.

One sample t-tests were used to evaluate whether CC and WW were significantly different than zero (see Table 1 and Fig. 4 & 5, black bar). Presentation state CC was not significantly different from zero, $t(41) = -1.89$, $p > .05$, $M = -0.08$, $SE = .04$, indicating that on average participants did not extend image boundaries and could veridically report seeing identical versions of the same image at study and test. However, presentation state WW was significantly different from zero, $t(41) = -2.03$, $p < .05$, $M = -0.09$, $SE = .04$, indicating, on average, participants were exhibiting some boundary extension by reporting the image at test looked closer up than the image at study, despite being identical. The interesting thing to note about these results is that they run contrary to a normalization account of memory. A normalization account suggests memory should regress to the average image size presented; for instance, wide angle images should show boundary restriction towards smaller angle view, and close angle images should show boundary extension towards a larger angle view (see Intraub et al., 1992; Intraub et al., 1998). However, in the case of these data, memory for close angle scenes was veridical rather than extended, and memory for wide angle scenes showed boundary extension not restriction, suggesting additional factors could be accounting for these results.

A paired- sample *t*-test was used to evaluate whether the absolute values of CW and WC were significantly different from each other, $t(41) = -2.2$, $p < .05$, $M = 0.37$, $SE = 0.6$ and $M = -0.54$, $SE = 0.7$, respectively. The results show the absolute values of CW and WC were significantly different from each other, indicating specifically that participants were more likely to report that a close-angle image at test looked much closer than a wide-angle image at test looked wider.

Experiment 1b

The global rating judgments made by participants were used to evaluate whether participants' boundary extension rating depended on their perceived depth or perceived navigability of an image. These perceived judgments should provide insight as to whether participants utilize these specific gist properties. Based on the value given on the 9 pt scale, responses were rank split into upper and lower quartiles. In this way, only images rated in the upper (high degree) or lower (low degree) quartiles based on the overall mean were analyzed. Given the subjective nature of this response, 28 subjects were included in the analyses for high depth, 41 subjects in low depth, 39 in high navigability, and 40 in low navigability.

Perceived depth. One sample *t*-tests indicated presentation states CC and WW in the low depth condition were significantly different from zero, $t(40) = -4.24$, $p < .0001$, $M = -0.28$, $SE = .07$, and $t(40) = -3.88$, $p < .0001$, $M = -0.24$, $SE = .06$, respectively (see Table 1 for all means). However, significant differences were not found in the high depth condition, *p's* > .05. Using paired sample *t*-tests, only the low depth condition compared with the high depth condition reached a significant asymmetry between CW and WC, $t(39) = -3.1$, $p < .01$. In other words, only the low depth conditions met the aforementioned criteria for exhibiting boundary extension. These results provide insight into the results found in Exp. 1a, namely, when the participants perceived the depth to be great, boundary extension was not found and when the participants perceived the depth to be small, boundary extension was exhibited.

A 2 (degree: high, low) x 4 (presentation state: CC, WW, CW, WC) repeated measures ANOVA was used to determine whether perceived depth influenced participants' boundary extension ratings differentially. A main effect of degree was found, $F(1, 27) = 21.91$, $p < .0001$, $\eta^2 = 0.45$. Overall, when images were rated as having a great amount of depth, participants tended to give more veridical responses than when images were rated as having less depth. As shown in Fig. 4, the light gray bars in the middle of each set, representing the high depth condition, were consistently greater than the dark gray bars on the right of each set, representing the low depth condition. The black bars on the left of each group represent the complete data from Exp. 1a. More boundary extension was found for images given a lower depth rating, keeping in mind that smaller values indicate test images were perceived closer than study images. These findings are consistent with previous studies showing a difference in boundary extension for close and wide angle images (Intraub et al., 1992), assuming close angle images are indicative of containing less depth and wide-angle images are indicative of containing a great amount of depth. In addition to supporting previous findings, these results would be one of the first exploring how an individual's perceived content influences boundary extension as well.

Furthermore, a main effect of presentation state was also found, $F(3,81) = 48.87$, $p < .0001$, $\eta^2 = 0.64$, suggesting that overall one or more of the presentation states were significantly different than the others. Specifically, presentation state CW was rated more positively than the other presentation states, *p's* < .0001, presentation state WC was rated more negatively than the others, *p's* < .0001, and presentation states CC and WW were not significantly different from each other, *p*'s > .05. Simply put, these results show a consistent pattern of responding when presenting different images at study and test, which produces varying responses across these conditions. Finally, an interaction was not found for degree and presentation state, suggesting that boundary extension ratings at each presentation state did not vary between perceived high or low depth, $F(3,81) = .852$, $p > .05$, $\eta^2 = 0.031$.

Overall, these results suggest a part of the pattern of responding in Exp. 1a was probably due to a difference in perceived depth; namely, when an image was perceived to have great depth (e.g., wider view), no boundary extension was elicited and responses tended to be veridical. When an image was perceived as being low in depth (e.g., close-up view), significant boundary extension was found.

Perceived navigability. One sample *t*-tests indicated the value for low degree images in presentation state CC was significantly different from zero, $t(39) = -2.16$, $p < .05$, compared with presentation state CC in the high navigability condition. Presentation state WW was not different from zero in either the high or low conditions, *p's* > .05 (see Table 1 for means). These results suggest that images with a low degree of navigability can elicit boundary extension by meeting the first two criteria for boundary extension, specifically CC less than zero and WW less than or no different from zero. Paired sample *t*-tests showed the low navigability condition reached a significant asymmetry between CW and WC, $t(38) = -2.3$, $p < .05$, whereas the high navigability condition did not reach significance, $p > .05$.

A 2 (degree: high, low) x 4 (presentation state: CC, WW, CW, WC) repeated measures ANOVA was used to determine whether perceived navigability influenced participants' boundary extension ratings differentially. Similar results to perceived depth were found for the main effect of presentation state and the interaction of presentation state and degree (see Fig. 5 and Table 1). Specifically, a significant main effect of presentation state was found, $F(3,102) = 32.87$, p < .0001, .492, indicating presentation state CW was rated more positively than the other presentation states, *p's* < .0001, presentation state WC was rated more negatively than the others, *p's* < .0001, and presentation states CC and WW were not significantly different from each other, *p*'s > .05. Additionally, the interaction between the degree of navigability and presentation state did not reach significance, $F(3,102) = .852$, $p > .05$, $\eta^2 = .012$. Interestingly, no significant differences were found for high and low degree images, $F(1, 34) = 2.54$, p > .05, $\eta^2 = 0.069$. Specifically, when participants viewed an image as easily navigable or not easily navigable, they responded similarly. However, a small amount of boundary extension was elicited for low navigability images (CC less than zero and a significant asymmetry). If perception of a highly cluttered scene (low navigability) elicits only a small amount of boundary extension, it might explain why the main effect of degree was not significant. Cluttered scenes of the kind used here may have only a weakened effect on boundary extension, in contrast with a previous study (Gottesman, 2011).

CHAPTER 3

EXPERIMENT 2

Overall, the purpose of Exp. 1 was to investigate whether natural, complex scenes could elicit boundary extension. The results from Exp. 1b, specifically those looking at perceived depth, suggested boundary extension could be elicited with these stimuli; however, participants had to perceive the scene as having a low degree of depth (i.e., a close-up view of a scene), otherwise, not all of the criteria would be met for boundary extension to exist (i.e., in Exp. 1a, CC was not significantly different from zero). Previous research has found wider angle images exhibit less boundary extension than close angle images (Gottesman & Intraub, 1993; Intraub et al., 1992) and wider angle images tend to exhibit greater depth, consistent with the results of Exp. 1. The results of Exp. 1 suggested mean depth could play an important role in influencing the magnitude of boundary extension. Perhaps, the statistical regularities inherent in scenes with a high or low degree of depth facilitate this effect through extraction of these gist properties.

To further explore this relationship and the relationship of boundary extension with gist processing in general, Exp. 2 used an incidental orienting task to focus participants' attention on the gist property depth during study by requiring a perceptual judgment about the depth in the scene. Images were specifically chosen to contain high or low levels of the gist property. If boundary extension relies on gist processing, or a subtype of gist processing, then presenting an image exhibiting high characteristics of that global property below the maximum exposure time should eliminate (or minimize) access to that gist property when forming a scene representation. Therefore, at a 46 ms encoding duration, access to the global property should be at its lowest availability and should contribute the least to the constructed representation. The representation created should be the most veridical in terms of actually presented sensory information. In other words, by removing the competing source of information during source monitoring, one should get a more veridical representation and less of a source monitoring error. Thereby less boundary extension should be elicited. Additionally, if access to a certain type of gist property contributes to boundary extension more, images that display a lower degree of the gist property (e.g., low depth vs. high depth) should show differences in boundary extension from their high degree counterparts. This finding would indicate the type of schematic information added to one's representation matters, as well.

Alternatively, shorter encoding times should also lead to a poorer memory representation of the original sensory input and as suggested previously, a poorer memory representation should lead to increases in boundary extension (Intraub, 2010, 2012). Therefore, if the magnitude of boundary extension continues to increase, instead of decrease, when the encoding duration is near threshold, then it can be assumed that the information added to the representation from processing of the gist does not contribute to boundary extension.

Method

Participants

Participants came from the University of Georgia research participant pool. Data from a total of 35 participants were collected for this experiment. All participants had signed informed consents, normal or corrected-to-normal vision (self-described and verified using a visual acuity test), and indicated no history of an attention deficit disorder (self-described).

Stimuli and Apparatus

Similar to Exp. 1, images were presented at a 256 x 256 pixel resolution in 24-bit color on an 18-in CRT monitor (85 Hz) using E-Prime v. 2. Participants sat 92.5 cm from the monitor, and images were centrally located, subtending a visual angle of 7° x 7°. Images used during the practice trials for Exp. 2 were taken from a subset of images from the SUN database piloted by 100 subjects to exhibit a high or low degree of depth (see Appendix B). However, images used in the experiment proper were taken from the original image set used in Greene and Oliva (2009a), which were calibrated by thousands of subjects to exhibit high or low characteristics of a specific global property. One hundred ninety eight of the original images were selected that exhibited either a high or low degree of depth, as discussed in Greene and Oliva (2009a, 2009b).

Procedure

Practice Trials. To ensure that participants could perform the task properly, practice trials were implemented. Participants practiced making speeded perceptual judgments, as well as the boundary extension task itself. The perceptual judgment task was crucial for the experiment in order to ensure participants were explicitly encoding the specific gist property of interest (e.g., depth). The task acted as an orienting procedure during encoding to focus attention on the depth of the scene and was used to determine whether participants had access to the particular property as encoding times decreased.

To practice the perceptual judgment task, participants viewed a series of scenes representing a high or low degree of depth. Participants made a two-alternative forced choice as to whether the image contained a great amount of depth, or not. If they believed the image contained a great amount of depth, measured in kilometers, they pressed 7 on the number pad or pressed 9 on the number pad if it did not contain a great amount of depth, measured in only a few meters. Participants were reminded of these ratings on a sheet of paper throughout the experiment. Feedback on accuracy and response times were provided on the screen only during these trials to help the participant understand the task. The viewing time gradually got shorter to simulate the timings of the experimental trials and train the participant how to make these speeded judgments. A total of 64 practice trials were viewed. The first eight trials showed each image for 257 ms. The image duration for the second set of eight trials were 116 ms, followed by three sets of 16 trials for 69 ms, 56 ms, and 46 ms durations, respectively. After an image was presented, a colored mosaic mask appeared for 2 s, while participants gave their response.

Once participants understood the perceptual judgment task with feedback, they continued onto practice trials combining the perceptual judgment task with a boundary extension task without feedback. These trials more closely resembled the actual experimental trials. Participants performed 10 practice trials. A fixation cross appeared at the center of the screen and a trial began when the participant pressed the spacebar. Similar to the practice perceptual judgment task, an image was presented for 257 ms, followed by a 1500 ms mosaic mask, during which time participants made their speeded perceptual judgment. Subsequently, the same image was always presented and participants made a boundary extension rating on a 5-pt scale, followed by a confidence rating. In that case, only one presentation view was shown during study and test (e.g., Daniels & Intraub, 2006; Gagnier et al., 2011).

Experiment Proper. Similar to the practice trials, participants viewed the study images for differing amounts of time, each followed by a colored mosaic mask and a test image. The test image was presented until participants made a boundary extension rating. Also similar to the practice trials, only one presentation view was shown during study and test. For the experiment proper, the study image was presented at 257 ms initially and gradually decreased to 116 ms, 69 ms, and finally 46 ms during later trials. Subsequently, a mask was presented, during which participants indicated whether the just-presented image contained a great amount of depth, or not, similar to the practice trials. Once again, the purpose of perceptual judgment orienting task during encoding was to focus attention on depth as a gist property. The mask was presented for 1500 ms to allow for the perceptual judgment to be made, and finally the test image was presented until a boundary extension rating was given.

The encoding duration of the study image was gradually decreased to reduce access to the gist (see Fig. 6). The first block of trials (24 trials each block) had an encoding duration of 257 ms, followed by a block of 116 ms, 69 ms, and 46 ms. No identifiable break between blocks was apparent to subjects. Greene and Oliva (2009a) found that participants needed about 26 ms to reach 75% accuracy, with a maximum duration of about 75 ms, therefore, the 257 ms block served as an encoding duration far above threshold. The 116 ms block served as a comparison group to Exp. 3, in which the images were above threshold for depth, but below maximum threshold for navigability. The 69 ms block served as an encoding duration above threshold but just below the maximum duration for depth, and the 46 ms block served as an encoding duration just barely above threshold for depth.

Results and Discussion

Perceptual judgment task

The purpose of Exp. 2 was to determine whether reliance on the gist during scene perception, specifically depth, influences boundary extension. To reduce access to the gist, encoding time of the study image was decreased to reach threshold levels of gist perception. It was predicted that there should be a noticeable decrease in performance on the perceptual judgment task, when access to the gist property had become compromised. During the experiment proper, participants gave a two-alternative forced choice response at encoding as to whether the image contained a great amount of depth or not. On average, participants responded to 95.1% (*sd* = .072) of trials. Outliers with a total response percentage less than 2.5 standard deviations away from this mean were removed from analyses. Only one subject was removed due to poor responding, creating a new overall average response total of 96.1%. Individual trials were removed from analyses if participants failed to give a response on the perceptual judgment task, which accounted for only 3.9% of all trials. Overall, participants performed above chance at each encoding duration on the perceptual judgment task (see Fig. 7 and Table 2), in which chance was 50%.

A repeated measures ANOVA (encoding duration: 257, 116, 69, 46 ms) was used to determine whether there was a noticeable decline in accuracy, at any given encoding duration. A significant difference between accuracy was found, $F(3,99) = 14.8$, $p < .0001$, $\eta^2 = 0.31$, indicating specifically that accuracy at 46 ms was significantly less than the other encoding durations, p*'s* < .001. These results suggest that access to depth as a gist property was reduced at 46 ms, consistent with the threshold values reported in Greene and Oliva (2009a) which found a maximum exposure of about 75 ms (*sd* = 4.9). Although, the 69 ms condition is technically below the maximum exposure duration as well, it is close enough to be within their margin of error and did not reach significance in this study. Furthermore, these results would predict that any change in the effect of gist on boundary extension will be most noticeable at the 46 ms duration.

Boundary extension rating

Trials missed by participants, as indicated by responding with “don’t remember/missed picture” on the confidence task, or giving no response, were excluded from the following analyses. One sample t-tests were used to determine whether boundary extension was elicited at each encoding duration. The longer encoding durations, 257 ms, 116 ms, and 69 ms, produced boundary extension, indicating the mean rating for each encoding duration was significantly different from zero, $t(33) = -4.20$, $p < .0001$, $t(33) = -3.55$, $p < .001$, and $t(33) = -3.97$, $p < .0001$, respectively (see Table 2 and Fig. 8, black bar). In other words, at longer encoding durations, boundary extension was elicited. At 46 ms, however, the mean boundary extension rating was not significantly different from zero, $M = -0.10$ (.05), $t(33) = -1.99$, $p > .05$, indicating that boundary extension was not elicited. Along with the results from the perceptual judgment task, which indicated that accuracy for the 46 ms condition was the worst, these result suggest that compromising access to the gist, specifically depth, can eliminate boundary extension. These results support the hypothesis that the gist facilitates boundary extension by adding competing sources of top-down information into our scene representation. Access to gist properties leads to greater source monitoring errors at recognition (i.e., greater boundary extension) and thereby, the reduction of gist access, as shown in the 46 ms condition, leads to an elimination of boundary extension.

To further investigate the relationship between depth and boundary extension, mean boundary extension ratings were calculated for high degree of depth images compared with low degree of depth images (see Table 2 for means and Fig. 8, light grey and dark grey bars). It was predicted that low degree of depth images, with their similarity to close-angle images, would elicit more boundary extension than high degree of depth images with their similarity to wide-angle images. Consistent with this prediction, only low degree images were significantly different from zero (i.e., elicited boundary extension) at each encoding duration, $t(33) = -5.37$, $p < .0001$, $t(33) = -4.50$, $p < .0001$, $t(33) = -5.98$, $p < .0001$, and $t(33) = -4.64$, $p < .0001$, respectively, whereas high degree images were not significantly different from zero at any of the encoding durations, $t(33) = -0.50$, $p > .40$, $t(33) = -0.85$, $p > .40$, $t(33) = 0.13$, $p > .40$, and $t(33) = 1.82$, $p > .05$, respectively. Note that at 46 ms, high degree images are approaching significance in the positive direction. Values significantly greater than zero indicate than viewers thought the image at test looked wider-angle than the image at study, despite being identical. It is possible that viewers were treating the high depth images similar to a WW condition (wide at study and wide at test). In this case, at 46 ms, when memory should be most verdical, they normalized to the mean size difference of the perceived depth of the presented images (Intraub et al., 1992; Intraub et al., 1998).

A 2 (degree: high; low) x 4 (encoding duration: 257; 116; 69; 46 ms) repeated measures ANOVA was used to evaluate whether boundary extension varied by time and degree of access to mean depth. As predicted, a main effect of degree was found, $F(1,33) = 43.35$, $p < .0001$, $\eta^2 = 0.57$, in which images with a high degree of depth elicited less boundary extension than images with a low degree of depth (see Table 2 and Fig. 8). A significant difference in boundary extension across encoding durations was not found, $F(3,99) = 1.70$, $p > .05$, $\eta^2 = 0.05$, nor was the interaction of encoding duration and degree significant, $F(3,99) = 1.29$, $p > .05$, $\eta^2 = 0.04$. Whereas the data did not vary significantly by encoding duration overall, there were significant differences between the mean boundary extension rating at 69 ms, $M = -0.20$, and 46 ms, $M = -0.09$, $p < .02$, providing further support that reducing access to the gist property depth has an influence on boundary extension.

Compared with Exp. 1, the CC boundary extension ratings in Exp. 2 were much larger overall (Exp. 1, 257 ms: $M = -0.08$; Exp. 2, 257 ms: $M = -0.19$). These data are consistent with the studies showing divided attention tasks lead to greater boundary extension (Intraub et al., 2008). The perceptual judgment task minimized focused attention on encoding of the scene and led to an increase in BE.

CHAPTER 4

EXPERIMENT 3

Overall, the purpose of Exp. 1 was to investigate whether natural, complex scenes could elicit boundary extension. The results from Exp. 1b, specifically those looking at perceived navigability, suggested a weakened form of boundary extension could be elicited with these stimuli; however, participants had to perceive the scene as having a low degree of navigability (i.e., a cluttered, hard-to-navigate scene), otherwise, not all of the criteria would be met for boundary extension to exist (i.e., in Exp. 1a, CC was not significantly different from zero). Previous research had found cluttered scenes elicited greater boundary extension than uncluttered scenes (Gottesman, 2011), however, “clutter” was not clearly defined. Exp. 3 used the gist property navigability to explore this relationship further. Similar to Exp. 2, participants’ attention was isolated and focused on the gist property navigability during study by using the perceptual judgment task. In this case, reliance on any other gist property could not be substituted as encoding time decreased and hypothetically access to the gist was reduced. If boundary extension relies on processing of navigability, then presenting an image exhibiting high characteristics of that global property below the maximum exposure time should eliminate (or minimize) access to that property when forming a scene representation. Therefore, at a 46 ms encoding duration, access to the global property should be at its lowest availability and should contribute the least to the constructed representation. The representation created should be the most veridical in terms of actually presented sensory information. In other words, by removing the competing source of information during source monitoring, one should get a more veridical representation and less of a source monitoring error. Thereby less boundary extension should be elicited. Alternatively, shorter encoding times should also lead to a poorer memory representation of the original sensory input and as suggested previously, a poorer memory representation should lead to increases in boundary extension (Intraub, 2010, 2012). Therefore, if the magnitude of boundary extension continues to increase, instead of decrease, when the encoding duration is near threshold, then it can be assumed that the information added to the representation from processing of the gist does not influence boundary extension.

Method

Participants and Stimuli

All methodological details were the same as Exp. 2, except a new group of 35 participants were used and images in the experiment proper exhibited either a high or low degree of the global scene property navigability, as discussed in Greene and Oliva (2009a, 2009b).

Procedure

Practice trials and experiment proper. All practice and experimental procedures were identical to Exp. 2, except the type of perceptual judgment made by the participant. Participants pressed 7 on the number pad if the scene was high in navigability, meaning the scene would be easy to navigate through, and contained a very obvious path free of obstacles. Participants pressed 9 on the number pad if the scene was low in navigability, meaning the scene contained many obstacles or a difficult terrain and would be hard to navigate through. Recall that Greene and Oliva (2009a) found participants needed about 36 ms to reach 75% accuracy, with a maximum duration of about 120 ms, therefore, the 46 ms block served as an encoding duration just above threshold, similar to Exp. 2. The 257 ms block served as a comparison set to Exp. 1

and 2, far above threshold. The 116 ms block served as an encoding duration just below maximum exposure, and 69 ms block served as an intermediary duration.

Results and Discussion

Perceptual Judgment

The purpose of Exp. 3 was to determine whether reliance on the gist during scene perception, specifically navigability, influences boundary extension. To reduce access to the gist, encoding time of the study image was decreased to reach threshold levels of gist perception. Similar to Exp. 2, it was predicted that there should be a noticeable decrease in performance on the perceptual judgment task, when access to the gist property had been eliminated. On average, participants responded to 95.0% (*sd* = .067) of trials. Outliers with a total response rate less than 2.5 standard deviations away from this mean were removed from analyses. Only one subject was removed due to poor responding, creating a new overall average response rate of 95.8%. Individual trials were removed from analyses if participants failed to give a response on the perceptual judgment task, which accounted for only 4.2% of all trials. Overall, participants performed above chance at each encoding duration on the perceptual judgment task (see Table 3 and Fig. 9), in which chance was 50%.

A repeated measures ANOVA (encoding duration: 257, 116, 69, 46 ms) was used to determine whether there was a noticeable decline in accuracy, at any given encoding duration (see Table 3 and Fig. 9). A significant difference in accuracy across encoding duration was found, $F(3,99) = 40.1$, $p < .0001$, $\eta^2 = 0.549$, indicating specifically that accuracy at 69 ms was significantly less than the two longer encoding durations, *p's* < .005, and accuracy at 46 ms was significantly less than all of the other encoding durations, *p's* < .0001. These results suggest that access to navigability as a gist property was reduced at 69 ms, and continued to decrease at 46 ms. These results are consistent with the threshold values reported in Greene and Oliva (2009a) indicating a maximum exposure of 120 ms (*sd* = 9.2). The 69 ms condition and the 46 ms conditions were both below this time. Similar to Exp. 2, although the 116 ms condition is technically below the maximum exposure duration, it was still within the margin of error. Furthermore, these results would predict that any change in the effect of gist on boundary extension will be most noticeable at the 69 ms duration, and should be lowest at 46 ms.

Boundary Extension Ratings

Similar analyses as Exp. 2 were used to evaluate whether boundary extension is differentially influenced by limiting access to navigability as encoding duration decreases. One sample t-tests were used to determine whether boundary extension was elicited at each encoding duration (see Table 3 for means and Fig. 10, black bars). Interestingly, in contrast to Exp. 2, the two longest encoding durations, 257 ms and 116 ms, did not produce significant boundary extension, indicating the mean rating for each encoding duration was not significantly different from zero, $M = -0.01$ (.07), $t(33) = -0.14$, $p > .8$, and $M = -0.11$ (.08), $t(33) = -1.33$, $p > .15$, respectively. When participants viewed these particular scenes, they were capable of correctly indicating the image they saw at test was the same image they saw at study. However, as presentation time decreased, 69 ms and 46 ms, the amount of boundary extension increased, $t(33) = -2.80$, $p < .01$, and $t(33) = -2.80$, $p < .01$, respectively. The perceptual judgment task suggested that access to the gist property navigability was reduced at 69 ms, and continued to become compromised at 46 ms. Instead of boundary extension decreasing when access to the gist was reduced, boundary extension increased. These results appear to suggest that the gist property navigability has a contradictory effect on boundary extension; access to gist property navigability helps to provide a more veridical memory representation of the scene, and when eliminated at shorter encoding durations, participants became less veridical in their judgment, producing boundary extension.

To further investigate the relationship between navigability and boundary extension, mean boundary extension ratings were calculated for high degree of navigability images compared with low degree of navigability images (see Table 3 for means and Fig. 10, light grey and dark grey bars). It was predicted that low degree images, with their similarity to cluttered images, would elicit more boundary extension than high degree images with their similarity to uncluttered images. Whereas high degree images were not significantly different from zero at any of the encoding durations, $t(33) = 1.47$, $p > .10$, $t(33) = -0.70$, $p > .40$, $t(33) = -1.67$, $p > .10$, and $t(33) = -1.82$, $p > .05$, respectively, the low degree images at 257 ms and 116 ms were also not significantly different from zero, $t(33) = -1.75$, $p > .05$, and $t(33) = -1.61$, $p > .10$, respectively. Only the 69 ms and 46 ms encoding durations elicited more, or any, boundary extension, $t(33) = -3.34$, $p < .005$, and $t(33) = -2.64$, $p < .02$, respectively. Taken in a different way, high navigability scenes never produced boundary extension. On the other hand, scenes with low navigability (i.e., cluttered scenes) were extended only at very short durations (69 ms and 46 ms).

A 2 (degree: high, low) x 4 (encoding duration: 257, 116, 69, 46 ms) repeated measures ANOVA further supported these findings. There was a significant main effect of degree, $F(1,33) = 10.27$, $p < .005$, $\eta^2 = 0.24$, in which low degree images elicited greater boundary extension overall. The non-significant main effect of encoding duration, $F(1,33) = 2.42$, $p > .05$, $\eta^2 = 0.07$, and the non-significant interaction, $F(1,33) = 0.74$, $p > .50$, $\eta^2 = 0.02$, suggest that overall reducing access to the gist property navigability by reducing encoding duration had only a minor effect on boundary extension.

A 2 (gist property: depth, navigability) x (2 (degree: high, low) x 4 (encoding duration: 257, 116, 69, 46 ms) repeated measures ANOVA was used to test whether one type of gist property influenced boundary extension differently overall. The only significant comparisons were a main effect of degree and the interaction of degree and type of gist property, $F(1,66) = 49.13$, $p < .0001$, $\eta^2 = 0.43$, and $F(1,66) = 7.05$, $p < .01$, $\eta^2 = 0.10$, respectively. Specifically, these results indicate that overall scenes displaying a high degree of either gist property elicited less boundary extension than scenes displaying a low degree of either gist property. Additionally, and not surprisingly given the individual results from Exp. 2 and 3, the interaction indicates that the most boundary extension was found when scenes displayed a low degree of depth, $M = -0.35$ (.06), compared with a high degree of depth, $M = 0.01$ (.05), in addition to images with a low and high degree of navigability, $M = -0.20$ (.06), and $M = -0.04$ (.05), respectively. Essentially close-up images produced the greatest boundary extension, consistent with previous literature. Together, these data suggest that the extraction of the gist property depth plays a more influential role in the mechanism of boundary extension than navigability.

CHAPTER 5

GENERAL DISCUSSION

The purpose of the current experiments was to examine how gist properties are utilized by participants during boundary extension, particularly, whether extraction of specific gist properties could explain previous findings in the boundary extension literature. In order to investigate this relationship between the gist and boundary extension, and be consistent with previous literature on gist processing, it had to first be determined whether boundary extension could be elicited with complex natural landscapes. The results from Exp. 1 showed that boundary extension could be elicited from complex natural, landscape scenes; however, it heavily depended on the participants' perceived depth rating and their perceived navigability of the scene. When all images were used, only two of the three suggested criteria for the existence of boundary extension were met, specifically, WW was less than zero and the absolute value of CW was not equal to the absolute value of WC, however, CC was not different from zero. Interestingly these results ran counter to previous evidence in the boundary extension literature showing normalization to the average set size. Once the images were parsed out by viewers' ratings, only then could the results be explained and a clear boundary extension trend noticed.

Specifically, when images rated in the upper and lower quartile of perceived depth were analyzed separately, a strong boundary extension effect could be seen for images containing a low amount of depth (see Fig. 4), whereas images with a great amount of depth showed more veridical memory. In this case, for target conditions, CC and WW, participants were able to correctly indicate that the image they were viewing at test was identical to the image at study for high depth images. This suggests that images high in depth may have limited access to providing additional schematic information. Similarly, when images rated in the upper and lower quartile of perceived navigability were analyzed separately, an attenuated boundary extension effect was shown for low navigability images, whereas high navigability images did not show boundary extension (see Fig. 5). These results allowed us to continue on with Experiments 2 and 3 to investigate the relationship further.

Overall, the current studies were able to show that varying gist properties contribute differentially to the boundary extension phenomenon. Specifically, Exp. 2 manipulated access to the mean depth gist property and showed that boundary extension decreased as access to it was reduced, whereas Exp. 3 manipulated access to the navigability gist property and found that boundary extension increased as access to it was reduced. Consistent with the multisource model of scene perception, we make the claim that certain types of gist properties influence boundary extension to a greater extent than others and, even at very short timespans, can determine whether additional information is incorporated into our scene representation or not.

Recall that according to the multisource model of scene perception (see Fig. 1), multiple factors, or sources during viewing are rapidly integrated into a single egocentric representation of the scene, so that within a single fixation the scene representation created includes input from the actual scene, as well as associated top-down and schematic information about the scene (Intraub, 2010, 2012). Given that boundary extension is suggested to be a reality monitoring error between the actual image presented and the representation created of the scene (Intraub, 2012), activation of schema related information via gist processing should increase the discrepancy between the actual image presented and the constructed representation. Therefore, as access to the gist was reduced (i.e., as encoding time decreased) one of two possibilities were predicted: (1) Because the gist enables quick access to the schematic information, reducing access to the gist would lead to more veridical memory (hence a decrease in boundary extension); or (2) as encoding duration decreases, previous literature would suggest that degrading the memory representation of the scene should increase error production and lead to increases in boundary extension. Fuzzy Trace Theory suggests that as participants increase their reliance on gist properties, they should also increase their false memory production (Brainerd & Reyna, 2002), leading to greater boundary extension. The current studies found both possibilities occurred depending on the type of gist property presented and focused on; specifically, the gist property depth appeared to support the first possibility, while the gist property navigability appeared to support the second possibility.

In Exp. 2, accuracy on the depth perceptual judgment task dropped below 75% between 69 ms and 46 ms, in correspondence to Greene and Oliva's (2009a) findings; therefore, access to the gist was assumed to be compromised at 46 ms. When access to the gist was reduced, there was a corresponding reduction in boundary extension (i.e., values were no longer significantly different from zero). By virtue of boundary extension being an error of memory, a reduction in boundary extension indicates memory became more veridical (see Fig. 8). Interestingly, when the images are split by high and low characteristics one can see that the effect is driven by high depth images progressively being viewed as wider angle. As encoding duration decreases, viewers treated the high depth images as if they were wider angle images (i.e., WW), indicating that the test image appeared wider than at study, despite no actual change in presentation view. In these cases, only the mean depth of the image was varied and when the gist was no longer available (at 46 ms) and identical presentation views were shown, viewers normalized to the average mean depth of the image, as opposed to the average presentation view of the set found in Intraub et al. (1996). Furthermore, viewers consistently reported that low depth images appeared closer at test than at study, similar to accounts of both extension and normalization (Intraub et al. 1996).

In Exp. 3, accuracy on the navigability perceptual judgment task indicated that a significant decrease in performance occurred between 116 ms and 69 ms, and continued to drop below 75% accuracy at 46 ms, in correspondence to Greene and Oliva (2009a). Given that, access to the gist was assumed to be compromised at 69 ms, and further reduced at 46 ms, as encoding duration decreased, a noticeable increase in boundary extension occurred, as opposed to the results in Exp. 2 (see Fig. 10). In other words, boundary extension ratings were significantly different from zero, indicating that an identical image at test appeared closer than at study. The most interesting aspect of these findings is that access to the gist property navigability, as indicted by above threshold encoding durations, appears to inhibit boundary extension. Only when access is reduced (at 69 and 46 ms) does boundary extension become apparent for low navigability images. Greene and Oliva (2009b) categorized navigability as measuring functional properties about a scene, for example, how easy it would be to form an escape path from a predator. In this case, it would make sense that the viewer would want access to an accurate spatial layout of the scene to prepare for action and not make errors of perception.

From the standpoint of the action perception hypothesis, when goal-directed action must be taken in a scene, the fast, dorsal, action pathway visually processes scenes in an egocentric way so that object positions and orientations are coded exactly (Goodale & Milner, 1992). On the other hand, the ventral, perception pathway processes scenes in an allocentric way and may lead to more errors in perception (Goodale & Milner, 1992). When navigability is low, and a direct pathway cannot be identified, it’s possible that viewers may extrapolate, or extend beyond the borders of the available scene in order to identify an escape path, modifying their scene representation. This additional information added to the scene representation would increase the discrepancy in the source monitoring error leading to increases in boundary extension. Future studies could examine how manipulations of goal-directed actions influence boundary extension.

Conclusion

In conclusion, the current studies were able to show that the gist of a scene can provide varying sources of information during scene perception that differentially influences boundary extension. Access to the mean depth of a scene appears to facilitate the existence of boundary extension, by providing structural information about the scene's size scaling which enhances the source monitoring error. On the other hand, access to the navigability of a scene appears to inhibit boundary extension by providing goal-directed functional information about the scene leading to a more veridical memory representation of the scene.

REFERENCES

Bertamini, M., Jones, L. A., Spooner, A., & Hecht, H. (2005). Boundary extension: The role of magnification, object size, context, and binocular information. *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1288.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, *11*(5), 164-169.

Dickinson, C. A., & Intraub, H. (2008). Transsaccadic representation of layout: What is the time course of boundary extension? *Journal of Experimental Psychology: Human Perception and Performance, 34*(3), 543.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*(3), 316-355.

Gagnier, K. M., Dickinson, C. A., & Intraub, H. (2013). Fixating picture boundaries does not eliminate boundary extension: Implications for scene representation. *The Quarterly Journal of Experimental Psychology*(just-accepted), 1-59.

Gagnier, K. M., Intraub, H., Oliva, A., & Wolfe, J. M. (2011). Why does vantage point affect boundary extension? *Visual Cognition, 19*(2), 234-257.

Gottesman, C. V. (2011). More Space Please! The effect of clutter on Boundary Extension. *Journal of Vision, 11*(11), 1130-1130.

Gottesman, C. V., & Intraub, H. (2002). Surface construal and the mental representation of scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28*(3), 589.

Gottesman, C. V., & Intraub, H. (2003). Constraints on spatial extrapolation in the mental representation of scenes: View-boundaries vs. object-boundaries. *Visual Cognition, 10*(7), 875-893.

Greene, M. R., & Oliva, A. (2009a). The briefest of glances: The time course of natural scene understanding. *Psychological Science, 20*(4), 464-472.

Greene, M. R., & Oliva, A. (2009b). Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology, 58*(2), 137.

Intraub, H. (1992). Contextual factors in scene perception. *Advances in Psychology, 88*, 45-72.

Intraub, H. (2010). Rethinking scene perception: A multisource model. *Psychology of Learning and Motivation, 52*, 231-264.

Intraub, H. (2012). Rethinking visual scene perception. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*(1), 117-127.

Intraub, H., Bender, R. S., & Mangels, J. A. (1992). Looking at pictures but remembering scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(1), 180.

Intraub, H., & Bodamer, J. L. (1993). Boundary extension: Fundamental aspect of pictorial representation or encoding artifact? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(6), 1387.

Intraub, H., Daniels, K. K., Horowitz, T. S., & Wolfe, J. M. (2008). Looking at scenes while searching for numbers: Dividing attention multiplies space. *Attention, Perception, & Psychophysics, 70*(7), 1337-1349.

Intraub, H., & Dickinson, C. A. (2008). False Memory 1/20th of a second later: What the early onset of Boundary Extension reveals about perception. *Psychological Science, 19*(10), 1007-1014.

Intraub, H., Gottesman, C. V., & Bills, A. J. (1998). Effects of perceiving and imagining scenes on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(1), 186.

Intraub, H., Gottesman, C. V., Willey, E. V., & Zuk, I. J. (1996). Boundary extension for briefly glimpsed photographs: Do common perceptual processes result in unexpected memory distortions? *Journal of Memory and Language*, *35*(2), 118-134.

Intraub, H., Hoffman, J. E., Wetherhold, C. J., & Stoehs, S.-A. (2006). More than meets the eye: The effect of planned fixations on scene representation. *Attention, Perception, & Psychophysics*, *68*(5), 759-769.

Intraub, H., & Richardson, M. (1989). Wide-angle memories of close-up scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(2), 179.

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality and Social Psychology*, *56*(3), 326.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*(1), 3.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review, 88*(1), 67-85.

Legault, E., & Standing, L. (1992). Memory for size of drawings and of photographs. *Perceptual and Motor Skills*, *75*(1), 121-121.

McDunn, B. A., Siddiqui, A. P., & Brown, J. M. (2013). Seeking the boundary of boundary extension. *Psychonomic Bulletin and Review, 20.* doi: 10.3758/s13423-013-0494-0

Oliva, A. (2005). Gist of the scene. *Neurobiology of Attention, 696*, 64.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, *42*(3), 145-175.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2*(5), 509-522.

Potter, M. C. (2012). Recognition and memory for briefly presented scenes. *Frontiers in psychology*, *3*.

Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, *81*(1), 10.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time-and spatial-scale-dependent scene recognition. *Psychological Science*, *5*(4), 195-200.

Siddiqui, A. P., McDunn, B., & Brown, J. M. (2012). Seeking the boundary for boundary extension. *Journal of Vision*, *12*(9), 1074-1074.

Siddiqui, A. P., McDunn, B., & Brown, J. M. (2013). Processing images in a masked BE paradogm: Does emotion matter? *Journal of Vision*, *13*(9), 1316.

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). *SUN database: Large-scale scene recognition from abbey to zoo.* IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010.

Footnotes

1. The author must acknowledge Beighley and Intraub (2012), a poster presented at the Vision Sciences Society. They presented inverted scenes for 125 ms, 250 ms, or 375 ms, and found boundary extension increased as viewing time decreased. However, exact mean values were not cited.

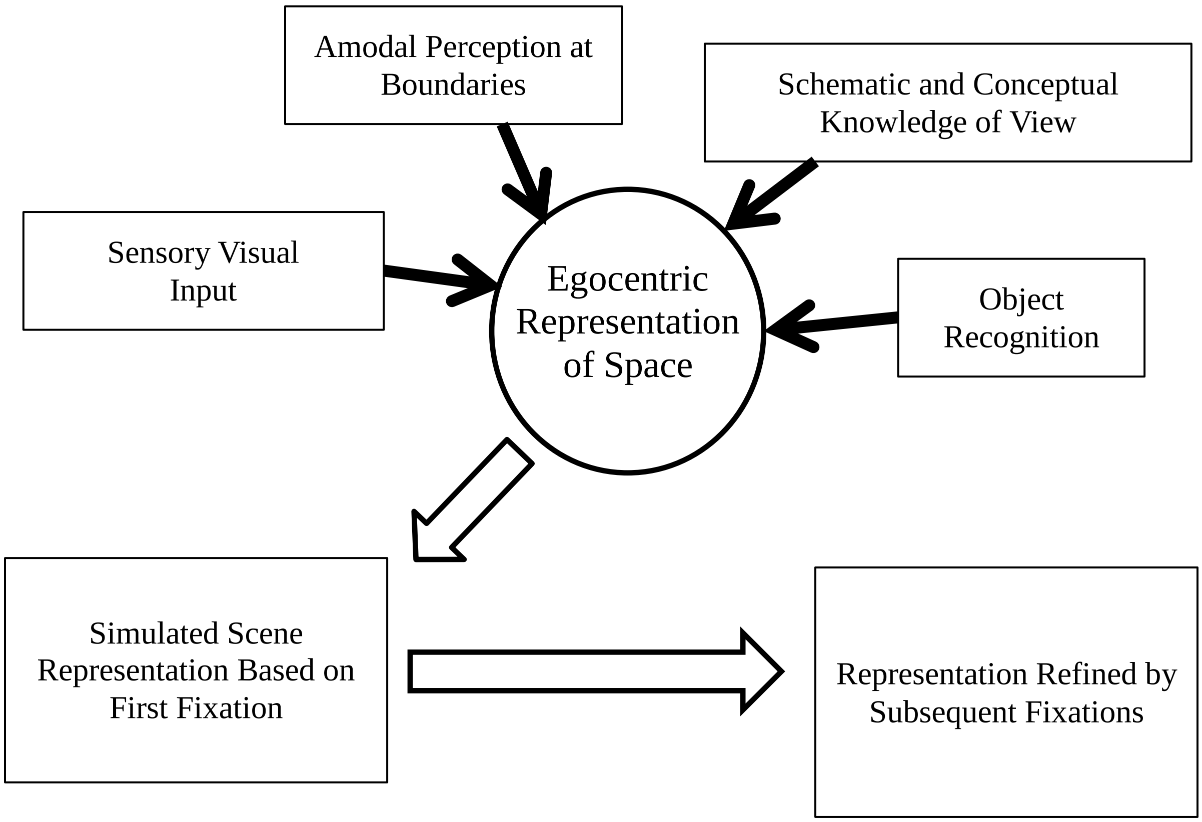

Figure 1. Example of the Multisource Model of Scene Perception.

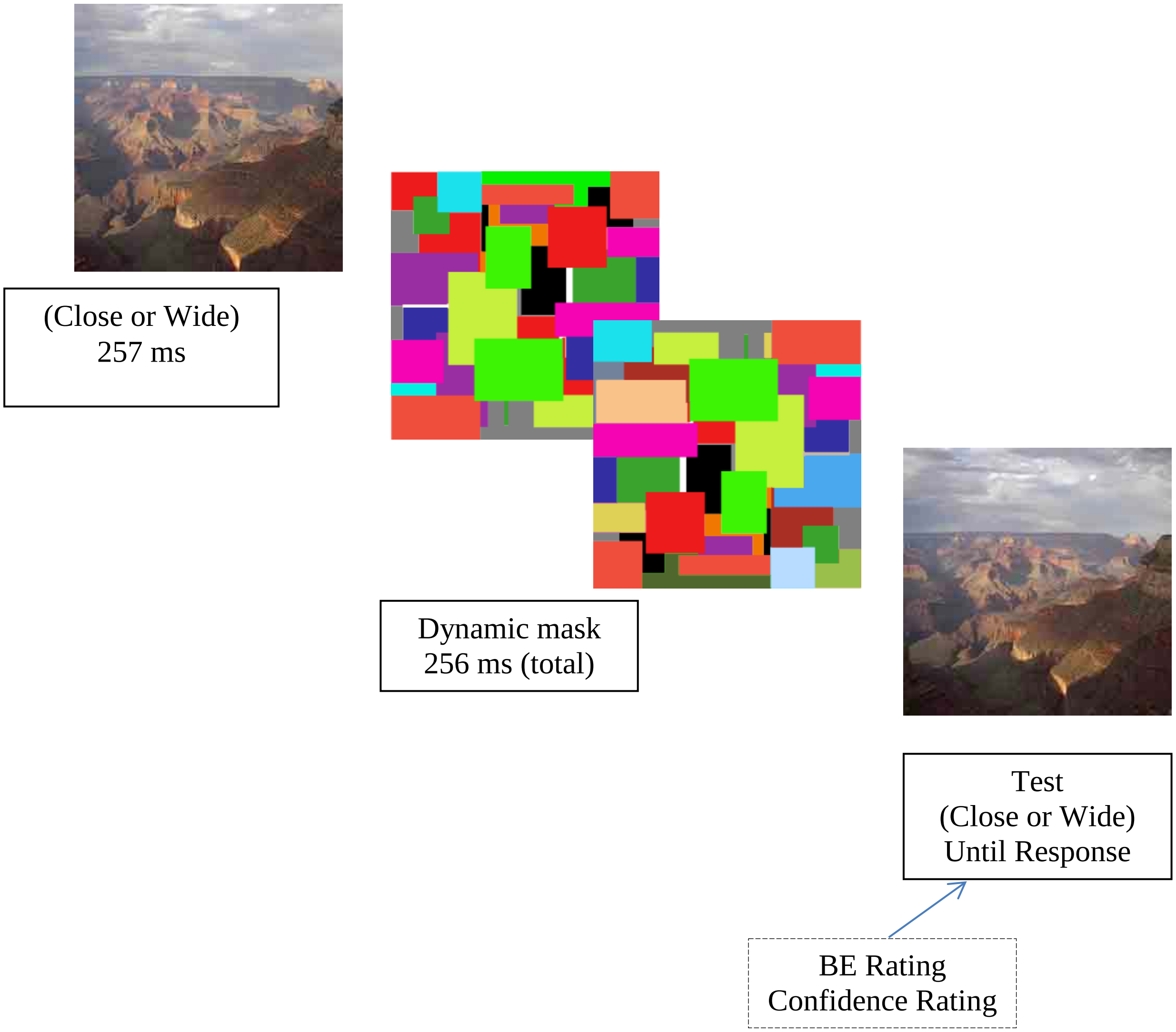

Figure 2. Example of the general procedure used in Exp. 1. This figure represents a presentation state of CW, close- angle image at study, wide angle image at test. The image at study is presented as a close-angle view but in the experimental trials could be close or wide-angle. The image at test is presented as a wide-angle view, but could be close or wide during the experimental trials.

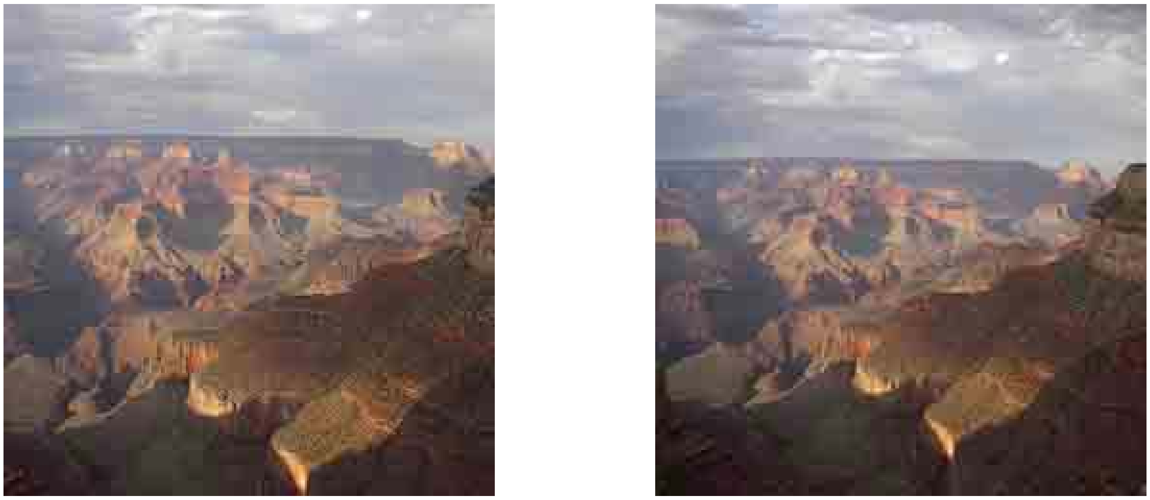

Figure 3. Example of a close *(on the left)* and wide-angle *(on the right)* image used in Exp. 1. The images were created by taking the original image and designating that as the wide angle image, and then reducing/cropping 16% to create the close angle image.

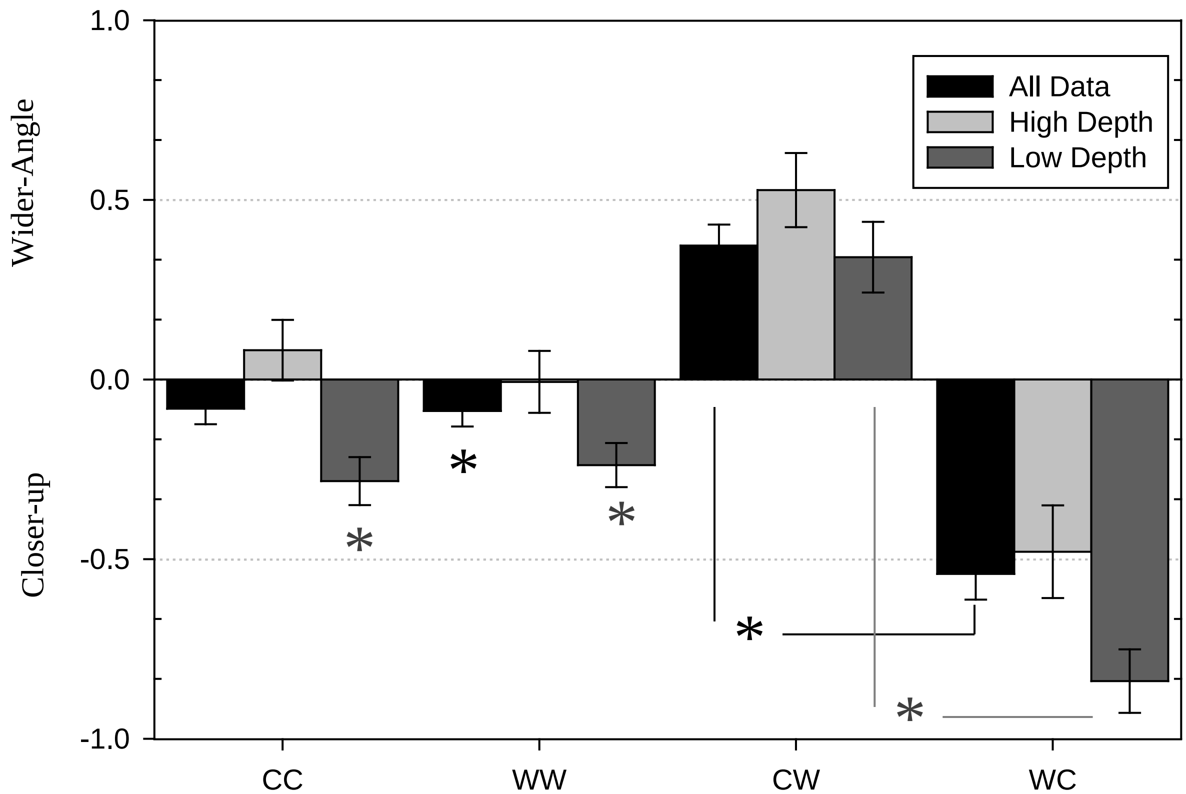

Figure 4. Boundary extension results based on perceived depth, in Exp. 1. The black bars represent data from Exp. 1a. The light and dark grey bars represent data from Exp. 1b.

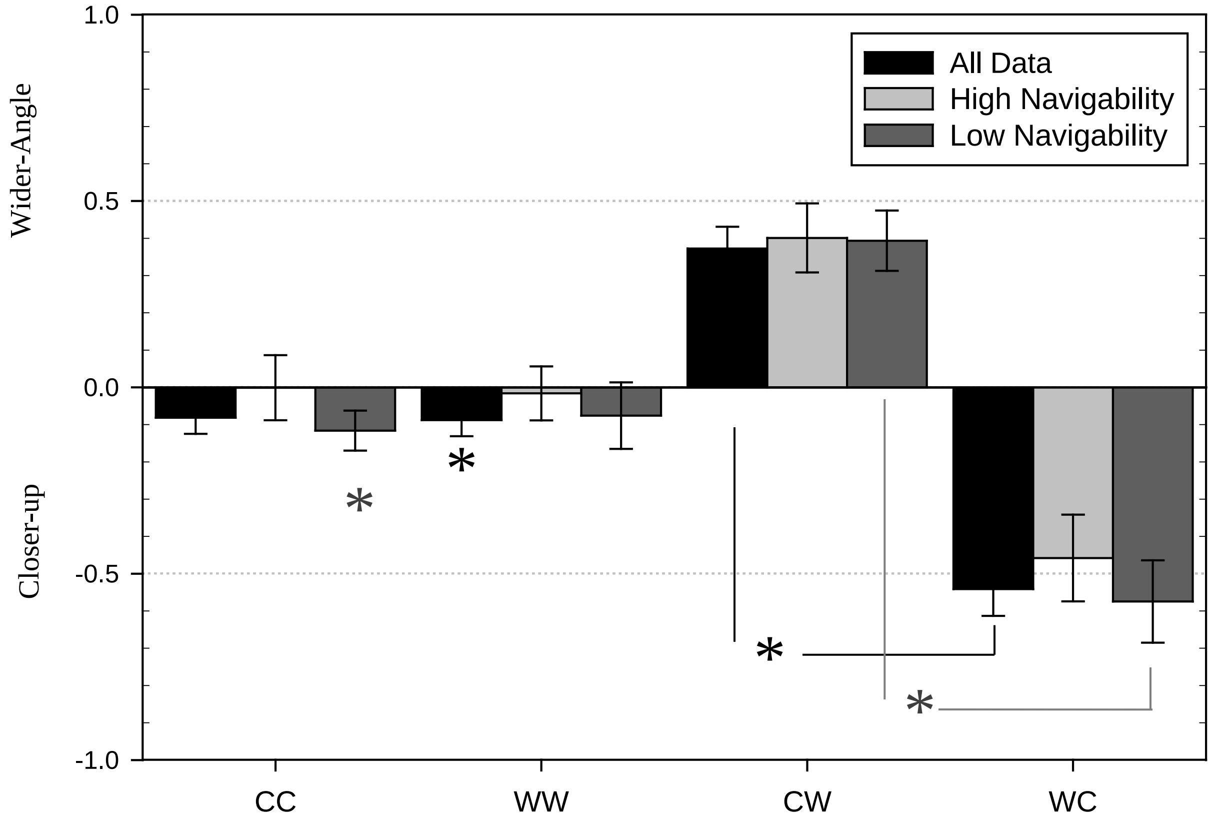

Figure 5. Boundary extension results based on perceived navigability, in Exp. 1. The black bars represent data from Exp. 1a. The light and dark grey bars represent data from Exp. 1b.

* $p < .05$

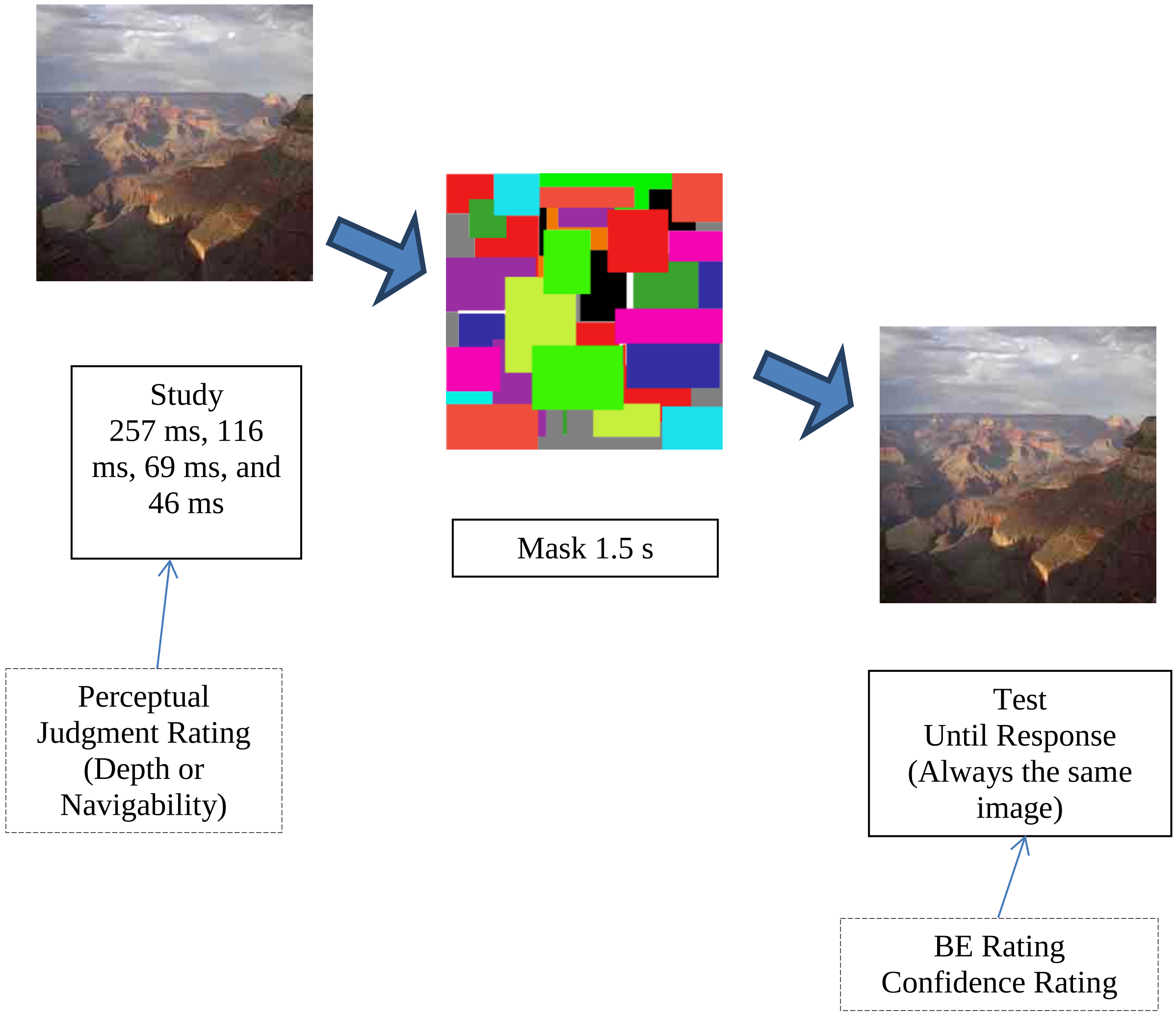

Figure 6. Example of the general procedure used in Exp. 2 & 3.

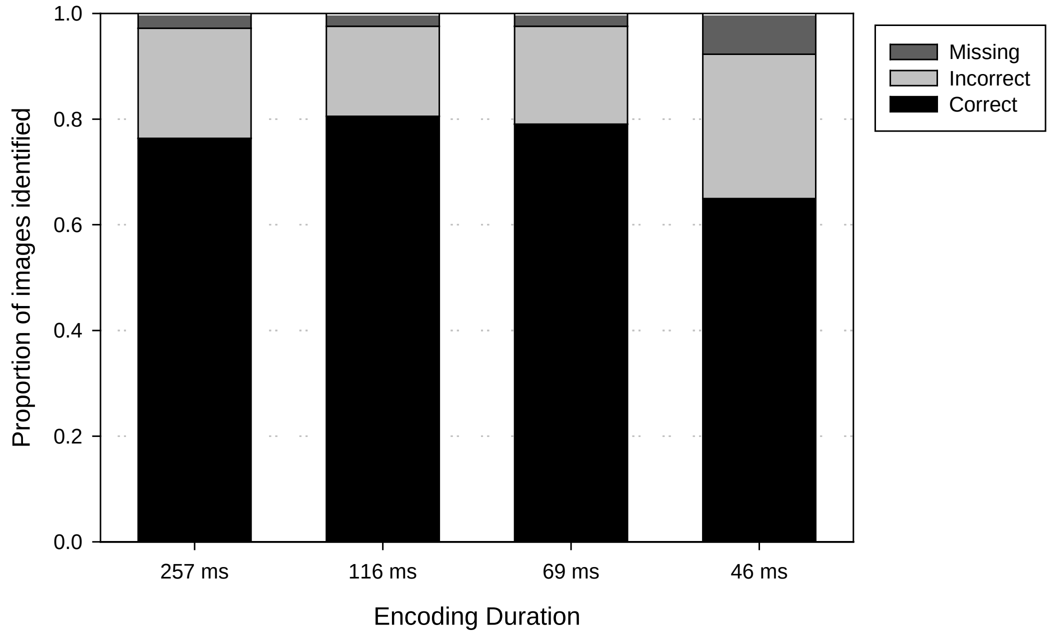

Figure 7. Proportion of trials identified as correct, incorrect, or missing during the perceptual judgment task by encoding duration, Exp. 2.

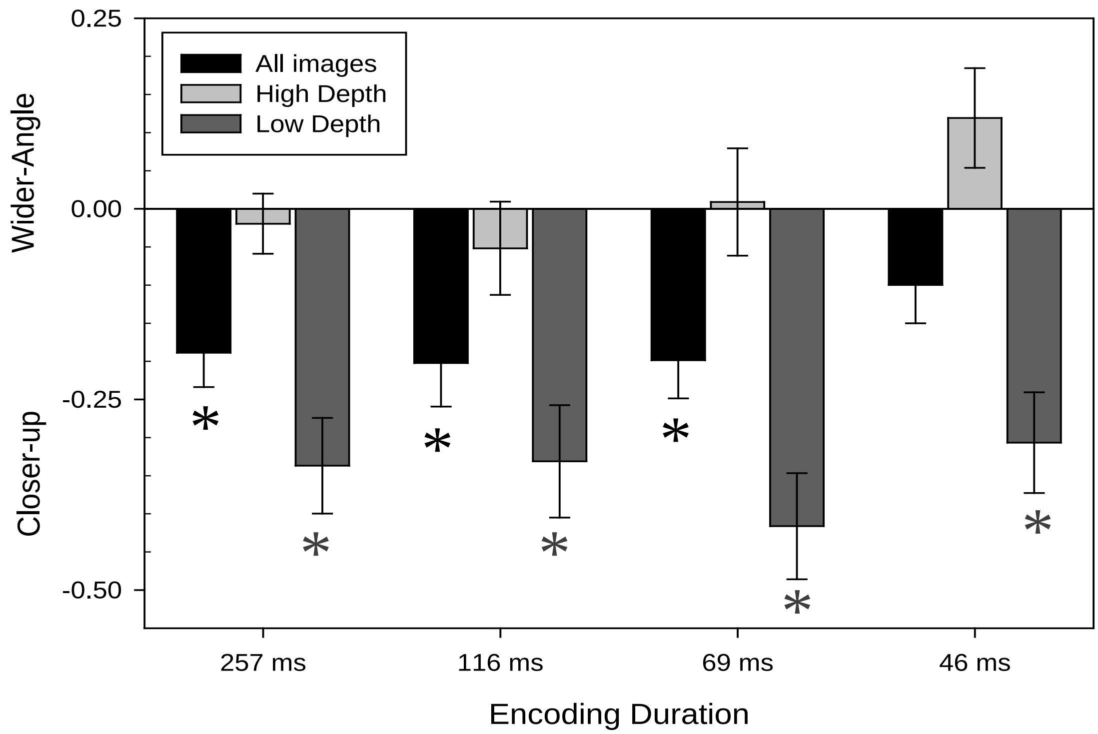

Figure 8. Boundary extension ratings by image type for Exp. 2 as a function of encoding duration. Black bars represent all data, excluding blank trials and images given a “don’t remember” response. The light grey bars in the middle of each set represent average boundary extension ratings given to scenes which displayed a high degree of depth, and the dark grey bars on the right of each set represent average boundary extension ratings given to scenes which displayed a low degree of depth.

* $p < .05$

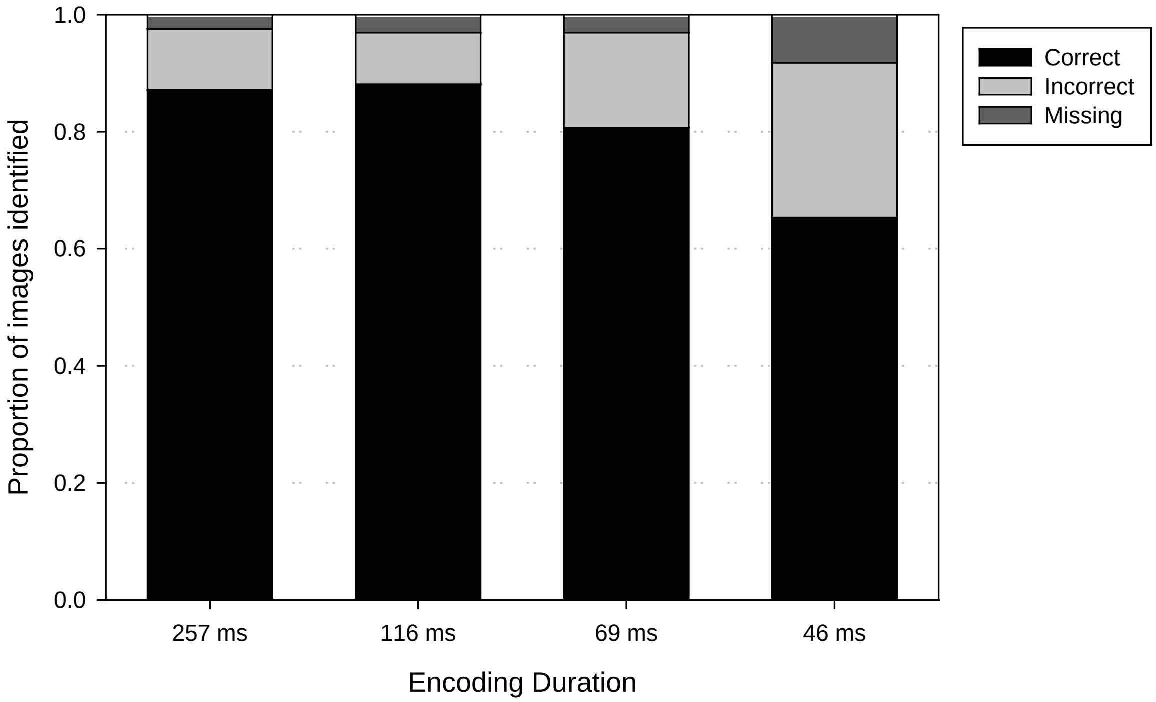

Figure 9. Proportion of trials identified as correct, incorrect, or missing during the perceptual judgment task by encoding duration in Exp. 3.

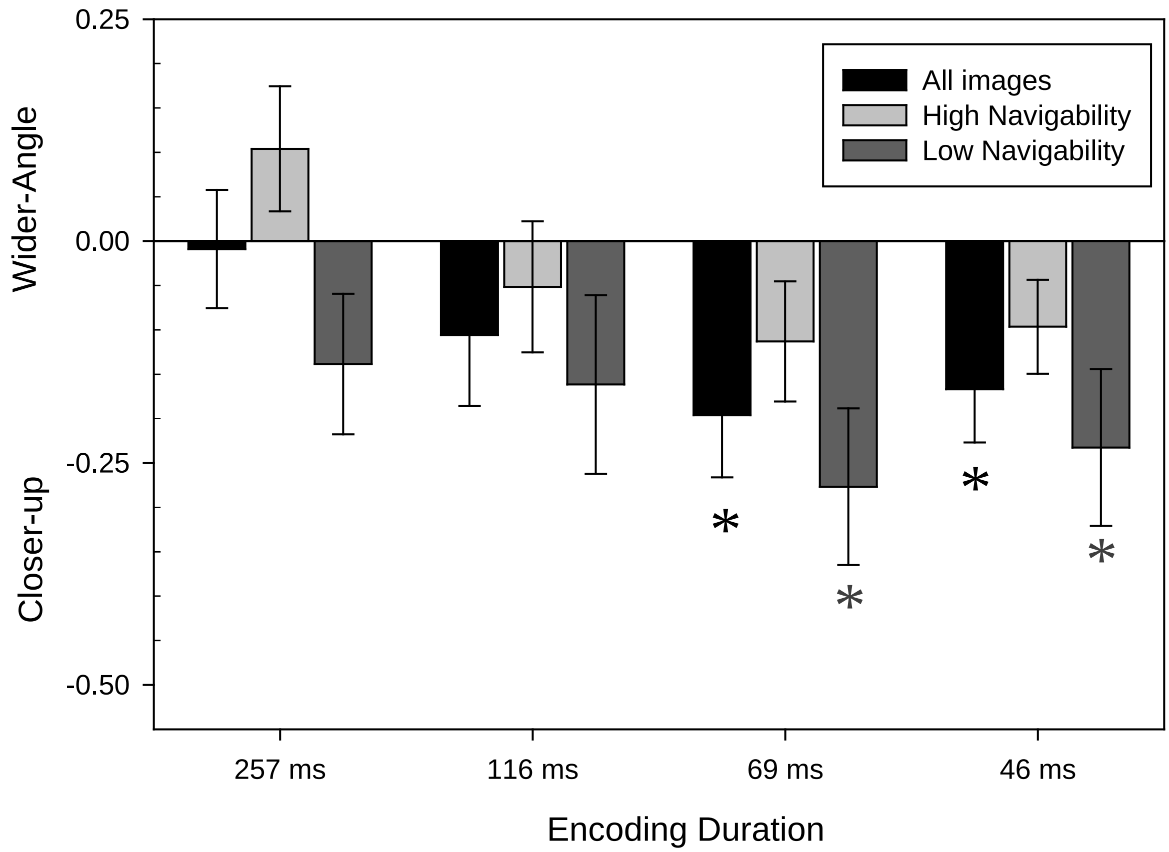

Figure 10. Boundary extension ratings by image type for Exp.3 as a function of encoding duration. Black bars represent all data, excluding blank trials and images given a "don't remember" response. The light grey bars in the middle of each set represent average boundary extension ratings given to scenes which displayed a high degree of navigability, and the dark grey bars on the right of each set represent average boundary extension ratings given to scenes which displayed a low degree of navigability.

* p < .05

Table 1. Mean values for boundary extension ratings in Experiment 1a and 1b. *(Standard error in parentheses)*

Presentation State	Exp. 1a	Exp. 1b: Depth		Exp. 1b: Navigability	
		High	Low	High	Low
CC	-0.08 (.04)	0.08 (.08)	-0.28 (.07)**	-0.001 (.09)	-0.12 (.05)*
WW	-0.09 (.04) *	-0.01 (.09)	-0.24 (.06)**	-0.02 (.07)	-0.08 (.09)
CW	0.37 (.06) [a]	0.53 (.10)	0.34 (.10) [a]	0.40 (.09)	0.39 (.08) [a]
WC	-0.54 (.07) [a]	-0.48 (.13)	-0.84 (.09) [a]	-0.46 (.12)	-0.57 (.11) [a]

** $p < .0001$

* $p < .05$

[a] Absolute values $p < .05$

Table 2. Mean values for the perceptual judgment task and boundary extension ratings in Experiment 2. Asterisks indicate the value is significantly different from zero. *(Standard error in parentheses)*

Encoding Duration	Proportion Correct	Exp. 2: Depth Boundary Extension Ratings		
		Overall	High	Low
257 ms	.763 (.03)	-0.19 (.04)**	-0.02 (.04)	-0.34 (.06)**
116 ms	.805 (.02)	-0.20 (.06)*	-0.05 (.06)	-0.33 (.07)**
69 ms	.790 (.02)	-0.20 (.05)**	0.01 (.07)	-0.42 (.07)**
46 ms	.650 (.02)	-0.10 (.05)	0.12 (.07)	-0.31 (.07)**

** $p < .0001$

* $p < .001$

Table 3. Mean values for the perceptual judgment task and boundary extension ratings in Experiment 3. Asterisks indicate the value is significantly different from zero *(Standard error in parentheses)*

Encoding Duration	Proportion Correct	Exp. 3: Navigability Boundary Extension Ratings		
		Overall	High	Low
257 ms	.871 (.01)	-0.01 (.07)	0.10 (.07)	-0.14 (.08)
116 ms	.881 (.02)	-0.11 (.08)	-0.05 (.07)	-0.16 (.10)
69 ms	.806 (.02)	-0.20 (.07)**	-0.11 (.07)	-0.28 (.09)**
46 ms	.653 (.03)	-0.17 (.06)**	-0.10 (.05)	-0.23 (.09)*

** $p < .01$

* $p < .05$

Appendix A

Greene and Oliva, 2009a, Table 1

TABLE 1

Descriptions of the Global Properties, as Presented to Participants in the Experiment

Global property	Target description	Nontarget description
Concealment	The scene contains many accessible hiding spots, and there may be hidden objects in the scene.	If standing in the scene, one would be easily seen.
Mean depth	The scene takes up kilometers of space.	The scene takes up less than a few meters of space.
Naturalness	The scene is a natural environment.	The scene is a man-made, urban environment.
Navigability	The scene contains a very obvious path that is free of obstacles.	The scene contains many obstacles or difficult terrain.
Openness	The scene has a clear horizon line with few obstacles.	The scene is closed, with no discernible horizon line.
Temperature	The scene environment depicted is a hot place.	The scene environment depicted is a cold place.
Transience	One would see motion in a video made from this scene.	The scene is not changing, except for patterns of daylight.

Greene and Oliva, 2009b, Table 1

Table 1

Description of the seven global properties of natural scenes used in Experiments 1,2 and 3

Structural properties
Openness [1,2,3,4] represents the magnitude of spatial enclosure. At one pole, there is a clear horizon and no occluders. At the other pole, the scene is enclosed and bound by surfaces, textures and objects. Openness decreases when the number of boundary elements increases
Expansion [1] refers to the degree of linear perspective in the scene. It ranges from a flat view on a surface to an environment with strong parallel lines converging on a vanishing point
Mean depth [1,3] corresponds to the scale or size of the space, ranging from a close-up view on single surfaces or object to panoramic scenes
Constancy properties
Temperature [2,4] refers to the physical temperature of the environment if the observer was immersed in the scene. In orther words, it refers to how hot or cold an observer would feel inside the depicted place.
Temperature [4,5,7] refers to the rate at which the environment depicted in the image is changing. This can be related to physical movement, such as running water or rustling leaves. It can also refer to the transience of the scene itself (fog is lifing, sun is setting). At one extreme, the scent identify is changing only in geological time, and at the other, the identity depends on the photograph being taken at the exact moment.
Functional properties
Concealment [4,6] refers to how efficiently and completely a human would be able to hide in a space, or the probability of hidden elements in the scene that would be difficult to search for. It ranges from complete exposures in a sparse space to complete concealment due to dense and variable surfaces and objects.
Navigability [2,4,5] corresponds to the ease of self-propelled movement through the scene. This ranges from complete impenetrability of the space due to clutter, obstacles or treacherous conditions to free movement in any direction without obstacle.

The numbers refer to additional references describing the properties ([1] Oliva and Torralba (2001); [2] Gibson (1979); [3] Torralba and Oliva (2002); [4] Greene and Oliva (2006); [5] Kaplan 91992); [6] Appleton (1975)).

Appendix B

Images used in the practice trials for Exp. 2 and 3 were collected from the SUN database and piloted to exhibit a high or low degree of a particular global property. One-hundred participants rated the mean depth and navigability of 60 images on a 9-pt scale, respectively. Images high in mean depth would be described as a scene that “takes up kilometers of space” (9), as compared with a scene taking up “less than a few meters of space” (1). Images high in navigability would be described as a scene containing “an obvious path that is free of obstacles” (9), as compared with a scene containing “many obstacles, or difficult terrain” (1). For each image, an average was taken for each property and rank ordered. Images in the top 25% (upper quartile) represented scenes with high characteristics of the given global property and images in the bottom 25% (lower quartile) represented scenes with low characteristics of the given global property.