Memory & Cognition

Object-based attention during scene perception elicits boundary contraction in memory

Journal:	*Memory & Cognition*
Manuscript ID	Draft
Manuscript Type:	Special Issue/Invited Manuscript
Date Submitted by the Author:	n/a
Complete List of Authors:	Hall, Elizabeth; University of California Davis, Psychology; UC Davis Geng, Joy; University of California Davis, Center for Mind and Brain; University of California Davis, Psychology
Keywords:	object-based attention, boundary transformation, functional visual field, scene memory, drawings

Title: Object-based attention during scene perception elicits boundary contraction in memory

Authors: Elizabeth H. Hall (1,2) & Joy J. Geng (1,2)

Affiliations:

1 – Department of Psychology, University of California Davis; Davis, CA 95616

2 – Center for Mind and Brain, University of California Davis; Davis, CA 95618

Corresponding Author: Elizabeth H. Hall, ehhall@ucdavis.edu

Conflict of interest statement: The authors declare no competing financial interests.

Acknowledgements: This work was supported by R01 MH113855 to JJG and a NDSEG fellowship to EHH. We thank Maya Tochimoto, Katherine Foray, and Zoe Hareng for their help scoring the drawings, and Tiffany Kim for her help running the experiment. We'd also like to thank the authors of Boettcher et al. (2018) for sharing their stimuli with us. Drawings and code will be made publicly available on the Open Science Framework (link TBD).

Abstract

Boundary contraction and extension are two types of scene transformations that are equally likely to occur in memory. In extension, viewers will extrapolate information beyond the edges of the image, whereas in contraction, viewers will forget information near the edges. Recent work suggests that image composition influences the direction and magnitude of boundary transformation. We hypothesize that the size of the functional visual field (FVF) at encoding may also drive transformation in memory, with constrained FVFs eliciting more contraction and wide FVFs eliciting more extension. We investigate whether attentional FVFs around an object during encoding leads to boundary contraction. One group (N=36) memorized 15 scenes while searching for targets, while a separate group (N=36) just memorized the scenes. Both groups then drew the scenes from memory with as much object and spatial detail as they could remember. We asked online workers to provide ratings of boundary transformations in the drawings, as well as how many objects they contained and the precision of remembered object size and location. We found that Search condition drawings showed significantly greater boundary contraction than drawings of the same scenes in the Memorize condition. Search drawings were significantly more likely to contain target objects, and the likelihood to recall other objects in the scene decreased as a function of their distance from the target. These findings suggest that tight FVFs around objects at encoding will lead to significant boundary contraction.

Introduction

Two types of boundary transformation, contraction and extension, are known to occur in memory. Boundary extension occurs when people remember objects and details that extend beyond the original edges of the studied image (Intraub & Richardson, 1989). In contrast, boundary contraction happens when people fail to remember objects located near the edge of a scene and report the boundaries of the scene as being more constricted than they were in the original image. Recent studies have shown that the direction of the transformation (extension or contraction) depends on the image composition (Bainbridge & Baker, 2020). Object-oriented images tend to elicit more extension in memory, whereas scene-oriented images tend to elicit more contraction (Bainbridge & Baker, 2020; Greene & Trivedi, 2022; Hafri et al., 2022; Lin et al., 2022; J. Park et al., 2021). Even more recent work suggests that boundary contraction and extension result from the magnitude of object and scene-based affordances revealed by the image (Bainbridge & Baker, 2020; Lin et al., 2022; J. Park et al., 2021). Object and scene-affordances revealed by the image also influence what visual information the viewer attends to, directly impacting the size and location of the functional field of view (FVF) at encoding (Wu & Wolfe, 2022). Putting these together, we hypothesize that the selective attention and the size of FVF at encoding is one of the primary drivers of boundary transformation effects, with large FVFs at encoding leading to boundary extension effects, and tight FVFs leading to boundary contraction. Therefore, in this current study, we ask participants to either engage in target search, or simply memorize an image, with the hypothesis that a difficult target search task will tighten the width of the FVF at encoding leading to significant boundary contraction errors in memory.

More than 30 years ago, Intraub and Richardson (1989) made the first report of a consistent pattern of errors where people remembered more details of a scene than what was actually in the original picture. Since then, this pattern of errors, dubbed "boundary extension," has been replicated in numerous studies (Candel et al., 2004; Chadwick et al., 2013; Chapman et al., 2005; Green et al., 2019; Intraub et al., 2008; Lin et al., 2022; Mathews & Mackintosh, 2004; McDunn et al., 2014; Munger & Multhaup, 2016; S. Park et al., 2007; Patel et al., 2022; Kong et al., 2010; Seamon et al., 2002; Wan & Simons, 2004). Recently, Bainbridge and Baker (2020) investigated what types of images elicit boundary transformation errors by testing participants' memory for a large set of object and scene-oriented images. While they replicated the boundary extension effect in some images, they also found that a distinct set of images evoked boundary contraction in memory. Images that resulted in boundary contraction often had a wide-angle view, containing several, dispersed and distinct objects. In contrast, images with extreme close-up views of a central object were more likely to elicit traditional boundary extension errors (Bainbridge & Baker, 2020). These results situated boundary contraction and extension as related phenomena, equally likely to occur depending on the content and composition of the image. Several subsequent studies have replicated the finding of bidirectional transformation errors based on image properties (Bainbridge & Baker, 2020; Lin et al., 2022; J. Park et al., 2021; Greene & Trivedi, 2022; Hafri et al., 2022). In one such study, Park and colleagues (2021) tested participants on a set of stimuli that included viewpoints of the same items taken from different distances, ranging from a very close-up viewpoint (taken a few centimeters away from the focus object) to fairly far away (roughly 4-5 meters from the object). Interestingly, the authors found there was no absolute viewing distance associated with the change between boundary extension and contraction. They instead found that the changepoint between transformation errors was dependent on the content and context of the image. Their results showed that distant viewpoints of object-rich environments tended to elicit more boundary contraction than object-spare environments. They then asked a separate group of participants to provide judgements as to what distance away from the focal point provided the best view of the scene. Participants preferred to view object-rich environments from closer distances, and the authors proposed that this closer viewpoint provided the viewer with higher resolution processing that allowed them to recognize the identity of individual objects. Object-rich environments could have presumably elicited boundary contraction in memory because visual processing at the time of encoding was object-centered. Additionally, participants preferred to view object sparse environments from farther distances, which could allow them to perceive more spatial information, such as the openness, concavity, and/or navigability of the scene (Bonner & Epstein, 2017; Cheng et al., 2019; Kravitz et al., 2011). These results suggested that both depth, and the affordances provided by the environment, could influence the direction of boundary transformation. The authors also raised the idea that endogenous factors like attention could influence the trend of transformation, along with the composition and the inherent affordances of the image.

In scene perception, attention is often guided to regions by exogenous cues (like image composition and salience) and endogenous motivation (Bonacci et al., 2020; Posner, 1980; Posner et al., 1980; Wolfe & Utochkin, 2019). Viewers will often move their eyes to fixate on new regions, as information within the region visible by the fovea is processed in higher acuity (Low, 1951). The portion of the scene that is visible around the point of fixation by both the foveal and parafoveal vision is known as the functional visual field or FVF (Sanders, 1970). This area is constrained by limits of visual acuity, or the extent of the field in which a target can be localized and identified (Hulleman & Olivers, 2017; Wu & Wolfe, 2022; Low, 1951; Daniel & Whitteridge, 1961; Wilson et al., 1990; Schütz et al., 2009). It can also be constrained by limits of attentional resolution when processing images that have a high degree of competition between representations (Belopolsky & Theeuwes, 2010; Hulleman & Olivers, 2017; Wu & Wolfe, 2022). The combined limits imposed by sensory and attentional constraints leads to a region within some distance of a point of fixation in which a set of items can be covertly attended and processed (Belopolsky & Theeuwes, 2010; Hulleman & Olivers, 2017; Wu & Wolfe, 2022). While labeled as both the FVF and the attentional FVF, here we refer to this region as the attentional functional visual field (AFVF). The size of the AFVF changes with the discriminability of objects within the field. If an object is difficult to distinguish, the AFVF will be small. If objects are easy to identify, the AFVF will be wider. A tightened AFVF helps to allocate visual processing resources to information within the AFVF. It does this by filtering out peripheral information preventing it from reaching the receptive fields of extrastriate neurons (Akbas & Eckstein, 2017; Moran & Desimone, 1985; Shamsi et al., 2021; Wu & Wolfe, 2022; Yao et al., 2011). This leads to a high resolution perception of information within the AFVF, and a more blurry or fuzzy perceptual representation of information peripheral to the AFVF (Akbas & Eckstein, 2017; Moran & Desimone, 1985; Shamsi et al., 2021; Wu & Wolfe, 2022; Yao et al., 2011). The visual system, in turn, is tuned to create continuous representation of an environment from discontinuous samples (Bonner & Epstein, 2017; Kanizsa & Gerbino, 1982; Kourtzi & Kanwisher, 2001; Mendola et al., 1999; S. Park et al., 2007). It forms a continuous representation of discrete samples by using available sensory information and statistically constrained predictions to “perceive” the information outside of the AFVF (Intraub, 2010; S. Park et al., 2007; Rosenholtz, 2016, 2020; Wu & Wolfe, 2022; Kersten et al., 2004) although these predicted perceptions are not always a veridical representation of the scene (Wu & Wolfe, 2022).

It is possible that the distinction between images that elicit boundary extension and those that elicit contraction relies on the type of visual information attended to within the AFVF. A single, close-up object in an image is most likely easily identifiable, and so would not require a highly constrained AFVF for the viewer to know the content of the image. Instead, because this single, easily identifiable object percept takes up most of the image, viewers will have a wide AFVF from the point of fixation. As information outside of the AFVF is perceived at a lower resolution, the viewer may encode peripheral information like the distance between the central object and the border with less veracity. While boundary extension effects are most easily elicited by images of a single, close-up object (Bainbridge & Baker, 2020; Lin et al., 2022; J. Park et al., 2021), an images of a distant object filtered to become more focal in the image will also elicit boundary extension (Hafri et al., 2022). Likewise, images of a few objects grouped together in the scene could also easily be processed within a single AFVF (Jeurissen et al., 2016; Kim & Cave, 2001; Schneider & Mattes, 2022). Distances between multiple contiguously attended objects will also be compressed during perception (Liverence & Scholl, 2011). Taken together, extension effects could derive from a wide AFVF at encoding that captures most of the object information in the scene under a single upweighted precept, eliciting extension because the information outside of that percept is encoded with less veracity.

The impact of selective attention on memory becomes more visible when there is competition between stimuli in the environment (Chun & Turk-Browne, 2007). When viewing images with wide-angle views distant objects become physically smaller on the screen and are less easily identifiable. Fine object details can only be discriminated efficiently in the fovea (Low, 1951; Daniel & Whitteridge, 1961; Wilson et al., 1990; Schütz et al., 2009). Selectively attending to fine object details tightens the constraints of the AFVF, by increasing foveal load, and decreasing allocation of attention to peripheral information leading to peripheral blur or tunnel vision (Akbas & Eckstein, 2017; Moran & Desimone, 1985; Shamsi et al., 2021; Wu & Wolfe, 2022; Yao et al., 2011; Pramod et al., 2022, Ringer et al., 2016). Importantly, fine object details need to be foveated during scene exploration to be encoded in to memory (Bridgeman et al., 1975; Grimes, 1996; Mack & Rock, 1998; Nelson & Loftus, 1980; Rensink et al., 1997; Simons & Rensink, 2005; Geringswald et al., 2016). With short scene presentation times, participants may only be able to foveate one or two objects in the scene. All these factors will limit the viewer to only be able to encode a small portion of the total scene image. As viewers typically encode upweighted representation of object information from closer vantage points, they may misremember the distance between the foveated object and their vantage point as closer than it originally appeared.

As noted by Park and colleagues (2021), factors outside of depth and the objects’ spatial layout can also elicit contraction in memory. Recently, Greene and Trivedi (2022) quantified the amount of semantic density of scenes as a component representing the median description length of the scene, as well as the average lexical entropy, and average pairwise similarity of semantic labels for the objects in the scene. Their results revealed that semantically rich scenes will also elicit substantial contraction effects in memory. These results are not necessarily contradictory to our hypothesis that the size of the AFVF at encoding is a primary driver of the boundary transformation. Indeed, semantic object information plays a special role when it comes to resource allocation in attention and working memory. Objects that share a semantic label can be grouped together in a multiple-identity tracking task, suggesting that shared semantic labels allow attention to extend across multiple objects more easily (Wei et al., 2018). Participants can also hold more objects in working memory when they can be indexed by a semantic label (Brady et al., 2016; Brady & Störmer, 2022). It is therefore probable that the semantic identity of objects plays a role along with image composition and the goals of the viewer in manipulating the size of the AFVF at encoding. Another property that has been known to elicit contraction effects is the emotional valence revealed objects in the image. Negatively valent stimuli have been shown in attention tasks to constrain the AFVF more tightly than positive or neutral stimuli (Fernandes et al., 2011; Masuda, 2015). Scenes with highly negatively valent items like weapons or injured body parts have also been shown to elicit significant contraction effects in memory (S. A. Christianson, 1984; S.-Å. Christianson & Loftus, 1987; Green et al., 2019; Safer et al., 2002; Takarangi et al., 2016; Wonning, 1994). Importantly, these studies also found heightened memory for negatively valent objects found in the scenes compared to neutrally valent objects (S. A. Christianson, 1984; S.-Å. Christianson & Loftus, 1987; Green et al., 2019; Safer et al., 2002; Takarangi et al., 2016; Wonning, 1994) and reduced memory for peripheral details in traumatic images compared to control-matched scenes with neutrally valent objects (S. A. Christianson, 1984). These findings suggest that boundary contraction in memory could have been elicited by constrained attention around the negatively valent objects at encoding. Taken together, these findings point to how factors like the visual and semantic density of objects in scenes, and their emotional valence can constrain the size of the attentional field, leading to boundary contraction in memory.

In this study we attempt to control for the influence of image composition by keeping the stimuli consistent across conditions and manipulating only the goals of the viewer. To do so, we recruited one group of participants to engage in a target search task while simultaneously memorizing the scene image. During search, attention will become significantly more constrained when the viewer is attempting to determine if the viewed object is the target object (Wu & Wolfe, 2022; Young & Hulleman, 2013; Yu et al., 2022). In our search task, participants are required to fixate the target object when they find it and give a button response that will immediately end the trial. We hypothesize that constraining attention around the target object at encoding will lead to boundary contraction effects in memory, namely that the participants will recall the scenes as being closer than they were originally portrayed, show an enhanced memory for target objects, and show a degraded memory for objects outside the target. We directly compare the memory performance of participants in the target search condition with a separate group of participants who were asked only to memorize the scenes, and had both groups draw the scenes from memory after a short delay. We analyzed their drawings for a variety of metrics including what objects were recalled, their drawn location and size, and transformation of scene boundaries. Results revealed significant boundary contraction in memory drawings from participants who engaged in Search, while drawings from the Memorize condition showed equal rates of contraction and extension. Further analyses of Search drawings compared, revealed enhanced memory for target objects and diminished memory for objects farther away from the target object. These findings provide evidence that constraining attention at encoding, such as those employed when making match decisions in target search, can lead to substantial boundary contraction in memory.

Method

Participants. Thirty-six undergraduate students (26 females, mean age = 19.44, SD = 1.34, range = 18-23 years) participated in the Search condition and thirty-six different students participated in the Memorize condition (27 females, mean age = 19.94, SD = 1.67, range = 18-25 years). Students were recruited from the University of California, Davis through the Sona research pool in exchange for research credit. Participants were native English speakers with normal or corrected vision. We also recruited online scorers to judge the drawings on a variety of metrics. Five-hundred and seventy-nine scorers were collected from Amazon Mechanical Turk and were monetarily compensated. One hundred and sixty-four scorers were collected from the SONA research pool to complete ratings on Testable for course credit. Each participant provided informed written consent in accordance with the local ethics clearance as approved by the National Institutes of Health.

Stimuli. The 15 scene images used in this study were initially constructed for an experiment assessing the role of anchor objects on eye movements in visual search (Boettcher et al., 2018). The stimulus images were created with ArchiCAD software version 18 (Graphisoft, Munich, Germany). All images were 1280 wide by 960 pixels tall. Each scene contained a visual search target (e.g., toilet paper). The scenes were selected so that there was no overlap in target objects across the scenes, and each scene could be identified by a unique categorical identifier (i.e., there was only one kitchen in our stimulus set).

Apparatus. Stimuli were presented on a ASUS MG279Q monitor with a 60 Hz refresh rate and a spatial resolution of 1920 x 1200 pixels. Participants were seated 60cm away from the screen and a computer running PsychoPy (Peirce, 2007) controlled all stimulus presentations. Eye movements were tracked using an EyeLink-1000 desktop mount, sampling from the right eye at 500 Hz (SR Research, Ontario, Canada).

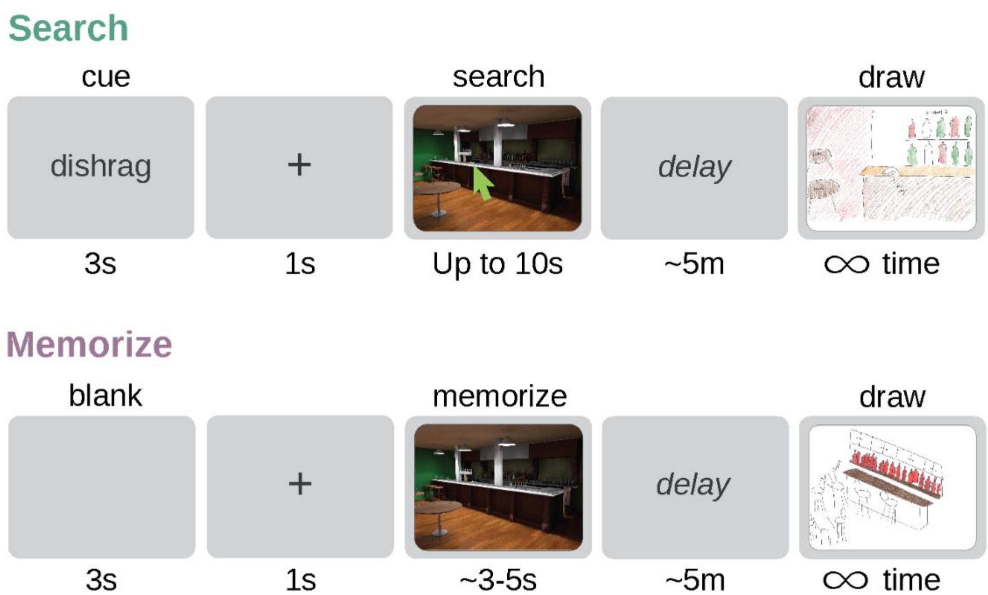

Figure 1. In the Search condition, participants (N=36) were given target cues and had up to 10 seconds to find the cued objects in 15 scene images. In the Memorize condition (N=36) participants memorized each scene for the average amount of time it was viewed by participants in the Search condition. After a delay, both groups of participants had an unlimited amount of time to draw the scenes from memory.

Experimental design. The visual Search group was run before the Memorize group so that scene exposure times from the Search group could be used to constrain viewing time in the Memorize group. All saw the same 15 computer-generated scenes and both groups completed 1 practice trial with a scene that was not from the main experimental set. The Search group was instructed to search for and click on a specific target. On each trial, a word describing the target appeared for 3s followed by a fixation cross that they were instructed to fixate for 1s. Once the image appeared, they were given up to 10 seconds to click on the target with the computer mouse. They were also instructed to memorize the scene in as much detail as possible since their memory for the images would later be tested, though specific details of the memory test were not provided. Participants were then asked to complete the Visual Vivid Imagery Questionnaire which contains questions regarding their ability to visualize images (Marks, 1973). This task was used to limit rehearsal of the scenes and items in memory and an average of 4.56 minutes (SD = 1.48 minutes) passed from the end of the eye-tracking phase to the start of the drawing phase. After the VVIQ, participants were instructed to draw as many scenes as they could recall, in no particular order and with no time limit, while their pen movements were tracked on a digital drawing pad. They were provided with 15 sheets of paper each with a 1280 x 960 black frame and were instructed to draw every detail they could remember about the scene within the frame. They were told that the drawings would not be scored on the basis of their drawing ability but would be scored on how accurately they were as a representation of the studied stimulus images. If they felt they could not accurately draw an item in the scene, they were instructed to try to draw the general shape, and they could label anything they felt was unclear. They could use color pencils to add any color they remembered. In the Memorize group, participants were instructed to memorize each scene in as much detail as possible as their memory would be tested later on, and they saw each scene for the average time that the participants viewed the scene in the Search experiment (M = 4.15s, SD = 0.96s, MIN = 3.23s, MAX = 6.06s). They were not instructed to search for the target or click the image and did not see the target word before each image but instead saw a blank screen for 3s followed by a 1s fixation cross. An average of 4.22 minutes (SD = 1.33 minutes) passed from the end of the eye-tracking phase to the start of the drawing phase.

Eye-tracking analysis. Fixations and saccades were defined from raw eye-tracking data using the Saccades package in R (von der Malsburg, 2015). Fixations could not be determined for one participant from each condition due to poor data quality. We included the drawings from these two participants in analyses but discarded their eye movement data.

Online scoring procedures

The 72 in-lab participants drew 601 scenes from memory. Three scorers, the first author and two undergraduate research assistants, matched each drawing to a scene image. A drawing was considered to be matched to an image if two out of three scorers agreed. If the scorers believed that a participant drew the same stimuli image more than once, the first drawing of that scene was considered a match, and subsequent drawings of the same image were not included in analyses. Drawings that were not matched to an original image by the experimenters were not scored (82 out of 601 drawings, or 13.64% of drawings), leaving 519 drawings for analyses. Of the 87 unmatched drawings, 32 were of the practice trial image. Three different measures were collected for each drawing. The code for these measures was adapted from Bainbridge et al. (2019).

Boundary transformation. Forty-four scorers were recruited from the SONA research pool to provide ratings of boundary transformation for each drawing on Testable. Scorers were shown the drawing and the originally viewed stimulus image side-by-side on the screen. Scorers were asked whether the drawing was “closer, the same, or farther than the original photograph,” and were told to ignore any extra or missing objects in the drawing. Scorers responded on a 5-item scale: much closer, slightly closer, the same distance, slightly farther and much farther, with the additional option to indicate “can’t tell” if they believed the drawing to be incomprehensible. Seven scorers provided boundary ratings for each drawing and boundary transformation scores for each drawing were calculated by the mean across the ratings normed on a scale of -1 (much farther) to +1 (much closer).

Object marking. One hundred and twenty scorers were recruited from the UC Davis SONA research pool to complete an online object marking task on Testable. The purpose of this task was to determine if an object from the original image was included in the drawing or not. All objects in the stimulus images were segmented using LabelMe, an online object annotation tool (Russell et al., 2008). There were 360 objects in the stimulus set and each image contained 24 objects on average (SD = 16.99, min = 8, max = 83). Objects were “nameable, separable, and visually distinct items” (Bainbridge et al., 2021). If multiple objects of the same type were touching, these objects were grouped together and given a plural label (e.g. “shampoos”). Object parts (e.g. “tire” on truck) were not segmented, but if an object was visually distinct and could be defined as a separate semantic label it was segmented separately (i.e. decorative “pillow” on a couch). Background segmentations (“grass”, “trees”, “floor”, “walls”, “ceiling”) were not included in analyses (Bainbridge et al., 2019). Scorers were shown the original image with an object outlined in red using the LabelMe annotations presented next to a drawing. Scorers were asked to indicate if the outlined object was included on the drawing. Scores were collected from three participants per object and an object was determined to be in the drawing if at least two out of three participants agreed that it was present.

Object location and size. Five-hundred and seventy-nine scorers were recruited on Amazon Mechanical Turk to complete an online object location task. The purpose of this task was to quantify the location and size of drawn objects. Only objects that had been determined to be present in the drawing by the object marking task were scored in this task. Scorers were shown an original image with an object outlined in red next to a drawing and asked to place and resize an ellipse around the same object in the drawing. Three scorers were asked to locate each object of interest in a drawing. Object location was calculated as the median centroid of the ellipses across the responses. Object size was calculated as the median radii of the ellipses across responses.

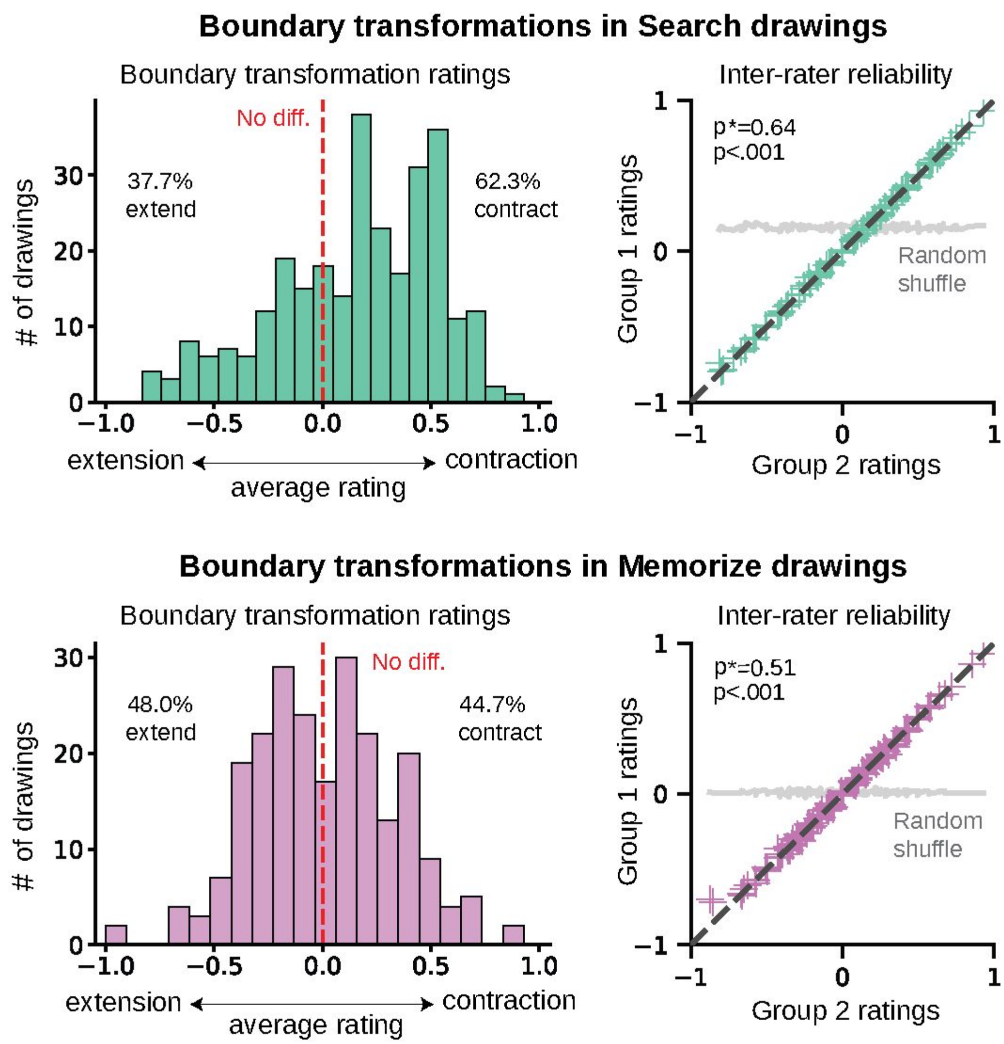

Figure 2. (left) Histograms of boundary transformations in the memory drawings by condition. On average, 62.3% of drawings in the Search condition showed boundary contraction, a significantly greater proportion than the 37.7% that showed extension. Only 44.7% of drawings in the Memorize condition showed contraction, while 48.0% showed extension, and there was no significant difference between the proportions. (right) Results of the split-half consistency analyses for each condition. Seven different raters

scored the amount of boundary transformation in each drawing. Each set of ratings was split in half, and we calculated the correlation between the average transformation score of each half. The gray line shows the other half of ratings sorted randomly. For both conditions, ratings between groups were highly similar and significantly correlated.

Results

The main question of this study was whether the patterns of object-based attention used in search would elicit contraction effects in memory above the rate elicited by the image alone. To investigate this we had an experimentally-naive group of online scorers rate the degree of boundary transformation in the Search and Memorize drawings. We visualized what percent of drawings from each condition showed either boundary contraction or extension (figure 2). To start, we found that a majority of Search drawings had contracted scene boundaries. On average, 62.29% of Search drawings showed boundary contraction, while only 30.73% showed boundary extension. The results from a chi-square test of independence confirmed that the difference between proportions was significant (χ^2 (1, N = 283) = 28.0, $p < .001$). Comparatively, only 44.67% of the drawings from the Memorize condition showed boundary contraction, with 47.95% of drawings showing boundary extension, and the chi-square test revealed no significant difference between the proportions (χ^2 (1, N = 232) = 0.04, $p > .5$). The transformations ratings for the Memorize drawings are consistent with the findings of Bainbridge and Baker (2021) who revealed that scene images have a fairly equal probability of eliciting either contraction or extension. To assess the reliability of the ratings we conducted a split-half analysis across 1,000 iterations and applied the Spearman-Brown correction formula (figure 2). Boundary transformation ratings were highly consistent across raters' responses for both the Search (ρ^* = 0.64; $p < .001$) and Memorize drawings (ρ^* = 0.51; $p < .001$). We then looked at the difference in boundary contraction scores averaged by scene image across conditions figure 3b). This analysis allows us to directly compare the effect of the task on memory, as both groups of participants studied the scenes for roughly the same amount of time. Results from a non-parametric Wilcoxon rank-sums test (WRST) confirmed that scene images were significantly more likely to elicit boundary contraction in memory when participants engaged in target Search during the encoding period (N = 15, Z = 2.32, $p = .020$). Taken together, these findings suggest that object-based attention during scene perception can elicit boundary contraction in memory.

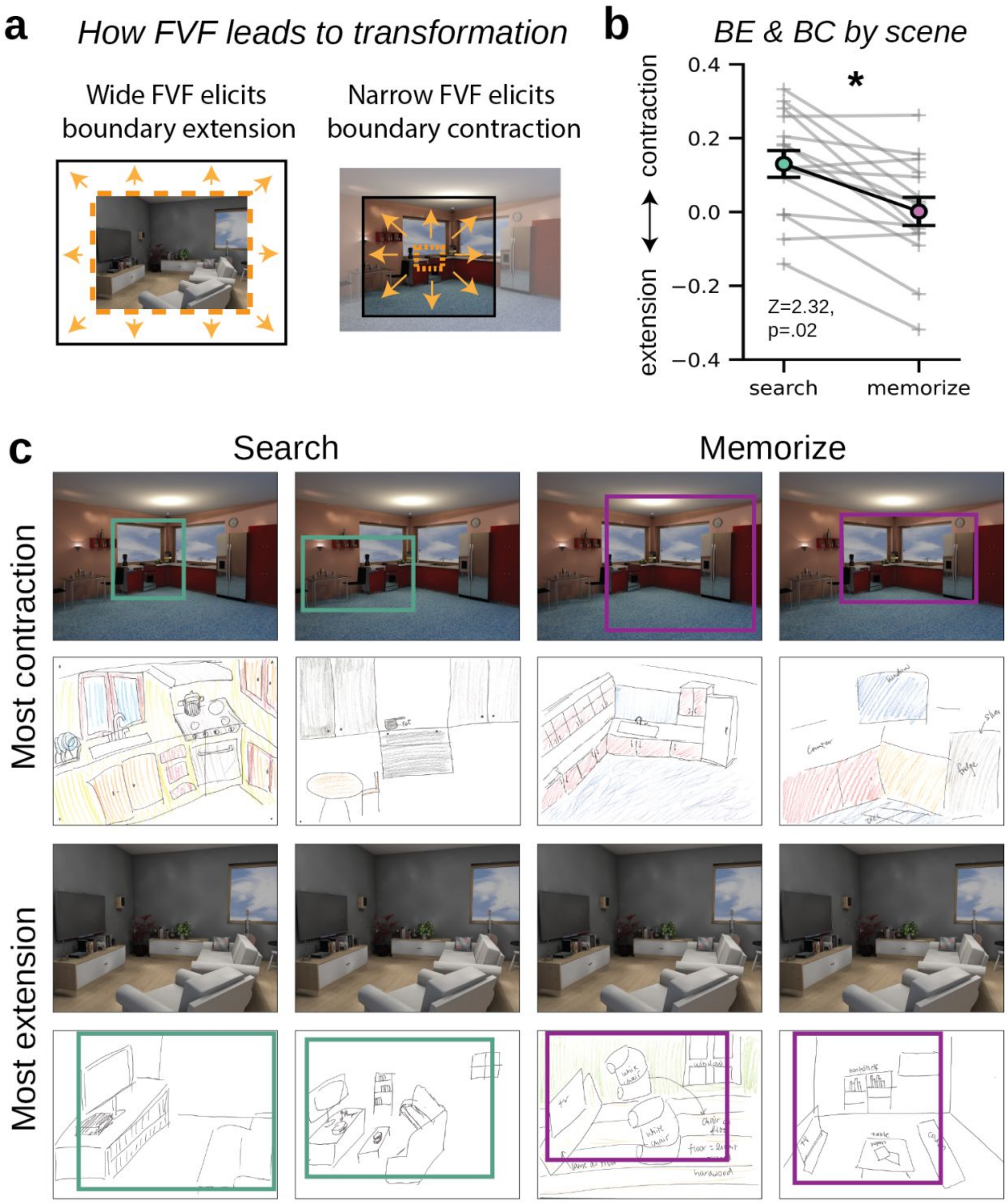

Figure 3. (a) Figure showing how the size of the attentional FVF at encoding can influence boundary transformations. The orange box is the FVF and the black box is what the participant remembers of the scene. When the FVF is narrow, boundaries could extend beyond target FVF in memory, but because the FVF is a relatively small proportion of the image, a majority of the scene is forgotten. **(b)** Plot of the average boundary transformation rating across drawings by image. Each gray line represents one of the 15 scene images. Error bars represent the standard error of the mean. Drawings of scenes done in the Search condition showed significantly more boundary contraction on average. **(c)** Example drawings of the scenes that elicited the most boundary contraction and boundary extension for the Search (left) and Memorize (right) conditions. For drawings with the most contraction (top), the colored outlines on the images show how much of the scene the participant recalled. For drawings with the most extension, the colored outlines on the drawings show the boundaries of the studied image. Area outside of the boundaries is what was extended in memory.

In boundary contraction, viewers often fail to remember objects located near the edge of the scene. However, for participants in our Search condition, their representation in memory for a scene might be centered around the target object instead of the image center. Therefore, they might have been more likely to fail to remember objects located farther away from the target object. We conducted a series of analyses to better understand what information is encompassed in a memory representation, when it is shifted off the center of the image. First, we looked to confirm that target objects were included in a significant amount of the memory representations for those who engaged in Search (figure 4b). To do so, we showed a separate group of online workers the drawing alongside the stimulus image and asked them to make a judgment as to whether an outlined object on the image was present in the drawing. From these judgments we confirmed that a majority of Search drawings contained the target object, with the target present in 80.36% (SD = 16.09%) of drawings on average. In comparison, memory representations from the Memorize group, who spent only 3% of the trial time looking at the target on average, contained the target only 19.39% (SD = 18.85%) of the time on average (N = 15, Z = 4.50, $p < .001$). We next tested whether a model of the memory representation focused around the target object could explain the detail revealed in the Search drawings (figure 4c). We fit a logistic regression model on the recall data from drawings done in the Search condition (N = 6,135 objects [1,115 objects drawn, 5,050 objects not drawn]). Proximity was defined as the distance from the center of the object segmentation to the center of the target object segmentation, z-scored across objects. We found that proximity to the target object significantly predicted whether an object would be included in the memory drawing ($\beta_0 = -1.51$, CI = [-1.58, -1.45], Z = -45.61, $p <.001$; proximity: $\beta = 0.11$, CI = [0.05, 0.18], Z = 3.33, $p = .001$). These findings suggest that some participants from the Search condition may have maintained a representation of the scene in memory that was focused around the target object instead of the image center.

Additionally, in a qualitative review of the drawings we saw that drawings from both conditions showed a tendency to include large, space-defining objects. Results of a logistic regression further confirmed that object memory was significantly predicted by both object size (n=12,230 objects [3,055 drawn / 9,175 not drawn]; $\beta_0= -0.87$, CI = [-0.93, -0.81], Z = -29.60, $p <.001$; size: $\beta = 0.41$, CI = [0.35,0.47], Z = 12.90, $p <.001$) and the Search condition ($\beta = -0.52$, CI = [-0.61, -0.44], Z = -12.05, $p <.001$). The interaction between size and condition was also significant in the model ($\beta = 0.10$, CI = [0.01, 0.19], Z = 2.28, $p = .022$). Looking at a plot of the model fit (figure 4d) we see that Memorize drawings show a stronger likelihood of including objects that fall within the first five standard deviations of mean object size. The interaction between object size and condition occurs as object size becomes greater than five standards above the mean, when Search drawings start to show a greater likelihood of drawing the object from memory.

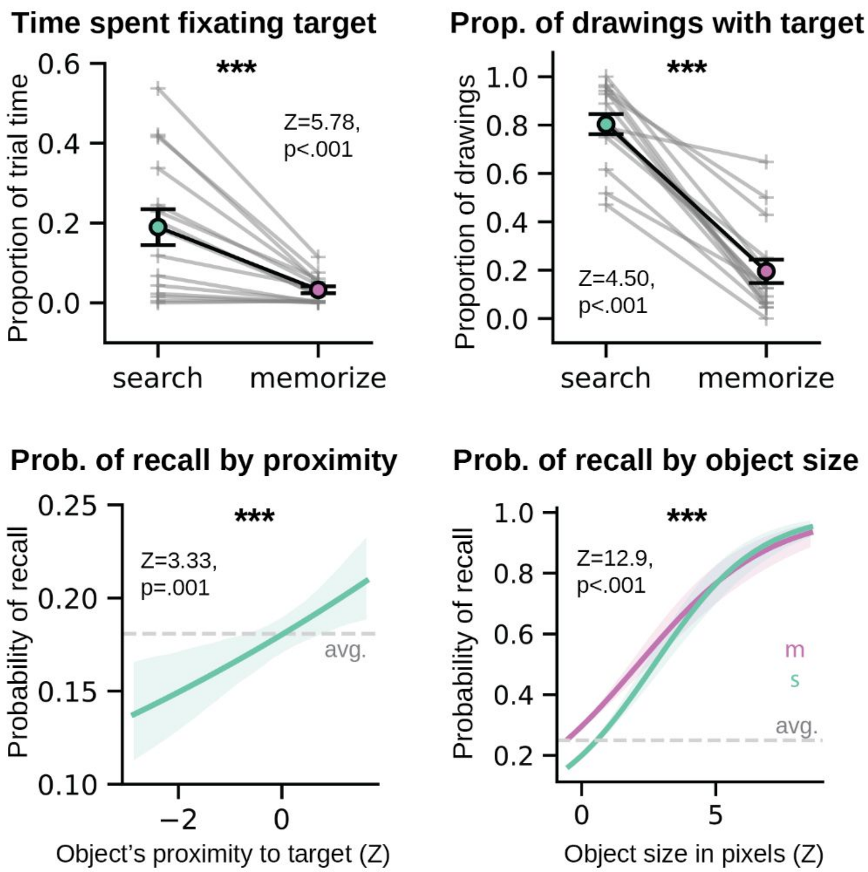

Figure 4. (top left) Plot of the proportion of trial time spent fixating the target object. Participants in the Search condition spent a significantly greater proportion of trial time fixating the target. (top right) Plot of the proportion of drawings that included the target object. A significantly greater proportion of Search drawings contained the target object. Each gray line represents one of the 15 scene images. Error bars represent the standard error of the mean. (bottom left) Regression line indicating the probability of recalling an object in the Search condition by its proximity to the target object. (bottom right) Regression lines indicating the probability of recalling an object by its size and condition. Shaded error bars are the confidence interval bootstrapped across 1000 iterations.

The drawings also provided some insight into how participants in the Search represented the targets' size and location in memory. We asked a group of online workers to provide judgements as to an object's height and width by asking them to draw an ellipse around each object in each drawing. We then defined the height and width of a drawn object as the median radii from a set of three judgements. From these judgments, we found that participants in the Search condition consistently drew the target objects from memory as wider (M = 4.55% wider, SD = 2.18% wider) and taller (M = 5.87% taller, SD = 1.76% taller) than they originally appeared. A paired sample signed-rank test of the ellipse size between target objects of the images and target objects in the Search drawings confirmed that target objects were drawn significantly larger than portrayed in the stimuli images (N = 15, Z = 3.30, p = .001). Memorize drawings that included the target object also showed a significant tendency to overestimate the target's size (wider: M = 2.14%, SD = 1.94%; taller: M= 2.84%, SD = 2.02%). However, results from paired sample signed-rank tests of the average increase in target height and width between conditions confirmed that targets were drawn as significantly taller and wider in memory representations from the Search condition compared to those from the Memorize condition (taller: N=14, Z = 2.44, p = 0.015; wider: N = 14, Z = 3.63, p < .001 [we excluded one scene from this analysis as no drawings from the memorize condition of this scene included the target object]). This result points to the idea that when memory representations for a scene are contracted, objects tend to be remembered as larger (or closer) than they originally appeared. The results of our recent study confirmed this prediction for size perception: when attention was allocated at the center of a circular stimulus, that stimulus was perceived as larger as compared with a neutral attentional condition (Kirsch et al., 2018). The finding that target objects were drawn as even larger in Search drawings compared to Memorize drawings lends support to the idea that these drawings represent memories where the boundaries were significantly contracted around the target object.

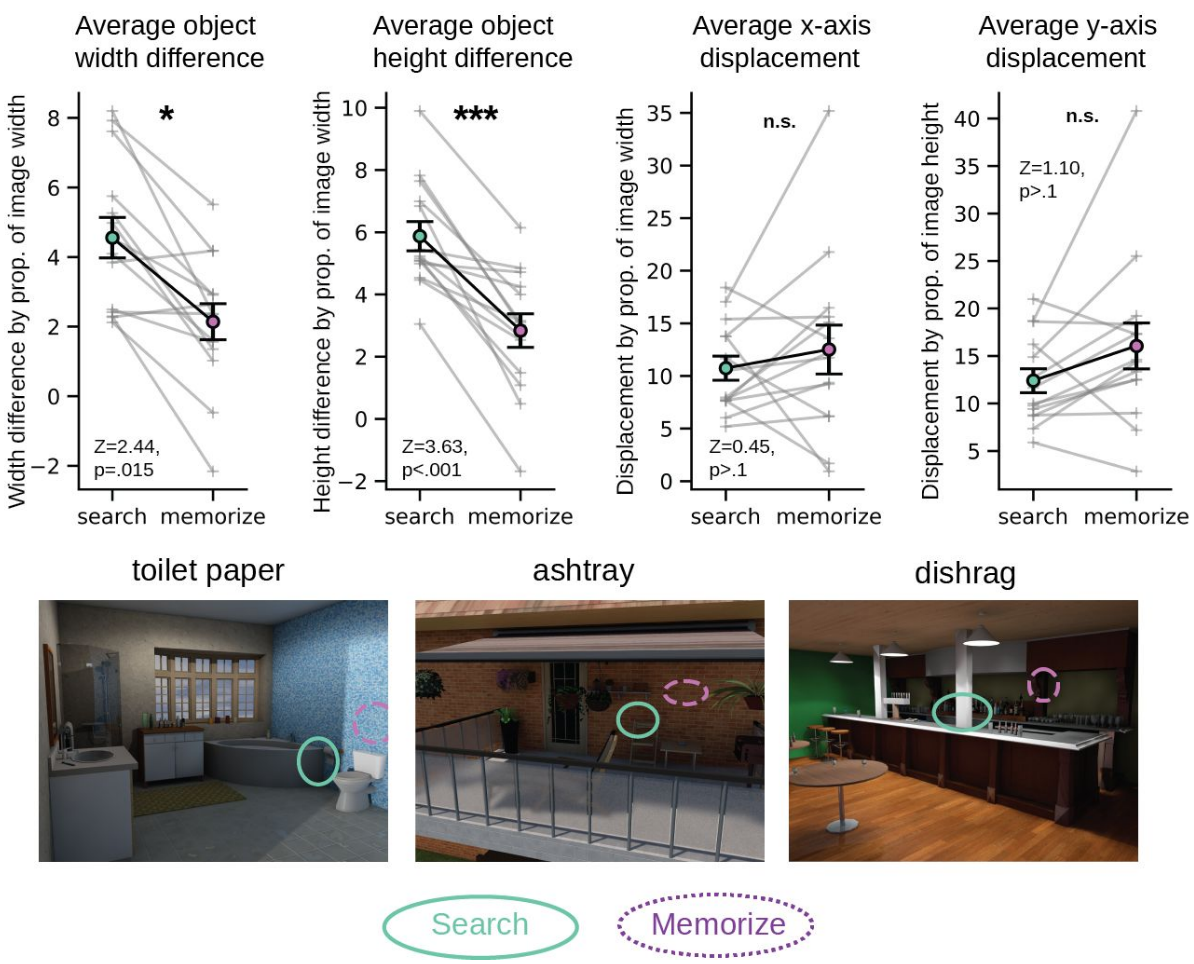

Figure 5. (top left) The mean height and width difference between objects of the different conditions and objects in the original image. (top right) The mean x-axis and y-axis distance between object centroids of the different conditions and object centroids in the original image. Each gray line represents one of the 15 scene images. Error bars represent the standard error of the mean. (bottom) Example maps of the average ellipse encompassing the target objects by condition. Target objects are noted above the scenes.

The ellipse judgements provided by the online raters also provided a metric for measuring whether participants had accurate memory for the objects' locations throughout the scene (figure 5). For this analysis we defined object location as the median centroid from a set of three ellipse judgments. Participants in the Search condition drew the target objects close to where they appeared in the stimulus image, with target objects centroids displaced on average by 10.74% (SD = 4.28%) of the scene overall in the x-direction, and by 12.40% (SD = 4.73%) in the y-direction. A paired sample signed-rank test found that there was no significant difference in where the object centroid was located on the x-axis compared to where it was drawn ($N = 15$, $Z = 0.97$, $p > .1$). A paired sample signed-rank test of the drawn and real target centroid location on the y-axis revealed that there was some significant displacement along that axis ($N = 15$, $Z = 2.22$, $p = 0.03$). Target objects were displaced slightly more in the Memorize drawings (x-axis: M = 12.51%, SD = 8.71%, y-axis: M = 16.05%, SD = 9.03%), although results from paired sample signed rank tests found that this difference in x and y-axis displacement between conditions was not significant (x-axis: $N = 14$, $Z = 0.46$, $p >.1$; y-axis: $N = 14$, $z = 1.10$, $p > .1$). The significant degree of displacement along the y-axis for target object in the Search condition and the lack of significant displacement along the x-axis is likely due to there being more pixels along the x-axis in the stimulus images (the stimuli 1280 pixels wide x 960 pixels tall). Therefore, we should not rule out that participants in the Search condition had some displacement overall for where the target object was located. However, the changes in location could be seen as slight compared to the magnitude of increase in size in memory for target objects.

Discussion

This study sought to provide greater insight into how scene memory is shaped by selective attention at encoding. Specifically, we looked to see if manipulating viewers to constrain spatial attention to support object recognition would impact the magnitude and trend of transformation effects. We hypothesized that the tightly constrained AFVF needed to efficiently search for object targets would lead to significant boundary contraction in memory as information outside of the focus of attention failed to be encoded. We found a high rate of boundary contraction in drawings from participants who engaged in Search, above the rate reported in previous studies (Bainbridge & Baker, 2020; Greene & Trivedi, 2022; Hafri et al., 2022; Lin et al., 2022; J. Park et al., 2021). Further, we found that a group of participants who viewed the same scenes for roughly the same amount of time had significantly less boundary contraction in their drawings, and in fact showed roughly equal rates of contraction and extension. Search drawings were also significantly more likely to contain target objects and objects located near the target object suggesting that boundaries were contracted in memory around the target. However, participants from both scenes were more likely to draw big objects from memory. Finally, we found that target objects were drawn significantly larger than they were portrayed in the stimulus image and showed a significantly greater increase in size in the Search drawings compared to the Memorize drawings. Taken together these results suggest that constrained attention on certain objects during encoding can lead to those objects having an exaggerated role in memory, with boundaries contracting around the object instead of the center of the image.

These findings point to a role for attention in eliciting boundary transformation effects. While scenes with several dispersed objects (like the ones used in our study) have been known to elicit boundary contraction (Bainbridge & Baker, 2020; Greene & Trivedi, 2022; Hafri et al., 2022; Lin et al., 2022; J. Park et al., 2021), we found an increased rate of contraction when we manipulated participants to have highly constrained attention during encoding. We suggest that the reason why images with several distant and dispersed objects elicit the most contraction effects is because viewers must constrain their attention to more effectively recognize these objects. To that point, one of the stimulus images that elicited the most contraction in both conditions was also the one of the images in which objects were the most distant on average (figure 3c). Further, from a deeper analysis of the Search drawings we saw that participants who searched for target objects were more likely to remember objects closer to the target object, and less likely to remember objects farther away from the target. We believe that these results could suggest that for participants in the Search condition boundaries contracted around the target object in memory. This contrasts with previous findings that show that objects near the edge of the image tend to be forgotten with boundary contraction (Bainbridge and Baker, 2020). A qualitative review of contraction effects in Memorize drawings showed that boundaries in their drawings tended to contract around the center of the image. However, participants in the Memorize condition could have been attending to the image center more because they were biased to attend to the most relevant or salient information as both professional and amateur photographs are biased to compose their photographs to have this information near the center of the image (Abeln et al., 2015; Locher et al., 1996). While our results suggest that endogenous attention may cause contraction effects off-center for the image, future studies could more empirically manipulate where salient or relevant (but less easily distinguishable) information is located within the scene. If a majority of participants then draw the scenes from memory as contracting around the salient information in the image, this would lend more precedence to the idea that memories are contracted around information that was preferentially upweighted for object processing.

We also saw that participants in both conditions tended to remember bigger objects more. We don’t believe that this discounts our hypothesis that constrained AFVF leads to contraction effects. First, lower visual acuity in peripheral vision can be compensated for by increasing stimulus size (Aubert & Foerster, 1857; Proulx, 2010). Second, a large body of work points to the specialization of the medial temporal lobe regions for the encoding of landmark objects for memory and navigation (Janzen, 2006; Li & Bonner, 2022; van Ekert et al., 2015). It is therefore not surprising that participants in the Search condition were able to perceive large objects and encode them in memory. More so we would expect that to see an interaction effect between object size and proximity to the target for participants who selectively attended to the target. As objects become more distant from the target, the size of the object will become more predictive of whether the participant will draw it from memory. Future studies could also more empirically test the relationship between these two factors.

Of note, a recent study by Gandolfo et al., (2023) found that many photographs that elicited boundary contraction had a deeper depth of field than can be processed by the human eye, while photographs that elicited more boundary extension tended to have a shallow depth of field. Depth of field is the range of distance that an object can be moved within that will not lead to a reduction in the resolution, and images with a deep depth of field may reflect unnaturalistic viewing conditions because they allow viewers to see distant objects in high resolution (Gandolfo et al., 2023; Artal, 2014; Middleton, 1957). While we believe that this study provides a large amount of evidence for what conditions cause boundary extension and contraction effects, we do not believe it contradicts our findings or disproves our claim that transformation effects are in part elicited by the AFVF. For one, in the discussion of their paper, Gandolfo and colleagues raise the idea that we propose here, namely that “photographs with a deep depth of field, shot from a great distance, and from a lateral vantage point typically contain many visible objects … to perceptually encode and remember such scenes requires greater attentional resources. This may lead to a loss of peripheral image content, resulting in boundary contraction…” (pg. 10). While they hypothesize that this idea can be disproven because the number of objects in their stimulus images does not explain more variance than depth of field in boundary contraction scores, we do not believe that the attentional constraints that elicit boundary contraction are caused by needing to encode and remember many objects. Instead, it is the attentional constraints required to process one difficult to distinguish object percept that we believe drives the contraction effect. Second, in a subsequent experiment they manipulated both the depth of field of the scene and the focus of the image in a crossed design. They saw extension for all four representations which they suggested provided evidence that boundary contraction was not a phenomenon elicited by naturalistic viewing conditions. However, while they manipulated the focus of the image across their experiments, they did not manipulate the endogenous attentional selection of the viewer. Indeed, there is no strong evidence that the large objects presented in the foreground of their images were not the endogenous focus of the viewers. These objects, while blurry, were most likely still distinguishable and would not require constrained AFVF to identify. Further, in our experiment participants saw the same stimulus images in both conditions, and the only difference was the endogenous guidance of their attention. Although our stimuli were computer generated, our target objects were designed to be difficult to distinguish without foveation, namely by being small in proportion to the rest of the image. Simply requiring the participants to focus and identify difficult to distinguish objects was enough to drive a greater rate of boundary contraction compared to the participants who were not required to distinguish small and distant objects.

Finally, we believe that the limits of spatial attention outside of the AFVF may drive interesting and seemingly contradictory effects dependent on the size of the AFVF in relation to the visual field. For example, a target object could be very distant and/or small in an image, but also be cut-off by an image border. Images that show the biggest effects of boundary extension in memory are images where the original objects are cut-off by the image border (Gagnier et al., 2013). The viewer is biased to remember the object as whole, which requires them to remember the target as farther away then it originally appeared. In this case, viewers may experience extension effects around a selected object, where they remember more information about the object then originally appeared, but still forget or contract the space around that extended area such that information from the object percept and the image border is compressed in their memory representation.

In summary, we believe that these results set-up an interesting account for how selective attention in visual processing can drive transformation effects. We found that requiring the participants to focus attention in order to identify objects during encoding lead memory drawings showing a higher rate of boundary contraction than found in those of participants not required to constrain attention. We propose that the size of AFVF in relation to the visual field is part of what determines the trend and magnitude of boundary transformation effects in memory, along with the composition of the image and the affordances it reveals. These findings should be weighed In light of other recent and fascinating discoveries on what factors influence transformation effects.

References

Abeln, J., Fresz, L., Amirshahi, S. A., McManus, I. C., Koch, M., Kreysa, H., & Redies, C. (2015). Preference for Well-Balanced Saliency in Details Cropped from Photographs. *Frontiers in Human Neuroscience*, *9*, 704.

Akbas, E., & Eckstein, M. P. (2017). Object detection through search with a foveated visual system. *PLoS Computational Biology*, *13*(10), e1005743.

Artal, P. (2014). Optics of the eye and its impact in vision: a tutorial. *Advances in Optics and Photonics*, *6*(3), 340.

Aubert, H., & Foerster, C. F. R. (n.d.). Beiträge zur Kenntniss des indirecten Sehens.(I). Untersuchungen über den Raumsinn der Retina. *Archiv Für Ophthalmologie*.

Bainbridge, W. A., & Baker, C. I. (2020). Boundaries Extend and Contract in Scene Memory Depending on Image Properties. *Current Biology: CB*, *30*(3), 537–543.e3.

Bainbridge, W. A., Hall, E. H., & Baker, C. I. (2019). Drawings of real-world scenes during free recall reveal detailed object and spatial information in memory. *Nature Communications*, *10*(1), 5.

Bainbridge, W. A., Kwok, W. Y., & Baker, C. I. (2021). Disrupted object-scene semantics boost scene recall but diminish object recall in drawings from memory. *Memory & Cognition*. https://doi.org/10.3758/s13421-021-01180-3

Belopolsky, A. V., & Theeuwes, J. (2010). No capture outside the attentional window. *Vision Research*, *50*(23), 2543–2550.

Boettcher, S. E. P., Draschkow, D., Dienhart, E., & Võ, M. L.-H. (2018). Anchoring visual search in scenes: Assessing the role of anchor objects on eye movements during visual search. *Journal of Vision*, *18*(13), 11.

Bonacci, L. M., Bressler, S., & Kwasa, J. A. C. (2020). Effects of visual scene complexity on neural signatures of spatial attention. *Frontiers in Human Neuroscience*. https://www.frontiersin.org/articles/10.3389/fnhum.2020.00091/full

Bonner, M. F., & Epstein, R. A. (2017). Coding of navigational affordances in the human visual system. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(18), 4793–4798.

Brady, T. F., & Störmer, V. S. (2022). The role of meaning in visual working memory: Real-world objects, but not simple features, benefit from deeper processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *48*(7), 942–958.

Brady, T. F., Störmer, V. S., & Alvarez, G. A. (2016). Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(27), 7459–7464.

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research*, *15*(6), 719–722.

Candel, I., Merckelbach, H., Houben, K., & Vandyck, I. (2004). How Children Remember Neutral and Emotional Pictures: Boundary Extension in Children's Scene Memories. In *The American Journal of Psychology* (Vol. 117, Issue 2, p. 249). https://doi.org/10.2307/4149025

Chadwick, M. J., Mullally, S. L., & Maguire, E. A. (2013). The hippocampus extrapolates beyond the view in scenes: an fMRI study of boundary extension. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *49*(8), 2067–2079.

Chapman, P., Ropar, D., Mitchell, P., & Ackroyd, K. (2005). Understanding boundary extension: Normalization and extension errors in picture memory among adults and boys with and without Asperger's syndrome. *Visual Cognition*, *12*(7), 1265–1290.

Cheng, A., Walther, D. B., Park, S., & Dilks, D. D. (2019). What is a scene? Concavity as an intrinsic property of a scene. *Journal of Vision*. https://jov.arvojournals.org/article.aspx?articleid=2750422

Christianson, S. A. (1984). The relationship between induced emotional arousal and amnesia. *Scandinavian Journal of Psychology*, *25*(2), 147–160.

Christianson, S.-Å., & Loftus, E. F. (1987). Memory for traumatic events. *Applied Cognitive Psychology*, *1*(4), 225–239.

Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current Opinion in Neurobiology*, *17*(2), 177–184.

Daniel, P. M., & Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *The Journal of Physiology*, *159*(2), 203–221.

Fernandes, M. A., Koji, S., Dixon, M. J., & Aquino, J. M. (2011). Changing the focus of attention: The interacting effect of valence and arousal. *Visual Cognition*, *19*(9), 1191–1211.

Gandolfo, M., Nägele, H., & Peelen, M. V. (2023). Predictive Processing of Scene Layout Depends on Naturalistic Depth of Field. *Psychological Science*, 9567976221140341.

Geringswald, F., Porracin, E., & Pollmann, S. (2016). Impairment of visual memory for objects in natural scenes by simulated central scotomata. *Journal of Vision*, *16*(2), 6.

Green, D. M., Wilcock, J. A., & Takarangi, M. K. T. (2019). The role of arousal in boundary judgement errors. *Memory & Cognition*, *47*(5), 968–982.

Greene, M., & Trivedi, D. (2022). Spatial scene memories contain a fixed amount of semantic information. In *PsyArXiv*. https://doi.org/10.31234/osf.io/r5fn9

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. *Perception*, *5*, 89–110.

Hafri, A., Wadhwa, S., & Bonner, M. F. (2022). Perceived Distance Alters Memory for Scene Boundaries. *Psychological Science*, *33*(12), 2040–2058.

Hulleman, J., & Olivers, C. N. L. (2017). The impending demise of the item in visual search. *The Behavioral and Brain Sciences*, *40*, e132.

Intraub, H. (2010). Chapter 6 - Rethinking Scene Perception: A Multisource Model. In *Psychology of Learning and Motivation* (Vol. 52, pp. 231–264). Academic Press.

Intraub, H., Daniels, K. K., Horowitz, T. S., & Wolfe, J. M. (2008). Looking at scenes while searching for numbers: Dividing attention multiplies space. In *Perception & Psychophysics* (Vol. 70, Issue 7, pp. 1337–1349). https://doi.org/10.3758/pp.70.7.1337

Intraub, H., & Richardson, M. (1989). Wide-angle memories of close-up scenes. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *15*(2), 179–187.

Janzen, G. (2006). Memory for object location and route direction in virtual large-scale space. *Quarterly Journal of Experimental Psychology* , *59*(3), 493–508.

Jeurissen, D., Self, M. W., & Roelfsema, P. R. (2016). Serial grouping of 2D-image regions with object-based attention in humans. *eLife*, *5*. https://doi.org/10.7554/eLife.14320

Kanizsa, G., & Gerbino, W. (1982). Amodal completion: Seeing or thinking? In *Organization and representation in perception* (pp. 167–190). Hillsdale.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271–304.

Kim, M. S., & Cave, K. R. (2001). Perceptual grouping via spatial selection in a focused-attention task. *Vision Research*, *41*(5), 611–624.

Kirsch, W., Heitling, B., & Kunde, W. (2018). Changes in the size of attentional focus modulate the apparent object's size. *Vision Research*, *153*, 82–90.

Kong, J.-G., Kim, G., & Yi, D.-J. (2010). Effects of scene inversion on boundary extension. In *Journal of Vision* (Vol. 9, Issue 8, pp. 952–952). https://doi.org/10.1167/9.8.952

Kourtzi, Z., & Kanwisher, N. (2001). Representation of Perceived Object Shape by the Human Lateral Occipital Complex. In *Science* (Vol. 293, Issue 5534, pp. 1506–1509). https://doi.org/10.1126/science.1061133

Kravitz, D. J., Peng, C. S., & Baker, C. I. (2011). Real-world scene representations in high-level visual cortex: it's the spaces more than the places. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(20), 7322–7333.

Li, D. S. P., & Bonner, M. F. (2022). Emergent selectivity for scenes, object properties, and contour statistics in feedforward models of scene-preferring cortex. In *bioRxiv* (p. 2021.09.24.461733). https://doi.org/10.1101/2021.09.24.461733

Lin, F., Hafri, A., & Bonner, M. F. (2022). Scene memories are biased toward high-probability views. *Journal of Experimental Psychology. Human Perception and Performance*, *48*(10), 1116–1129.

Liverence, B., & Scholl, B. (2011). Sustained selective attention warps perceived space: Parallel and opposing effects on attended and inhibited objects. *Journal of Vision*, *11*(11), 223–223.

Locher, P., Gray, S., & Nodine, C. (1996). The Structural Framework of Pictorial Balance. *Perception*, *25*(12), 1419–1436.

Low, F. N. (1951). Peripheral visual acuity. *A.M.A. Archives of Ophthalmology*, *45*(1), 80–99.

Mack, A., & Rock, I. (1998). Inattentional blindness: Perception without attention. *Visual Attention*, *8*, 55–76.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology* , *64*(1), 17–24.

Masuda, N. (2015). Effects of emotional valence (positive or negative visual images) and arousal levels (high or low arousal levels) on the useful field of view. *Psychology* , *6*(4), 478–483.

Mathews, A., & Mackintosh, B. (2004). Take a closer look: emotion modifies the boundary extension effect. *Emotion* , *4*(1), 36–45.

McDunn, B. A., Siddiqui, A. P., & Brown, J. M. (2014). Seeking the boundary of boundary extension. *Psychonomic Bulletin & Review*, *21*(2), 370–375.

Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., & Tootell, R. B. (1999). The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *19*(19), 8560–8572.

Middleton, W. E. K. (1957). Vision through the Atmosphere. In J. Bartels (Ed.), *Geophysik II / Geophysics II* (pp. 254–287). Springer Berlin Heidelberg.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*(4715), 782–784.

Munger, M. P., & Multhaup, K. S. (2016). No imagination effect on boundary extension. *Memory & Cognition*, *44*(1), 73–88.

Nelson, W. W., & Loftus, G. R. (1980). The functional visual field during picture viewing. *Journal of Experimental Psychology. Human Learning and Memory*, *6*(4), 391–399.

Park, J., Josephs, E. L., & Konkle, T. (2021). *Systematic transition from boundary extension to contraction along an object-to-scene continuum*. https://doi.org/10.31234/osf.io/84exs

Park, S., Intraub, H., Yi, D.-J., Widders, D., & Chun, M. M. (2007b). Beyond the edges of a view: boundary extension in human scene-selective visual cortex. *Neuron*, *54*(2), 335–342.

Patel, S. D., Esteves, C. V., So, M., Dalgleish, T., & Hitchcock, C. (2022). More than meets the eye: emotional stimuli enhance boundary extension effects for both depressed and never-depressed individuals. *Cognition & Emotion*, 1–9.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1), 8–13.

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

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

Pramod, R. T., Katti, H., & Arun, S. P. (2022). Human peripheral blur is optimal for object recognition. *Vision Research*, *200*, 108083.

Proulx, M. J. (2010). Size matters: large objects capture attention in visual search. *PloS One*, *5*(12), e15293.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To See or not to See: The Need for Attention to Perceive Changes in Scenes. *Psychological Science*, *8*(5), 368–373.

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

Rosenholtz, R. (2016). Capabilities and Limitations of Peripheral Vision. *Annual Review of Vision Science*, *2*, 437–457.

Rosenholtz, R. (2020). Demystifying visual awareness: Peripheral encoding plus limited decision complexity resolve the paradox of rich visual experience and curious perceptual failures. In *Attention, Perception, & Psychophysics* (Vol. 82, Issue 3, pp. 901–925). https://doi.org/10.3758/s13414-019-01968-1

Safer, M. A., Levine, L. J., & Drapalski, A. L. (2002). Distortion in Memory for Emotions: The Contributions of Personality and Post-Event Knowledge. *Personality & Social Psychology Bulletin*, *28*(11), 1495–1507.

Sanders, A. F. (1970). Some aspects of the selective process in the functional visual field. *Ergonomics*, *13*(1), 101–117.

Schneider, I. K., & Mattes, A. (2022). The effect of spatial distance between objects on categorization level. *Psychonomic Bulletin & Review*, *29*(1), 230–242.

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2009). Object recognition during foveating eye movements. *Vision Research*, *49*(18), 2241–2253.

Seamon, J. G., Schlegel, S. E., Hiester, P. M., Landau, S. M., & Blumenthal, B. F. (2002). Misremembering pictured objects: people of all ages demonstrate the boundary extension illusion. *The American Journal of Psychology*, *115*(2), 151–167.

Shamsi, F., Chen, V., Liu, R., Pergher, V., & Kwon, M. (2021). Functional Field of View Determined by Crowding, Aging, or Glaucoma Under Divided Attention. *Translational Vision Science & Technology*, *10*(14), 14.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends in Cognitive Sciences*, *9*(1), 16–20.

Takarangi, M. K. T., Oulton, J. M., Green, D. M., & Strange, D. (2016). Boundary Restriction for Negative Emotional Images Is an Example of Memory Amplification. *Clinical Psychological Science*, *4*(1), 82–95.

van Ekert, J., Wegman, J., & Janzen, G. (2015). Neurocognitive development of memory for landmarks. *Frontiers in Psychology*, *6*, 224.

von der Malsburg, T. (2015). Saccades: Detection of fixations in eye-tracking data. *R Package Version 0.1-1. URL http://CRAN. R-Project*.

Wan, X., & Simons, D. J. (2004). Examining boundary extension in recognition memory for a large set of digitally edited images. In *Journal of Vision* (Vol. 4, Issue 8, pp. 872–872). https://doi.org/10.1167/4.8.872

Wei, L., Zhang, X., Li, Z., & Liu, J. (2018). The semantic category-based grouping in the Multiple Identity Tracking task. *Attention, Perception & Psychophysics*, *80*(1), 118–133.

Wilson, H. R., Levi, D., Maffei, L., Rovamo, J., & DeValois, R. (1990). THE PERCEPTION OF FORM. In *Visual Perception* (pp. 231–272). https://doi.org/10.1016/b978-0-12-657675-7.50016-8

Wolfe, J. M., & Utochkin, I. S. (2019). What is a preattentive feature? *Current Opinion in Psychology*.

Wonning, T. S. (1994). *Susceptibility to misleading post-event information: Emotionality or uniqueness*. Butler University.

Wu, C.-C., & Wolfe, J. M. (2022). The Functional Visual Field(s) in simple visual search. *Vision Research*, *190*, 107965.

Yao, J.-G., Gao, X., Yan, H.-M., & Li, C.-Y. (2011). Field of attention for instantaneous object recognition. *PloS One*, *6*(1), e16343.

Young, A. H., & Hulleman, J. (2013). Eye movements reveal how task difficulty moulds visual search. *Journal of Experimental Psychology. Human Perception and Performance*, *39*(1), 168–190.

Yu, X., Johal, S. K., & Geng, J. J. (2022). Visual search guidance uses coarser template information than target-match decisions. *Attention, Perception & Psychophysics*, *84*(5), 1432–1445.