Scene layout priming relies primarily on low-level features rather than scene layout

Anna Shafer-Skelton

Department of Psychology, University of California, San Diego, CA, USA ✉

Timothy F. Brady

Department of Psychology, University of California, San Diego, CA, USA

The ability to perceive and remember the spatial layout of a scene is critical to understanding the visual world, both for navigation and for other complex tasks that depend upon the structure of the current environment. However, surprisingly little work has investigated how and when scene layout information is maintained in memory. One prominent line of work investigating this issue is a scene-priming paradigm (e.g., Sanocki & Epstein, 1997), in which different types of previews are presented to participants shortly before they judge which of two regions of a scene is closer in depth to the viewer. Experiments using this paradigm have been widely cited as evidence that scene layout information is stored across brief delays and have been used to investigate the structure of the representations underlying memory for scene layout. In the present experiments, we better characterize these scene-priming effects. We find that a large amount of visual detail rather than the presence of depth information is necessary for the priming effect; that participants show a preview benefit for a judgment completely unrelated to the scene itself; and that preview benefits are susceptible to masking and quickly decay. Together, these results suggest that "scene priming" effects do not isolate scene layout information in memory, and that they may arise from low-level visual information held in sensory memory. This broadens the range of interpretations of scene priming effects and suggests that other paradigms may need to be developed to selectively investigate how we represent scene layout information in memory.

Introduction

One of the central challenges in understanding our visual experience is understanding what information about the world we hold in visual memory across brief delays and interruptions, like eye movements and blinks. Visual memory is critical for many tasks we perform every day, like visual search and spatial navigation, and given our limited ability to process everything from a single fixation, visual memory is necessary to build up an experience of a coherent and complete visual scene (e.g., Hollingworth, 2004, 2005). Countless studies investigate memory for discrete *objects*, including the capacity limit of visual memory for objects (e.g., Brady, Störmer, & Alvarez, 2016), the format of the representations for objects and how precision and the number of objects held in mind trade off (Zhang & Luck, 2008), and what neural mechanisms are responsible for storing objects in working memory (Serences, 2016).

However, our visual environment is made up both of discrete objects and also of extended surfaces which form a spatial layout, and there is significant evidence that our visual system processes these types of information separately. For example, fMRI studies in humans show evidence for regions of the brain that respond selectively to scenes compared to objects (Epstein, 2005; Epstein & Kanwisher, 1998; Kravitz, Saleem, Baker, & Mishkin, 2011) and which seem to represent features of a scene's spatial layout rather than the objects it contains (Epstein, 2005; Park, Brady, Greene, & Oliva, 2011). In addition, it is possible to recognize briefly presented scenes even without being able to recognize any of the objects in those scenes (Oliva & Torralba, 2001; Schyns & Oliva, 1994), providing evidence of the independence of scene recognition from object recognition. Greene and Oliva (2009) proposed that this ability could arise from the representation of global properties of scenes, such as the "perspective" or "openness" of a scene. Past research has also drawn distinctions between other types of scene information that may be represented; for example, scene meaning (sometimes called "gist"; e.g., if the scene is a beach, a dining room, etc.; Oliva, 2005, Oliva & Torralba, 2006) and the spatial layout of scenes (Epstein, 2005). Finally, evidence suggests that scene

structure, including the spatial layout of a scene, is crucial to guiding our attention during visual search for objects, and may be represented in a global way independent of object processing (e.g., Torralba, Oliva, Castelhano, & Henderson, 2006; Wolfe, Võ, Evans, & Greene, 2011). However, despite this evidence for distinct representations of scenes (separate from those of objects), little work has investigated how scene-specific spatial layout information is maintained across saccades or brief delays, with most work on scene memory focusing on the role of memory for objects within scenes (Hollingworth, 2004, 2005).

One technique used to study memory for natural scenes in general is to test whether a preview of a scene facilitates subsequent processing related to that scene. For example, a preview of a real-world scene image facilitates subsequent visual search for an object present in that scene (Castelhano & Henderson, 2007; Võ & Henderson, 2010). Whereas there is evidence that the memory representations retained in these studies are abstracted from the exact visual features (e.g., Castelhano & Henderson, 2007 show size invariance), these studies do not make it clear what specifically about the scene is remembered across the delay or to what extent this memory reflects the spatial layout per se as opposed to hypotheses about particular objects and their locations. Work by Sanocki and colleagues has asked more directly about the extent to which the spatial layout of a scene is held in memory by examining the conditions under which a preview of a scene facilitates a depth judgment within that scene (e.g., Sanocki, 2003, 2013; Sanocki & Epstein, 1997; Sanocki, Michelet, Sellers, & Reynolds, 2006). Deciding which of two things is closer in depth specifically targets scene layout representation as it requires participants to have processed and held in mind information about which parts of a scene are near or far from the observer, as opposed to only having held in mind a distribution of possible locations of objects. This "scene-priming" paradigm is widely cited as an example of scene layout information being maintained in memory (e.g., by Chun & Jiang, 1998; Oliva & Torralba, 2001). However, while existing experiments show that the effect persists when some low-level information is varied (e.g., Sanocki, 2003), the effect is often diminished, and it remains possible that low-level visual information (e.g., patterns of orientation across the image; e.g., Brady, Shafer-Skelton, & Alvarez, 2017) could be driving the effect without an abstract representation of the spatial layout of a scene.

In the present experiments, we sought to better characterize the robustness and content of the memory representations responsible for scene-priming effects. In particular, we ask (a) whether scene-priming paradigms are able to isolate the effects of scene layout information held in memory, and (b) whether scene-

priming effects are primarily driven by information held in maskable memory stores, such as iconic memory, or more robust memory stores, such as visual working memory. In our first experiment, we reasoned that if "scene-priming" benefits reflect memory for scene layout, we would expect them to persist when scene previews contain layout information (boundaries of major surfaces or large objects), even if these previews have no identifiable objects and little extraneous visual detail. However, in Experiment 1 we find that whereas previews consisting of full photographs of target scenes are able to speed depth judgments on the target scenes, sparse line drawings of the scenes, which contain only the boundaries of major surfaces or objects and lack semantic information, are unable to speed depth judgments despite containing significant depth information. In Experiment 2 we find that even in a task that doesn't require the usage of the scene at all—and particularly not its layout—photo preview benefits are still present, suggesting they are not a selective index of scene layout or even scene processing. In Experiment 3, we test whether scene-priming benefits are due to a memory store robust to visual masking (e.g., working memory). We find a preview effect for the more detailed line drawings used by Sanocki and Epstein (1997), which contain identifiable shapes as well as extra visual detail, and we find that the effect is abolished with a mask and a longer delay. This suggests that even line drawing preview benefits may be due to a maskable memory store, such as iconic memory. Compared to previous interpretations, these results broaden the possibilities for how the preview is speeding participants' judgments—arguing that low-level information held in iconic memory may be sufficient to facilitate the detection of sudden onsets of the target shapes rather than giving participants a head start on processing scene layout.

Experiment 1: Preview benefit for photos but not sparse line drawings

In a first experiment we tested whether participants were faster at making a depth judgment (i.e., which of two regions of a scene would be closer in depth) when they first saw a preview of either a photograph of the scene or a line drawing of the scene, as compared to an uninformative rectangle presented with the same timing as the two scene-specific previews. The main task for participants was to judge which of two red dots on a scene was on the position in the scene that was closer in depth to the viewer (Figure 1; see Sanocki, 2003). Just before each scene was presented, participants saw one of the preview images. Because line drawings share minimal low-level visual features with the target
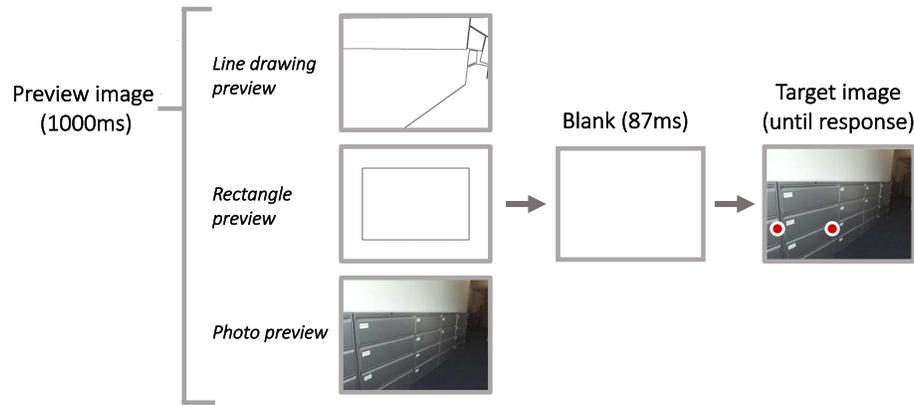
Figure 1. Trial timing and conditions for Experiment 1. Each trial started with a preview image from one of the three preview conditions—a photo preview without the red probe dots present, a rectangle preview, or a line drawing preview that contained information about the spatial layout of the scene but not about the identity of individual objects. As in previous work, these previews were visible for 1,000 ms. After an 87 ms blank, a target image was then presented, and participants were instructed to respond which of the locations cued by the two red probe dots would be closer to the viewer in depth in real life. (Red dots enlarged here for visibility.) In Experiment 1, preview conditions were intermixed, and participants were given no special instructions regarding the previews.

images, a line drawing preview benefit might indicate that scene-priming effects are due to abstract information stored in memory about the spatial layout of the surfaces in the scene. To best assess this idea, the line drawings we selected for this experiment contained the boundaries of the major surfaces and objects in a scene but were screened to ensure they contained no recognizable objects. Because they were automatically generated from the boundaries dividing labeled regions of a scene, they also did not contain extraneous visual detail (e.g., blades of grass, artistic details).

Method

The design, number of participants, and analysis plan for this experiment were preregistered (URL for this experiment: https://aspredicted.org/yw5bg.pdf; also see Appendix Figures A9–A15 for all preregistrations).

Participants

To complete a full counterbalance (see Design and procedure for details), we had 102 participants (6 groups of 17 each). Participants were Mechanical Turk workers who participated in exchange for monetary compensation. Previous literature finds that Mechanical Turk workers are representative of the adult American population (Berinsky, Huber, & Lenz, 2012; Buhrmester, Kwang, & Gosling, 2011) and provide similar data to participants run in laboratory visual cognition studies (Brady & Alvarez, 2011). We recorded timing information in order to ensure

consistency across individual participants' computers and monitors.

Stimuli

Fifty-four images of indoor scenes were selected from the SUNRGB-D database (Song, Lichtenberg, & Xiao, 2015), which includes RGB images of scenes as well as corresponding semantic segmentations and maps of ground-truth depth. Because we didn't want participants to be able to use the vertical position of the target dots as a depth cue, the two target dots placed on each image always had the same vertical position and different horizontal position in the image. Left-right depth-asymmetric scenes ensured a wider variety of possible target dot locations. Thus, to select the scenes to use as target images, we first ordered the images by asymmetry in the mean depth between the left and right halves of the image. Starting with the most depth-asymmetric scenes, line drawings were then created in MATLAB (MathWorks, Natick, MA) by tracing the borders of the semantic segmentations of these same images, and the first ~500 line drawings were screened for identifiable objects, as we wished our line drawing preview images to contain information about spatial layout but not about the identity of particular objects. Participants were asked to list any objects they could identify in the images (excluding major surfaces, like "wall" or "floor"), and an image was selected for the main experiments if neither author AS nor any of 10 pilot participants per image reported being able to identify any objects. This step resulted in 54 images. One set of probe locations was chosen for each image, and target images were created by using MATLAB to

add red dots with white outlines at the chosen probe locations. MATLAB was also used to create the rectangle preview. Scene photograph previews were the original scene images used to create target images. All images were cropped and down-sized, if necessary, to 561 × 427 pixels.

Design and procedure

Participants' task on every trial was to judge which of two red probe dots was on the part of the scene image that would be closer to the viewer in depth in real life. Each trial began with a preview from one of three conditions: (a) a line drawing of the scene photo (line drawing preview); (b) the black outline of a rectangle (rectangle preview), as used in Sanocki and Epstein (1997); and (c) the exact same scene photograph that was used to create the target image (photo preview). Each preview image was presented for one second. Following a brief blank (87 ms, as in Sanocki & Epstein, 1997), the target image was presented until participants responded (see Figure 1 for a schematic of a trial). Participants were instructed to respond as quickly as possible while still getting most trials correct, and feedback was given for incorrect answers.

Each image appeared once in each of the three conditions. The order images appeared was randomized with the constraint that each target image was presented for the first time before any images were presented for the second time. Six possible counterbalance conditions ensured that across all participants, each image appeared equally often in each of the six possible orders of preview conditions (e.g., line drawing, then photo, then rectangle, etc.).

Analyses

Our exclusion criteria and analyses were decided in advance (see preregistration). We excluded individual trials if reaction times were faster than 150 ms and only included correct trials in reaction time analyses. Participants were excluded and replaced with a new participant from the same counterbalance condition if any of the following applied: overall accuracy more than 3 SD below the mean accuracy; overall accuracy below 55%; same response key used on more than 80% of trials; median RT slower than 2 s for any of the three preview conditions; or fewer than 50% of trials included in the main analysis, either because of RTs below 150 ms, or because of incorrect responses. These criteria resulted in the exclusion of 15 participants (14 participants for accuracy, one of whom also had too many RTs faster than 150 ms and another of whom also had median RTs slower than 2 s; as well as one participant for having median RTs slower than 2 s).
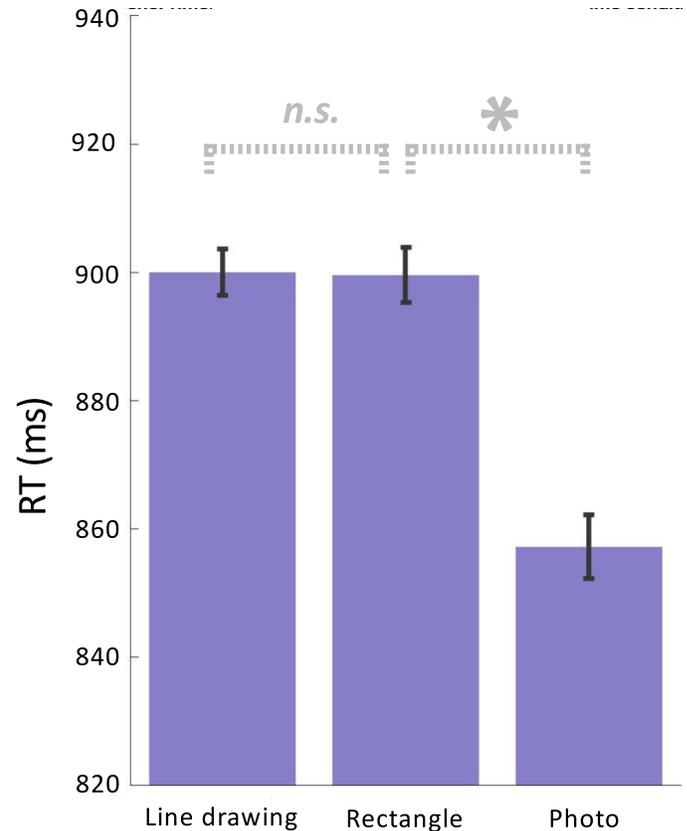

Figure 2. Participants' reaction times in each preview condition in Experiment 1. Bars represent means over participants. Error bars are within-participant SEM. N = 102.

In all experiments, our statistics were performed based on each participant's median reaction time in each of the three preview conditions. The critical analyses were two t tests between participants' median RTs in the photo preview condition and the rectangle preview condition, and between the line drawing preview condition and the rectangle preview condition. Effect sizes were calculated using Cohen's d.

Results

Participants were faster with photo previews, $M = 857$ ms, than with rectangle previews, $M = 900$ ms; $t(101) = 4.91$, $p < 0.001$, $d = 0.49$, indicating that participants were making use of the previews. However, we did not see facilitation for the line drawing preview condition, $M = 900$ ms) compared to the rectangle preview condition, $M = 900$ ms; $t(101) = -0.07$, $p = 0.94$, $d = -0.06$. The photo preview benefit was also significantly larger than the line drawing preview benefit, $t(101) = 5.64$, $p < 0.001$, $d = 0.56$. See Figure 2 for full pattern of results and Figure A1 for single-subject data points.
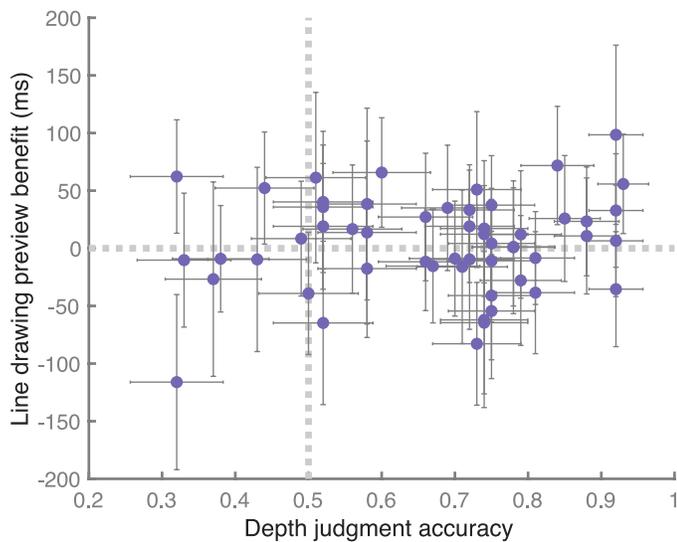
Figure 3. For each image, proportion of participants who correctly made the depth judgment in Experiment A1, plotted against the size of the line drawing preview benefit for that image in Experiment 1. Error bars on depth judgment accuracy are standard error of the proportion, and error bars on the line drawing preview benefit are *SEM*. Gray dotted lines indicate a line drawing preview benefit of 0 (horizontal) and chance performance on the depth judgment task (vertical).

Because we designed the task to have as many usable trials as possible for the reaction time analysis, mean accuracies were high and within a 0.7% range (line drawing: 97.5%; rectangle: 97.0%; photo: 96.8%). Uncorrected posthoc t tests showed one significant accuracy difference (line drawing vs. photo) and small effect sizes in each comparison: rectangle versus photo, $t(101) = 0.49$, $p = 0.62$, $d = 0.05$; line drawing versus rectangle, $t(101) = -1.92$, $p = 0.06$, $d = -0.19$; line drawing versus photo: $t(101) = -2.32$, $p = 0.02$, $d = -0.23$. Because there are no large accuracy differences, speed-accuracy tradeoffs are unlikely to have affected our pattern of RT data. See Appendix Figures A4 through A6 for accuracy data, including individual subject accuracies.

To verify that our line drawings contained information about the spatial structure of each scene, we performed a supplemental experiment (see Experiment A1), in which the red target dot locations were placed directly on the line drawings, and participants judged which regions of the line drawings would be closer in real life. Participants saw the line drawings for the same timing as they saw them during the preview in Experiment 1 (1,000 ms). Participants were 67% accurate at this task, significantly above chance, $t(99) = 17.46$, $p < 0.001$, $d = 1.75$, and in a posthoc analysis, when we reanalyzed Experiment 1 using only the line drawings with significantly above-chance performance (lowest: 66%; mean: 78%), we again did not find a line drawing preview benefit, $t(101) = 0.21$, $p = 0.83$, $d =$

0.02. Again, the photograph preview benefit and the interaction between the line drawing and photograph preview benefits were both significant: photo preview benefit, $t(101) = 4.98$, $p < 0.001$, $d = 0.49$; interaction, $t(101) = 5.17$, $p < 0.001$, $d = 0.51$. In order to further explore the relationship between depth information in the sparse line drawing previews and the line drawing preview benefit, we also plotted the size of the line drawing preview benefit for each image against the proportion of participants who correctly judged depth in that image. If our lack of a preview benefit were due to lack of depth information in the previews, we would expect a positive relationship between depth judgment accuracy and line drawing preview benefits. Instead, we find no evidence of a relationship ($r = 0.13$, $p = 0.35$; see Figure 3 for plot). See Figure A8 for line drawing stimuli ordered by the percent of participants who correctly performed the depth judgment.

Discussion

We found that while previews of the full photograph provided a significant benefit in a subsequent depth judgment task, sparse line drawing previews did not provide a benefit (relative to uninformative rectangle previews). This was true despite the presence of significant depth information in the line drawing previews and held even when we limited our analysis to only those line drawings that provided the best depth information.

In additional experiments reported in the Appendix, we replicated the photograph preview benefit (Experiments A1 through A3) and the lack of a line drawing benefit (Experiments A1 through A2; no line drawings were included in Experiment A3). These replications were originally designed to address the role of mirroring the photo or line drawing preview to distinguish representations of spatial layout from more global scene properties. In all experiments conducted using our sparse line drawing stimuli, we found the same pattern of results: a significant preview benefit for the photo previews, but none for the sparse line drawings in any of the three experiments in which they were included. This was despite the fact that these line drawings contain enough information for participants to make depth judgments.

Thus, despite the presence of depth information in our sparse line drawings, they did not lead to a preview benefit. Previous work (e.g., Sanocki & Epstein, 1997) has found reliable preview benefits from a different set of line drawings, an effect we successfully replicate in Experiment 3. There are two important differences between these stimulus sets. First, whereas Sanocki and Epstein's original (1997) drawings contained semantic information, we specifically chose line drawings that did not contain identifiable objects. This was because we wanted to be able to differentiate between effects

due to the presence of semantic information vs. the presence of spatial layout. The second difference is that the original line drawings share much more local orientation information with the target images (e.g., from blades of grass, small and medium-sized objects) than the sparse line drawings used in Experiment 1. Critically, Experiment 3 of Sanocki and Epstein (1997) does show a scene-priming benefit for artificially generated stimuli that lack semantic information (as our line drawings do) but also share much of the same local orientation information with the target images (which our line drawings do not). This difference led us to believe that the lack of a line drawing benefit in Experiment 1 was not due to the lack of semantic information or participants' inability to categorize our line drawings—instead, one important possibility to consider was whether the amount of visual detail (e.g., orientation information) shared between the previews and targets is critical to finding a line drawing preview effect, and that such a preview effect might not result from processing of scene layout.

Given the very brief delay in our experiment (87 ms, based on previous scene-priming paradigms), it is possible that low-level visual information about the preview image may be stored in a high-capacity visual memory store, such as iconic memory, and that a preview image that is sufficiently similar to the target image (simply missing the probe dots) might allow participants to find the probe dots more efficiently. In other words, rather than giving participants a head-start on layout processing, it is also possible that when more visual detail is shared between the preview image and the target image, the sudden onset of the probe dots becomes more salient, speeding participants' judgments by speeding their detection of the probe dots (e.g., Jonides & Yantis, 1988; Theeuwes, 1991). To address this theory, we conducted two further experiments. Experiment 2 tests whether the photo preview benefit remains for a task in which participants' judgments on the target image should not be sped by knowledge of scene layout, as the target scene is irrelevant to the task, but could be sped by faster detection of the probe dots. Experiment 3 tests whether previews with more detailed line drawings facilitate depth judgments and tests how robust this is to longer delays and visual masking.

Experiment 2: Photo preview benefit even when layout information is irrelevant

The sudden onset of an object tends to draw attention (Jonides & Yantis, 1988; Theeuwes, 1991),

and thus the appearance of probe dots may draw attention even when the preview scene is in iconic memory rather than present on the screen. For example, empty-cell localization tasks and other related tasks show evidence for integration—and detection of new information—across brief delays (Di Lollo, 1980; Eriksen & Collins, 1967).

In particular, evidence suggests that if the delay between two stimuli is less than 80–100 ms, visual persistence of the first overlaps with the initial sensory processing of the second, allowing participants to perceptually combine the two stimuli (Di Lollo, 1980; Eriksen & Collins, 1967), as in the case of two sets of dots forming a letter string (Eriksen & Collins, 1967). Even at slightly longer delays, participants may be able to use informational persistence in iconic memory to notice the sudden onset of the probe dots (e.g., Hollingworth, Hyun, & Zhang, 2005). Thus, given the short delay used in typical scene-priming experiments, it may be that much of the scene-priming benefit arises as a result of faster detection of the probe items following the informative previews rather than faster processing of the target scene.

If preview benefits for more visually detailed preview images are driven by something other than scene layout information (e.g., speedier detection of the probe dots when more visual detail is shared between the preview and target images), we should find a preview benefit for a task that does not require scene layout information at all, or even the use of the target scene at all.

Thus, in Experiment 2, we used the same scene images and target shape locations as in Experiment 1, but rather than seeing two red circles and making a depth judgment about the scene regions underlying these two circles, participants saw a red square and a red diamond and judged whether the left or right of these two target shapes was a square—a judgment for which the background scene was completely irrelevant. If participants' responses in scene-priming experiments like Experiment 1 were speeded due to ease in locating the target shapes, we should also find a photo preview benefit here. On the other hand, if the scene-priming paradigm effectively isolates a head-start in processing layout information, we should not expect a photo preview benefit, since layout information and scene information in general is not informative for this task.

Method

The design, set size, and analysis plan for this experiment were preregistered (https://aspredicted.org/8g5v2.pdf; also see Appendix Figures A9–A15 for all preregistrations).
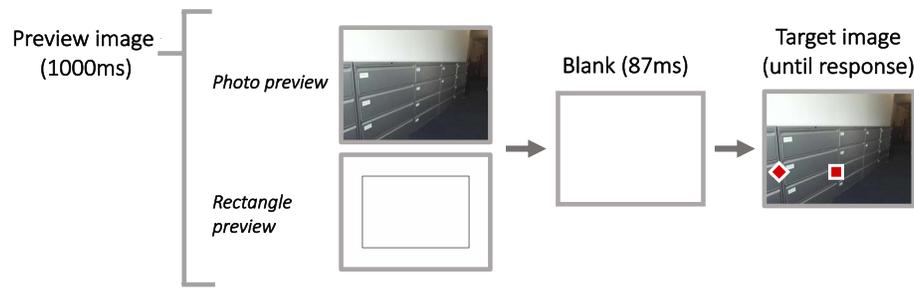
Figure 4. Trial timing and conditions for Experiment 2. As in Experiment 1, a preview image appeared for 1,000 ms, followed by an 87 ms blank. In this experiment, each preview image was either the photo preview (without the square/diamond) or an uninformative rectangle preview. After the delay, a target image was presented, and participants were instructed to indicate which of the two shapes (left or right) was a square. Square and diamond enlarged here for visibility.

Participants

Participants were 100 Mechanical Turk workers (25 in each of 4 counterbalance conditions) who participated in exchange for monetary compensation. No participants participated in the previous experiment.

Stimuli

Stimuli were the same as Experiment 1, except (a) we did not include a line drawing condition, since we did not find a line drawing preview benefit in Experiment 1, and (b) we replaced each set of the target dots with a square and a diamond.

Design and procedure

See Figure 4 for example trial. The design of this experiment was the same as for Experiment 1, except that there was no line drawing preview condition. This resulted in four counterbalance groups, since each target image was repeated with the opposite placement of squares and diamonds across groups, and each variation of each target image was presented either in the rectangle condition first or in the photo condition first across groups. Rectangle and photo previews were intermixed.

Participants' task was to judge whether the square was the left of the two shapes or the right of the two shapes.

Analyses

Analyses were the same as in Experiment 1, and exclusion criteria were the same as in the other two experiments. The preregistered exclusion criteria resulted in the exclusion of one participant for having an overall accuracy lower than 3 SD below the mean accuracy. This participant was replaced with a participant from the same counterbalance condition.

Results and discussion

Participants were significantly faster in the photo preview condition, $M = 777$ ms, compared to the rectangle preview condition, $M = 814$ ms; $t(99) = 4.36$, $p < 0.001$, $d = 0.44$ (see Figure 5), indicating the presence of photo "scene-priming" effects even for a task that does not require any scene layout information or any use of the background scene in the task. Accuracies in the two conditions were high and very similar (rectangle: 98.5%; photo: 98.6%), and a posthoc uncorrected t test showed no significant difference between them, $t(99) = 0.39$, $p = 0.70$, $d = 0.04$. See Figure A2 for single-subject data points.

Because square versus diamond targets are randomly assigned to either target location (with this assignment counterbalanced across participants for each image), the effects here cannot be the result of layout information being predictive of the locations of squares versus diamonds, or of the visual features of these targets, or of the response participants need to make. Instead, the
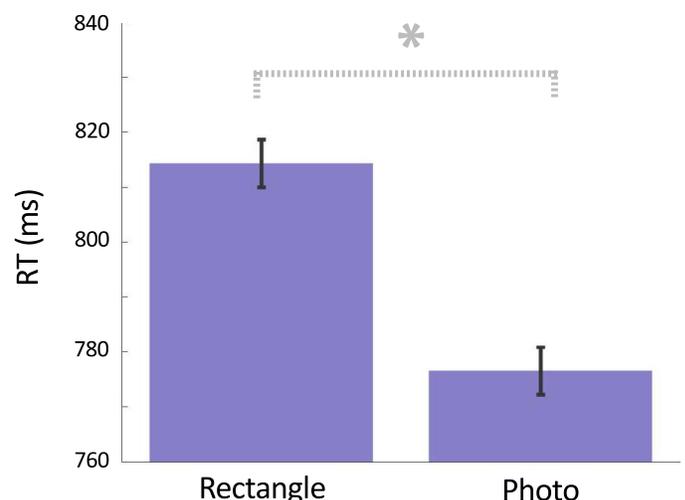

Figure 5. Means of reaction times in each preview condition in Experiment 2. Error bars are within-participant SEM. $N = 100$.

results support the hypothesis that scene-priming with photograph previews can result from participants being faster to localize the probes; in other words, that response times are facilitated by the sudden onsets of the probe shapes when detailed visual information is shared by the preview image and the target image. Because the preview images do contain layout information, we cannot rule out the hypothesis that participants obligatorily process this information. However, because of the absence of a relationship between the layout of the scene and the shape task, there is no plausible explanation for how faster processing of the background scene's layout could speed shape judgments. Further, overall faster reaction times in this experiment compared to Experiment 1 are consistent with the task in the current experiment not requiring any processing of the background scenes. (By contrast, in Experiment 1, once the dots were localized in each condition, a depth task also needed to be performed.)

This hypothesis that the source of scene-priming effects may be the detection of the onset of the target shapes provides a potential explanation for the lack of scene priming in the line drawings we used in Experiment 1. That is, whereas the sparse line drawings contained significant depth information, they were more abstract and considerably less visually detailed than Sanocki and Epstein's (1997) line drawings, causing them to share less low-level visual information with the target images. Thus, it may be that this lack of visual detail prevented participants from detecting the onset of the probes efficiently. To examine this hypothesis, we next sought to test the source of the scene-priming effects found using Sanocki and Epstein's (1997) original stimuli, and in particular the robustness of these effects to visual masking and increased delay, both of which should severely curtail participants' ability to quickly detect the onset of the probes if such detection relies on iconic memory (Irwin & Thomas, 2008).

Experiment 3: Replication using original Sanocki and Epstein stimuli; effects abolished using 200 ms masked delay period

In Experiment 3, we asked what type of memory store drives scene-priming effects. Since these effects may be dependent on the amount of visual detail present shared between preview images and target images, and appear to occur even when the background scene is irrelevant, this condition raises the possibility that they could arise from integration between the preview and the target scene and the improved ability of participants to detect the probes that results from this integration. Thus we hypothesized that they may

be driven not by a robust working memory representation but by a high-capacity but fragile visual memory like iconic memory.

A classical distinction in visual memory is between iconic memory and visual working memory, with high-capacity sensory memory ("iconic" memory) decaying quickly and being easily disrupted by masks, and visual short-term memory being relatively robust to longer delays and visual masks (Irwin & Thomas, 2008). Thus, we reasoned that if the benefits of detailed line drawing previews and photograph previews arose from integration between the preview scene and the target scene in iconic memory, this memory should be interrupted by a visual mask and/or by a longer delay period, even if this delay period remains quite short. By contrast, if the preview benefit reflects a head-start in scene layout processing or participants' ability to hold scene layout in working memory, the preview benefit should remain even after a brief visual mask and a 200 ms delay.

Thus, using Sanocki and Epstein's original (1997) stimuli and timing, we first replicated both the photo preview benefit and the line drawing preview benefit. Critically, we included two delay period conditions: an unmasked delay period of the same duration as the original experiments (87 ms) and a masked delay period of 200 ms. If Sanocki and Epstein's scene-priming effects were driven by information held in iconic memory, the mask and the longer delay between the preview and target image should abolish the preview benefits. On the other hand, if scene-priming effects are driven by information in a more robust form of visual memory, such as visual working memory, the scene-priming benefits should remain.

Method

The design, set size, and analysis plan for this experiment were preregistered (preregistration for this experiment here: https://aspredicted.org/rk6f6.pdf; also see Appendix Figures A9–A15 for all preregistrations).

Participants

Participants were 306 Mechanical Turk workers (102 in each counterbalance condition) who participated in exchange for monetary compensation. We sought (and preregistered) greater power in this experiment as we were predicting a smaller or absent effect of scene previews in the masked conditions.

Stimuli

Stimuli were the original Sanocki and Epstein (1997) target images, scene photographs, and line drawings. The rectangle preview was created in MATLAB. In
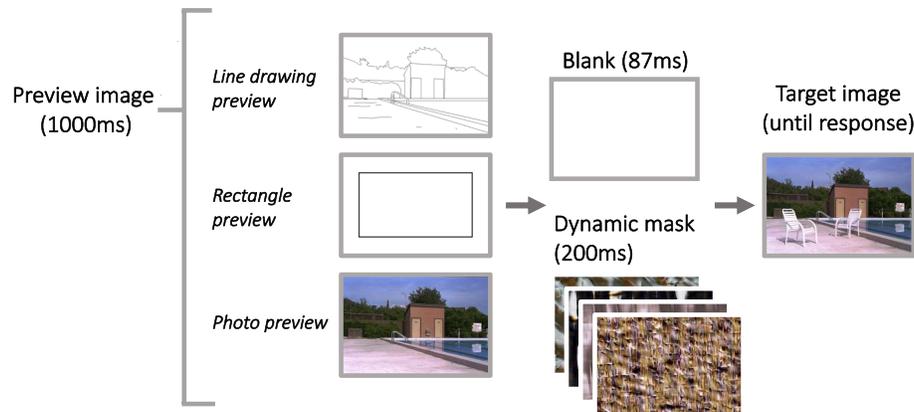
Figure 6. Trial timing and conditions for Experiment 3. The line drawing and photo previews do not have the chairs present that are present in each of the target images, and the judgment required on the target image is which of two chairs would be closer to the viewer in depth in real life. In the task, first, a preview image appeared for 1,000 ms. It was either followed by an 87 ms blank, as in the first two experiments (and as in Sanocki & Epstein, 1997), or a dynamic visual mask, for 200 ms. Preview and target images were the same as in Sanocki and Epstein (1997).

addition to these three preview conditions, which we focus on here, the experiment also contained mirrored line drawing previews, as our original interest was to examine the role of spatial layout versus more global scene properties in scene priming (see also Experiment 1 replications in the Appendix). In this experiment, we do not focus on the mirrored line drawing condition because in this particular set of stimuli the images are extremely symmetrical (with only the exception of the pool image), and thus there is no real difference in the informativeness of the original line drawings and the mirrored line drawings (see Appendix Figure A7).

Design and procedure

See Figure 6 for example trial. Preview conditions were blocked, with the order of blocks counterbalanced across participant groups using a balanced Latin square. In this experiment, following Sanocki and Epstein (1997), participants' task was to judge which of two chairs was closer in depth to the viewer (rather than the red dots in the previous experiments).

Analyses

Analyses and exclusion criteria were the same as for Experiment 1, except that we were now also interested in how any line drawing or photo preview benefits changed according to mask condition. Our preregistered exclusion criteria resulted in the exclusion of 17 participants (15 for accuracy, one of whom also had too many trials faster than 150 ms; there were also two participants with median RTs slower than 2 s in at least one condition).

Results and discussion

In the unmasked condition, we found benefits for both line drawings, $M = 807$ ms, and photographs, $M = 800$ ms, over rectangle previews, $M = 826$ ms; line drawings versus rectangles, $t(305) = 3.18$, $p < 0.002$ $d = 0.18$; photos versus rectangles: $t(305) = 3.88$, $p < 0.001$, $d = 0.22$ (see Figure 7). However, both effects were abolished in the masked condition, line drawings versus rectangles, $t(305) = -1.26$, $p = 0.21$, $d = -0.07$; photos versus rectangles, $t(305) = -0.56$, $p = 0.57$, $d = -0.03$, with the direction of means for both being in the direction of the preview slowing response, and with Bayes factors showing substantial evidence favoring the
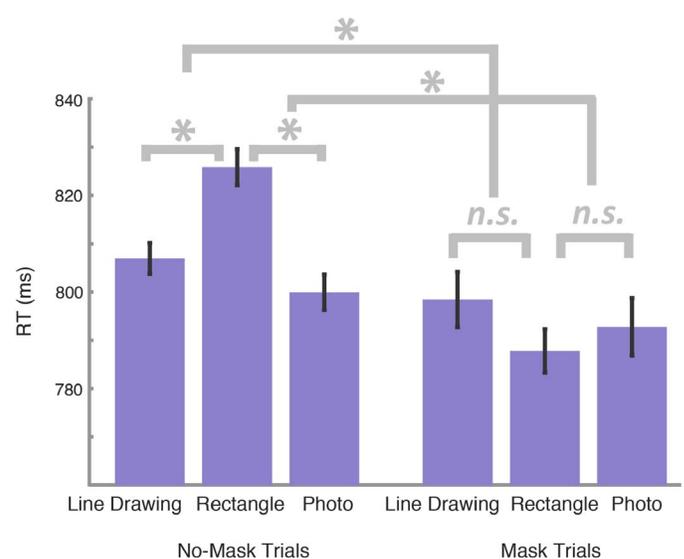

Figure 7. Reaction times in each preview condition in Experiment 3. Bars represent means over all participants. Error bars are within-participant *SEM*. $N = 306$.

null hypothesis in both cases (Scaled JZS Bayes Factor = 7.1 line drawing vs. rectangles; 13.4 photos vs. rectangles; using default of $r = 0.707$ and the method of Rouder, Speckman, Sun, Morey, & Iverson, 2009). A posthoc power analysis suggests that if the preview benefits in the masked condition were of the same effect size as in the unmasked conditions ($\sim d = 0.20$), we had 96.7% power to detect this in the current study with our sample size. Comparing the benefit in the masked versus unmasked conditions, both drawing versus rectangle and photo versus rectangle were significantly smaller in the masked compared to the unmasked conditions: line drawing benefit, $t(305) = 2.97$, $p = 0.003$, $d = 0.17$; photo benefit, $t(305) = 3.02$, $p = 0.003$, $d = 0.17$. Mean accuracies in each combination of mask and preview condition ranged between 98.0% and 98.6%. Posthoc uncorrected t tests showed no significant differences in any pairs of conditions within either mask/delay condition, or for either of the two critical interactions across mask/delay conditions. See Figure A3 for single-subject data points.

Note that in this experiment using the Sanocki and Epstein (1997) stimuli, rather than making a depth judgment on a pair of red circles, participants had to make a depth judgment on two large chairs that appear in the target scene but are not present in the previews. Thus, the raw reaction times are numerically faster than in Experiment 1, likely reflecting easier localization of the larger chair targets compared to the smaller dot targets. The faster overall reaction times in the 200 ms masked condition are consistent with participants benefiting from a longer preparation time compared to the 80 ms no-mask condition. While this possibility does not detract from our main conclusions, it prevents us from making any additional conceptual claims based on the overall RT differences in the 80 ms no-mask condition versus the 200 ms masked condition. Because the reaction times in our study are well within the range reported for previous scene-priming effects (as fast as 562 ms in Sanocki & Epstein, 1997 and as slow as 1,029 ms in Sanocki, 2013), this argues that the lack of scene priming in our masked condition is not due to ceiling effects.

The fact that both effects were abolished by a longer but still short (200 ms) delay and a mask argues that the original preview benefits were due to visual information held in high-capacity sensory memory (e.g., iconic memory). Because a wide variety of information can be stored in iconic memory, including low-level visual information such as patterns of orientation across an image, the results of the present experiment further argue that scene-priming paradigms are not able to isolate the effects of scene layout information stored across a delay period. Instead, these results are also consistent with the interpretation that preview images facilitate participants' search for the probes rather than giving them a head-start on layout processing.

General discussion

In three experiments, we showed that the effects of scene previews on subsequent depth judgments (termed "scene priming"; Sanocki & Epstein, 1997) are (a) present for visually detailed preview images, but not for sparser preview images that still contain depth information; and (b) are driven by information held in iconic memory or another short-term and maskable memory store. In particular, we showed that while both photograph previews (Experiments 1 and 3) and visually detailed line drawings (Experiment 3) produced scene-priming benefits, abstract line drawings (containing only the boundaries of major objects and surfaces; Experiment 1 did not, despite containing significant depth information. This is not what we would expect if scene previews facilitated performance by giving participants a head start on layout processing. Further arguing against the idea that scene previews primarily facilitate layout processing, we found a photograph preview benefit even for a task in which the background scene was completely irrelevant (Experiment 2). Finally, we found that the scene-priming effects from Sanocki and Epstein's original (1997) photographs and detailed line drawings both disappeared when the delay period is masked, suggesting that scene-priming effects are driven by information held in iconic memory. Together, our data suggest that scene previews may primarily speed participants' localization of the probe shapes on the target image.

Relationship to previous scene-priming findings

Our results are in line with previous studies showing benefits of a scene preview on subsequent processing of a scene. For example, a preview of a real-world scene image facilitates subsequent visual search in that scene (Castelhano & Henderson, 2007; Võ & Henderson, 2010), and both scene photograph and detailed line drawing previews speed subsequent depth judgments on scenes (Sanocki & Epstein, 1997). We consistently replicated photograph preview benefits, and we replicated line drawing preview benefits when using the same line drawings as the original experiment (Sanocki & Epstein, 1997).

However, our results are at odds with the argument that these effects are due to abstract visual information about a scene's layout that speeds participants' judgments by giving them a head start on processing scene layout information. Previous support for this argument is based on a few experiments: First, in Sanocki and Epstein (1997), a previewed line drawing of a scene photograph facilitates 3D depth judgments on the photograph. Because the line drawing has less

low-level information in common with the target image than a full photograph preview and facilitates depth judgments, they reasoned that layout information is stored across the delay. Second, Sanocki (2003) showed scene priming with moderate retinal shifts between previews and targets (experiment 5), and Sanocki and Epstein (1997) argue that the viewpoint shifts present in their experiment 4 are evidence of a more abstract, higher-level representation. Finally, Sanocki (2003; experiments 2 through 5) varies lighting direction between preview and target images, disrupting some low-level visual information.

However, while the above experiments show that scene-priming benefits persist when some low-level information is varied, the effect is often diminished, and remaining low-level visual information (e.g., the orientation information present in each part of the image) could be driving the preview benefit. Even line drawing previews, which perhaps share the least pixel-by-pixel information with target photographs, still preserve some of the important orientation information in the target photographs, especially the detailed line drawings used in Sanocki and Epstein (1997). Orientation and edge information is well known to be relevant to scene information. Both local orientations, curvatures, and angles (e.g., Walther & Shen, 2014) and the global distribution of orientation information (e.g., Brady et al., 2017; Oliva & Torralba, 2001) are critical to scene recognition. Furthermore, detailed line drawings elicit remarkably similar brain activity in scene regions to real scene photographs (Walther, Chai, Caddigan, Beck, & Fei-Fei, 2011). Thus, it may be that line drawing preview benefits in fact reflect the preservation of these important low-level or midlevel features of a scene that are necessary for participants to notice the onset of a new set of objects, rather than reflecting the representation of more abstract properties such as spatial layout.

Another study using scene previews (Castelhano & Pollatsek, 2010) shows the limited viewpoint tolerance of scene-priming effects, and it is notable that the viewpoints that give the largest scene-priming benefits are also the ones with the most low-level overlap with the target images. This is in line with the results we report here. Prior work by Gottesman (2011) has the potential to demonstrate the maintenance of more abstract information from scene previews, but the conclusions of that work rest on the particular details of the stimuli they used and how specific the effects of boundary extension are to higher levels of the visual hierarchy. Future work could investigate the potential of their paradigm for specifically investigating scene layout information stored in memory.

Our findings are also consistent with arguments made in Germeys and d'Ydewalle (2001), but whereas their results call into question scene-priming results with significant pixel-by-pixel overlap between preview and target images, ours argue that even studies designed to reflect a more abstract memory store, such as those using line drawings as previews (e.g., see Sanocki 2003), may instead be picking up on the speedier detection of target shapes.

Implications for representations of space in visual memory

There is a long-running and broad debate over how much information we maintain about the world in memory (O'Regan & Noë, 2001), whether and when we are able to integrate information from successive fixations into a more complete picture of our surroundings (Henderson, 1997; Irwin, Yantis, & Jonides, 1983; Irwin, 1991), and what format these representations are in. Investigating the types of scene information retained in memory has the potential to shed light on how much information we maintain in memory about the world and how we combine information across successive fixations to build a more complete picture of our surroundings. Whereas a good deal of work has been done on the maintenance of object information across brief delays and eye movements, less is known about whether scene layout information persists across eye movements, and if so, how this type of memory fits into the process of maintaining a stable representation of the world. The flash-preview-moving-window paradigm (Castelhano & Henderson, 2007; Võ & Henderson, 2010) demonstrates memory for a size-invariant representation of some information about a natural scene, but it is unclear what the content of this representation is. Change blindness effects (Carlson-Radvansky & Irwin, 1995; Franconeri & Simons, 2003; Grimes, 1996; Luck & Vogel, 1997; McConkie & Currie, 1996; Phillips, 1974; Rensink, O'Regan, & Clark, 1997; Simons, 1996) argue that when we are unable to rely on iconic memory (as is often the case in the real world), visual details are often lost. It is an important question the extent to which people store detailed spatial layout information in memory—and particularly working memory, which is quite capacity limited. Because the current findings call into question one of the main literatures used to support the existence of spatial layout representations, it remains an open question the extent of the layout of specific surfaces in a scene (scene layout) that we are capable of maintaining in working memory.

One of the challenges for future work is understanding how scene layout representations can be quantified and incorporated into existing models of working memory. In particular, although working memory is known to be quite capacity limited, there is

significant debate in the visual working memory literature about whether the units of working memory capacity are discrete "slots" or a more continuous resource that can be used to remember fewer objects with more precision or more objects with less precision (Luck & Vogel, 2013; Ma, Husain, & Bays, 2014). Because the layout of a scene is not obviously broken down into discrete objects, it is a challenge to conceptualize how to incorporate it into these primarily object-based models of working memory. Existing models that incorporate both individual objects as well as higher level information like ensemble structure may be adaptable to incorporate other information like scene layout (Brady & Alvarez, 2011; Brady & Tenenbaum, 2013).

Neural models of working memory more easily accommodate the representation of scene layout information. For example, the occipital place area (OPA) and parahippocampal place area (PPA) are generally seen as perceptual areas, but many neural models of working memory are based on the idea that "perceptual" areas can be recruited for working memory storage (Awh & Jonides, 2001; Chelazzi, Miller, Duncan, & Desimone, 1993; Curtis & D'Esposito, 2003; D'Esposito, 2007; D'Esposito & Postle, 2015; Harrison & Tong, 2009; Lara & Wallis, 2015; Magnussen, 2000; Miller, Li, & Desimone, 1993; Pasternak & Greenlee, 2005; Serences, Ester, Vogel, & Awh, 2009; Sreenivasan, Curtis, & D'Esposito, 2014). The neuroimaging literature shows evidence of scene-specific representations in perceptual contexts (Dilks, Julian, Paunov, & Kanwisher, 2013; Epstein & Kanwisher, 1998; Maguire, 2001), including boundary information in the OPA (Julian, Ryan, Hamilton, & Epstein, 2016). Thus, future work could examine working memory delay period activity or patterns in these regions to quantify working memory for spatial layout and examine how it interacts with other working memory capacity limits.

Conclusion

The ability to perceive and remember the spatial layout of a scene is critical to understanding the visual world, both for navigation and for other complex tasks that depend upon the structure of the current environment. The present studies offer a new interpretation of scene-priming effects, which are one of the primary tools used to study the representation of spatial layout. We find that scene-priming effects are driven by visual detail held in iconic memory that does not necessarily isolate scene layout information. Studying scene layout information in memory has the potential to offer fresh insight into several long-standing questions about visual memory, and the current studies are a critical first step towards this goal.

Keywords: scene perception, object recognition, visual memory, spatial layout

Acknowledgments

Corresponding author: Anna Shafer-Skelton.
Email: annashaferskelton@ucsd.edu.
Address: Department of Psychology, University of California, San Diego, CA, USA.

References

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, 5(3), 119–126, https://doi.org/10.1016/S1364-6613(00)01593-X.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. *Political Analysis*, 20(3), 351–368, https://doi.org/10.1093/pan/mpr057.

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science*, 22(3), 384–392, https://doi.org/10.1177/0956797610397956.

Brady, T. F., Shafer-Skelton, A., & Alvarez, G. A. (2017). Global ensemble texture representations are critical to rapid scene perception. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 0–17, https://doi.org/10.1037/xhp0000399.

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, USA*, 113(27), 7459–7464, https://doi.org/10.1073/pnas.1520027113.

Brady, T. F., & Tenenbaum, J. B. (2013). A probabilistic model of visual working memory: Incorporating higher order regularities into

working memory capacity estimates. *Psychological Review, 120*(1), 85–109, https://doi.org/10.1037/a0030779.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk. *Perspectives on Psychological Science, 6*(1), 3–5, https://doi.org/10.1177/1745691610393980.

Carlson-Radvansky, L. A., & Irwin, D. E. (1995). Memory for structural information across eye movements. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*(6), 1441–1458.

Castelhano, M. S., & Henderson, J. M. (2007). Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 33*(4), 753–763, https://doi.org/10.1037/0096-1523.33.4.753.

Castelhano, M. S., & Pollatsek, A. (2010). Extrapolating spatial layout. *Memory & Cognition, 38*(8), 1018–1025, https://doi.org/10.3758/MC.38.8.1018.

Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993, May 27). A neural basis for visual search in inferior temporal cortex. *Nature, 363*(6427), 345–347, https://doi.org/10.1038/363345a0.

Chun, M. M., & Jiang, Y. H. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*(1), 28–71.

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences, 7*(9), 415–423, https://doi.org/10.1016/S1364-6613(03)00197-9.

D'Esposito, M. (2007). From cognitive to neural models of working memory. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 362, pp. 761–772), https://doi.org/10.1098/rstb. 2007.2086.

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology, 66*, 115–142, https://doi.org/10.1146/annurev-psych-010814-015031.

Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General, 109*(1), 75–97. Retrieved from http://wexler.free.fr/library/files/di lollo (1980) temporal integration in visual memory.pdf

Dilks, D. D., Julian, J. B., Paunov, A. M., & Kanwisher, N. (2013). The occipital place area is causally and selectively involved in scene perception. *Journal of Neuroscience, 33*(4): 1331–6a,

https://doi.org/10.1523/JNEUROSCI.4081-12.2013.

Epstein, R. (2005). The cortical basis of visual scene processing. *Visual Cognition, 12*(6), 954–978, https://doi.org/10.1080/13506280444000607.

Epstein, R., & Kanwisher, N. (1998, April 9). A cortical representation of the local visual environment. *Nature, 392*(6676), 598–601, https://doi.org/10.1038/33402.

Eriksen, C. W., & Collins, J. F. (1967). Some temporal characteristics of visual pattern perception. *Journal of Experimental Psychology, 74*(4, Part 1), 476–484, https://doi.org/10.1037/h0024765.

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics, 65*(7), 999–1010. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14674628

Germeys, F., & Ydewalle, G. (2001). Revisiting scene primes for object locations. *Quarterly Journal for Experimental Psychology, 54A*(3), 683–693, https://doi.org/10.1080/0272498004200036.

Gottesman, C. V. (2011). Mental layout extrapolations prime spatial processing of scenes, *37*(2), 382–395, https://doi.org/10.1037/a0021434.

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science: A Journal of the American Psychological Society, 20*, 464–472, https://doi.org/10.1111/j.1467-9280.2009.02316.x.

Harrison, S. A., & Tong, F. (2009, April 2). Decoding reveals the contents of visual working memory in early visual areas. *Nature, 458*(7238), 632–635, https://doi.org/10.1038/nature07832.

Henderson, J. M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science, 8*(1), 51–55, https://doi.org/10.1111/j.1467-9280.1997.tb00543.x.

Hollingworth, A. (2004). Constructing visual representations of natural scenes: The roles of short-and long-term visual memory. *Journal of Experimental Psychology: Human Perception and Performance, 30*(3), 519, https://doi.org/10.1037/0096-1523.30.3.519.

Hollingworth, A. (2005). The relationship between online visual representation of a scene and long-term scene memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 31*(3), 396–411, https://doi.org/10.1037/0278-7393.31.3.396.

Hollingworth, A., Hyun, J.-S., & Zhang, W. (2005). The role of visual short-term memory in empty cell

localization. *Perception & Psychophysics*, *67*(8), 1332–1343, https://doi.org/10.3758/BF03193638.

Irwin, D. E., & Thomas, L. E. (2008). Visual sensory memory. In S. J. Luck & A. Hollingworth (Eds.), *Visual memory* (pp. 9–43). New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195305487.003.0002.

Irwin, D. E., Yantis, S., & Jonides, J. (1983). Evidence against visual integration across saccadic eye movements. *Perception & Psychophysics*, *34*(1), 49–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6634358

Irwin, E. (1991). Information Integration across saccadic eye movements. *Cognitive Psychology*, *23*, 420–456.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, *43*(4), 346–354, https://doi.org/10.3758/BF03208805.

Julian, J. B., Ryan, J., Hamilton, R. H., & Epstein, R. A. (2016). The occipital place area is causally involved in representing environmental boundaries during navigation. *Current Biology*, *26*(8), 1104–1109, https://doi.org/10.1016/j.cub.2016.02.066.

Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews Neuroscience*, *12*(4), 217–230, https://doi.org/10.1038/nrn3008.

Lara, A. H., & Wallis, J. D. (2015). The role of prefrontal cortex in working memory: A mini review. *Frontiers in Systems Neuroscience*, *9*: 173, https://doi.org/10.3389/fnsys.2015.00173.

Luck, S. J., & Vogel, E. K. (1997, November 20). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281, https://doi.org/10.1038/36846.

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400, https://doi.org/10.1016/j.tics.2013.06.006.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356, https://doi.org/10.1038/nn.3655.

Magnussen, S. (2000). Low-level memory processes in vision. *Trends in Neurosciences*, *23*(6), 247–251, https://doi.org/10.1016/S0166-2236(00)01569-1.

Maguire, E. A. (2001). The retrosplenial contribution to human navigation: A review of lesion and neuroimaging findings. *Scandinavian Journal of Psychology*, *42*, 225–238.

McConkie, G. W., & Currie, C. B. (1996). Visual

stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(3), 563–581, https://doi.org/10.1037/0096-1523.22.3.563.

Miller, E. K., Li, L., & Desimone, R. (1993). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *The Journal of NeuroscienceJournal of Neuroscience*, *13*(4), 1460–1478, https://doi.org/10.1016/j.conb.2004.03.013.

O'Regan, J. K., & Noë, a. (2001). A sensorimotor account of vision and visual consciousness. *The Behavioral and Brain Sciences*, *24*(5), 939–973; discussion 973–1031, https://doi.org/10.1017/S0140525X01000115.

Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Encyclopedia of neurobiology of attention* (pp. 251–256). San Diego, CA: Elsevier. Retrieved from https://s3.amazonaws.com/academia.edu.documents/30821187/oliva04.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1533077700&Signature=uo7Thqhg6Pxgx8OSajOnw%2FpOFH0%3D&response-content-disposition=inline%3Bfilename%3DGist_of_the_scene.pdf

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.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, *155*, 23–36.

Park, S., Brady, T. F., Greene, M. R., & Oliva, A. (2011). Disentangling scene content from spatial boundary: Complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. *The Journal of Neuroscience*, *31*(4), 1333–1340, https://doi.org/10.1523/JNEUROSCI.3885-10.2011.

Pasternak, T., & Greenlee, M. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, *6*, 97–107, https://doi.org/10.1038/nrn1637.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics*, *16*(2), 283–290.

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, https://doi.org/10.1111/j.1467-9280.1997.tb00427.x.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic*

Bulletin & Review, *16*(2), 225–237, https://doi.org/10.3758/PBR.16.2.225.

Sanocki, T. (2003). Representation and perception of scenic layout. *Cognitive Psychology*, *47*, 43–86, https://doi.org/10.1016/S0010-0285(03)00002-1.

Sanocki, T. (2013). Facilitatory priming of scene layout depends on experience with the scene. *Psychonomic Bulletin & Review*, *20*(2), 274–281, https://doi.org/10.3758/s13423-012-0332-9.

Sanocki, T., & Epstein, W. (1997). Priming spatial layout of scenes. *Psychological Science*, *8*(5), 374–378.

Sanocki, T., Michelet, K., Sellers, E., & Reynolds, J. (2006). Representations of scene layout can consist of independent, functional pieces, *68*(3), 415–427.

Schyns, P., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, *5*(4), 195–200. Retrieved from http://pss.sagepub.com/content/5/4/195.short

Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. *Vision Research*, *128*, 53–67, https://doi.org/10.1016/j.visres.2016.09.010.

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, *20*(2), 207–214, https://doi.org/10.1111/j.1467-9280.2009.02276.x.

Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science*, *7*(5), 301–305, https://doi.org/10.1111/j.1467-9280.1996.tb00378.x.

Song, S., Lichtenberg, S. P., & Xiao, J. (2015). SUN RGB-D: A RGB-D scene understanding benchmark suite. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (June), 567–576, https://doi.org/10.1109/CVPR.2015.7298655.

Sreenivasan, K. K., Curtis, C. E., & D'Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends in Cognitive Sciences*, *18*(2), 82–89, https://doi.org/10.1016/j.tics.2013.12.001.Revisiting.

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, *49*(1), 83–90. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.457.2361&rep=rep1&type=pdf

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*(4), 766–786, https://doi.org/10.1037/0033-295X.113.4.766.

Võ, M. L.-H., & Henderson, J. M. (2010). The time course of initial scene processing for eye movement guidance in natural scene search. *Journal of Vision*, *10*(3):14, 1–13, https://doi.org/10.1167/10.3.14. [PubMed] [Article]

Walther, D. B., Chai, B., Caddigan, E., Beck, D. M., & Fei-Fei, L. (2011). Simple line drawings suffice for functional MRI decoding of natural scene categories. *Proceedings of the National Academy of Sciences, USA*, *108*(23), 9661–9666, https://doi.org/10.1073/pnas.1015666108.

Walther, D. B., & Shen, D. (2014). Nonaccidental properties underlie human categorization of complex natural scenes. *Psychological Science*, *25*(4), 851–860, https://doi.org/10.1177/0956797613512662.

Wolfe, J. M., Võ, M. L. H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, *15*(2), 77–84, https://doi.org/10.1016/j.tics.2010.12.001.

Zhang, W., & Luck, S. J. (2008, May 8). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235, https://doi.org/10.1038/nature06860.

Appendix

Experiment A1: Verifying sparse line drawings contain layout information

The design, set size, and analysis plan for this experiment were preregistered (see below for pre-registrations).

Participants

Participants were 100 Mechanical Turk workers who participated in exchange for monetary compensation. No participants participated in any other experiments using these line drawings.

Stimuli

Stimuli were the line drawing images used in Experiments 1 and 2, with target dots placed on them in the locations corresponding to the photo target images from Experiments 1 and 2.

Design and procedure

In this experiment, there were no preview images, and participants saw each target line drawing once. During practice, participants were shown examples of line drawings created from photographs, and they practiced choosing which dot would indicate the closer part of the line drawing if the scene existed in three dimensions. Participants were given feedback for correct and incorrect answers in the practice, but only for incorrect answers during the main experiment.

Analyses

In this experiment, we analyzed average performance as well as performed a two-tailed binomial test on each image to determine whether participants' depth judgments were significantly above chance.

Results

Participants were 67% accurate at this task, significantly above chance, $t(99) = 17.46$, $p < 0.001$, $d = 1.75$. We found that 35 of the 54 images had above-chance depth judgments in the binomial test, and these are the images that are the focus of the posthoc analysis in Experiment 1.

Experiment A2: Mirrored and unmirrored line drawing previews

The design, set size, and analysis plan for this experiment were preregistered (see below for preregistrations).

Participants

Participants were 100 Mechanical Turk workers who participated in exchange for monetary compensation. No participants participated in any other experiments using these line drawings.

Stimuli

Stimuli were the same as in Experiment 1, except there was an additional preview condition using left/right mirror-reversed line drawings, which were created using MATLAB.

Design and procedure

In this experiment, there were four preview conditions: line drawing preview, mirrored line drawing preview, uninformative rectangle preview, and photo preview. The order images appeared in was randomized with the constraint that each target image was presented for the first time before any images were presented for the second time, for each of four presentations of each image (one per preview condition).

Analyses

Our preregistered comparison was a t test between the mirrored line-drawing condition and the unmirrored line drawing condition. Based on Sanocki and Epstein (1997), we also expected at least the unmirrored line drawing condition to be facilitated relative to the rectangle baseline condition.

Results and discussion

We found no significant benefit for either of the line drawing preview conditions compared to the uninformative rectangle baseline (unmirrored significantly slower than baseline, $t(99) = \times 2.93$; $p = 0.004$; $d = -0.29$; mirrored no difference, $t(99) = -0.70$, $p = 0.49$, $d = -0.07$, making any difference between the two line drawing conditions uninterpretable. We did, however, find a photograph preview benefit, $t(99) = 7.66$, $p < 0.001$, $d = 0.77$, suggesting that the lack of line drawing benefit was not due to participants ignoring previews altogether or lack of trying at the task.

Because of a mistake in counterbalancing, the mappings between condition order and target image was not changed across participants as intended (That is, all participants saw a particular target image first in the photograph condition, and another particular target image first in the unmirrored line drawing condition, etc.).

Experiment A3: Mirrored and unmirrored line drawings, blocked design

The design, set size, and analysis plan for this experiment were preregistered (see below for preregistrations).

Participants

Participants were 100 Mechanical Turk workers (25 in each counterbalance condition) who participated in exchange for monetary compensation. No participants participated in any other experiments using these line drawings.

Stimuli

Stimuli were the same as in Experiment A2.

Design and procedure

We reasoned that in Experiment A2 the intermixing of unmirrored and mirrored line drawings may have caused participants to pay less overall attention to both types of line drawing previews. For this reason, we blocked the preview conditions in Experiment A3. Thus, preview conditions were blocked in this experiment, with the order of blocks counterbalanced across participant groups using a balanced Latin square. Other aspects of the design were the same as those in Experiment A2.

Analyses

Again, our preregistered comparison was a *t* test between the mirrored line-drawing condition and the unmirrored line drawing condition; based on Sanocki and Epstein (1997), we again expected at least the unmirrored line drawing condition to be facilitated relative to the rectangle baseline condition.

Results and discussion

We found no significant benefit for either of the line drawing preview conditions compared to the uninformative rectangle baseline: unmirrored versus rect, $t(99) = -0.48$, $p = 0.64$, $d = -0.05$; mirrored versus rect, $t(99) = -0.31$, $p = 0.76$, $d = -0.03$, making any difference between the two line drawing conditions uninterpretable. Again, we found a photograph preview benefit, $t(99) = 3.08$, $p = 0.003$, $d = 0.31$, suggesting that the lack of line drawing benefit was not due to general inattention to preview images. Because the preview types were blocked and introduced at the beginning of each block, the lack of a line drawing benefit was unlikely to be due to participants ignoring all line drawings because mirrored line drawings were unhelpful.

Experiment A4: Unmirrored photograph previews facilitate depth judgments better than mirrored photograph previews

The design, set size, and analysis plan for this experiment were preregistered (see below for preregistrations).

Participants

Participants were 102 Mechanical Turk workers (17 in each of 6 counterbalance conditions) who participated in exchange for monetary compensation. No participants participated in any other experiments using these line drawings.

Stimuli

Target images were the same as in Experiment A2 and A3, and preview images were rectangle previews, photograph previews, or mirror-reversed photograph previews created in MATLAB.

Design and procedure

Preview conditions were blocked in this experiment, with every possible order of blocks equally likely across the six participant groups. Other aspects of the design were the same as in Experiments A2 and A3.

Analyses

Our preregistered comparison was a *t* test between the mirrored and unmirrored photograph preview conditions (note there is a small inconsistency in the preregistration, which says *line drawings* rather than *photographs* in the analysis section, despite the fact that there were no line drawings in this study); we also expected to replicate the unmirrored photograph preview benefit we found in Experiments A2 and A3.

Results and discussion

Our critical analysis found that subjects' median RTs were significantly faster in the unmirrored photo prime condition compared to the mirrored photo prime condition, $t(101) = 2.27$, $p = 0.026$, $d = 0.22$. There was again a benefit of the unmirrored photo prime compared to the rectangle prime, $t(101) = 2.44$, $p = 0.016$, $d = 0.24$, but not for the mirrored photo prime compared to the rectangle prime, $t(101) = -0.24$, $p = 0.81$, $d = -0.02$). This argues against scene-priming benefits originating solely from memory for *global properties* (Oliva, 2005) of scenes, such as openness or amount of perspective, since the photograph and mirrored photograph previews had identical global properties, but only the unmirrored photographs facilitated depth judgments.
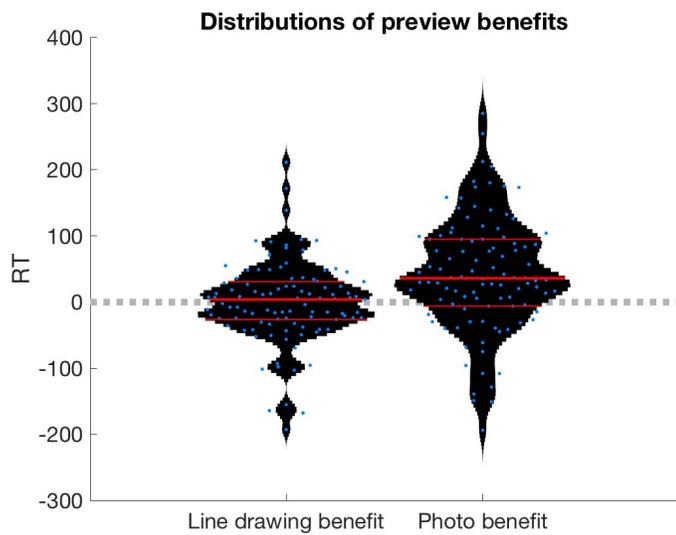
Figure A1. Distributions of individual participants' line drawing and photo preview benefits in Experiment 1. Red lines mark the boundaries of quartiles, and blue points are individual participants' preview benefits in each condition. Note that because we collected many participants but with relatively few trials per participant (to avoid repeating scenes too often), the spread of participants data is larger than in a typical psychophysics study, whereas our power to estimate the grand average across participants and the variation across participants is higher.
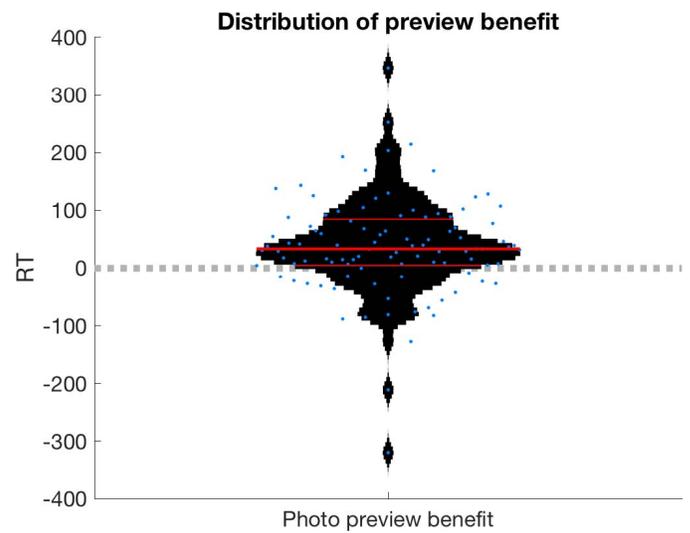
Figure A2. Distributions of individual participants' photo preview benefits in Experiment 2. Red lines mark the boundaries of quartiles, and blue points are individual participants' preview benefits. Note that because we collected many participants but with relatively few trials per participant (to avoid repeating scenes too often), the spread of participants data is larger than in a typical psychophysics study, whereas our power to estimate the grand average across participants and the variation across participants is higher.
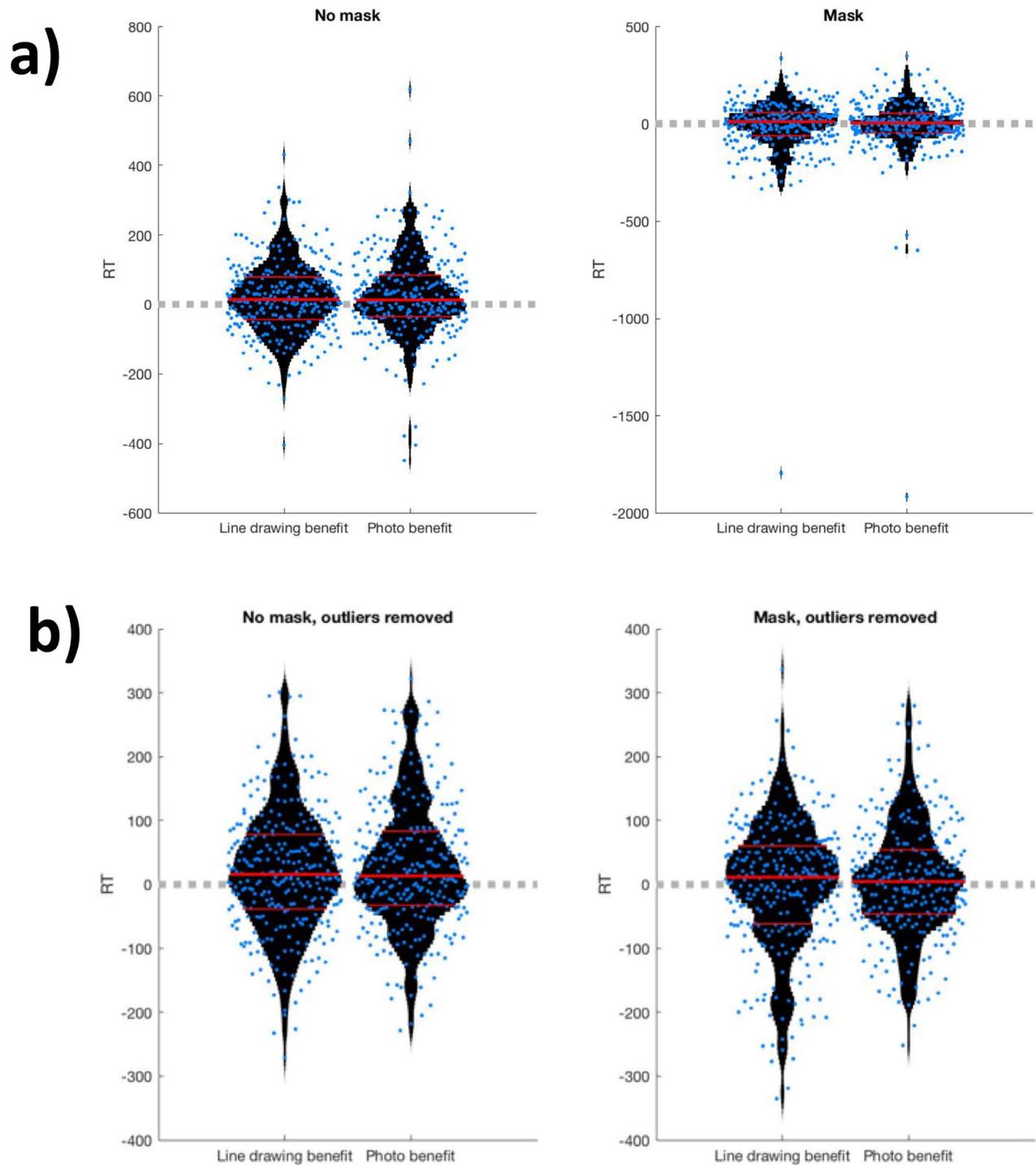
Figure A3. (a) Distributions of individual participants' photo and line drawing preview benefits in Experiment 3 show a few outliers. We identified outliers as any participants who had a median RT in any condition that was 3 *SD* more extreme than the mean. Note that because we collected many participants but with relatively few trials per participant (to avoid repeating scenes too often), the spread of participants data is larger than in a typical psychophysics study, whereas our power to estimate the grand average across participants and the variation across participants is higher. Supplemental analyses show that posthoc removal of these outliers gives the same pattern of results for our main analyses: line-drawing preview benefit, no mask, $t(295) = 3.43$, $p < 0.001$, $d = 0.20$; photo preview benefit, no mask, $t(295) = 4.88$, $p < 0.001$, $d = 0.28$; line-drawing preview benefit, mask, $t(295) = -0.22$, $p = 0.83$, $d = -0.01$; photo preview benefit, mask, $t(295) = 1.45$, $p = 0.15$, $d = 0.08$; line-drawing benefit diminishes with mask, $t(295) = 2.96$, $p = 0.003$, $d = 0.17$; photo benefit diminishes with mask, $t(295) = 2.80$, $p = 0.005$, $d = 0.16$. (b) Distributions of preview benefits with outliers removed.
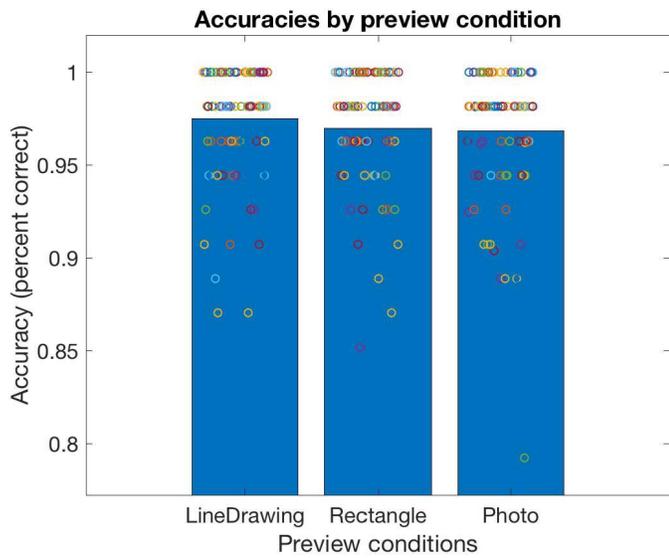
Figure A4. Accuracy data for Experiment 1. Circles are individual participants.
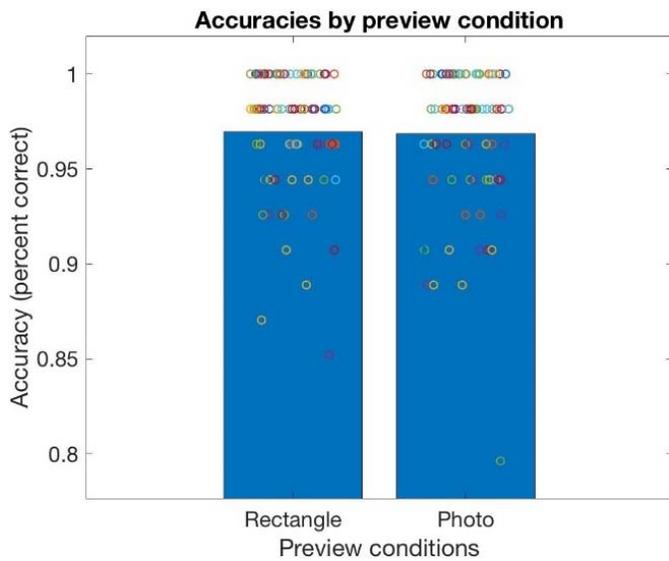
Figure A5. Accuracy data for Experiment 2. Circles are individual participants.
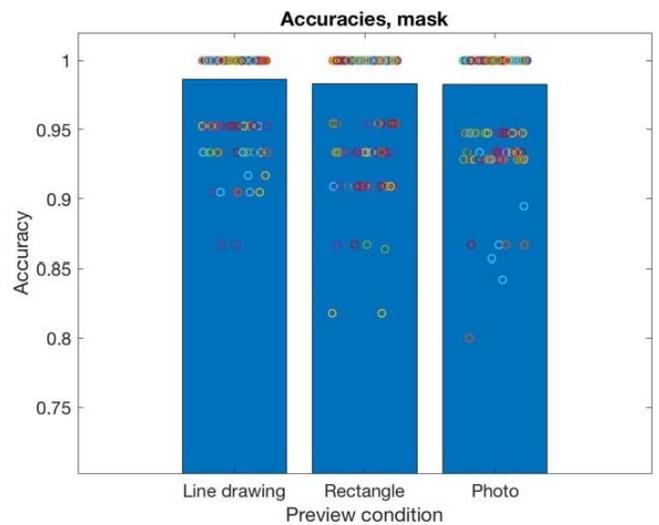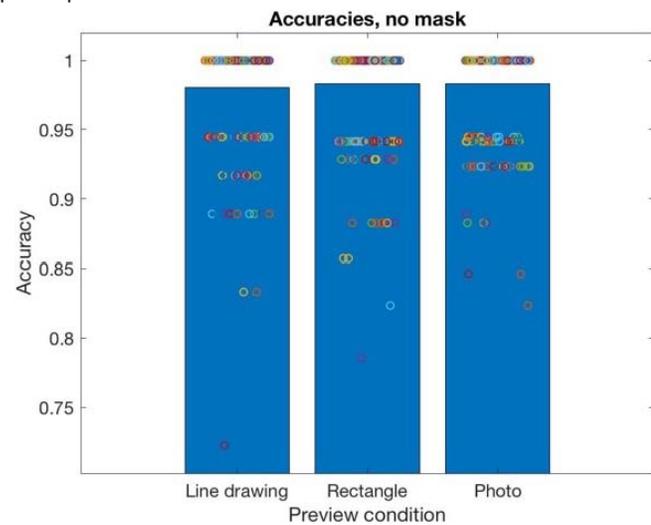
Figure A6. Accuracy data for Experiment 3. Circles are individual participants.
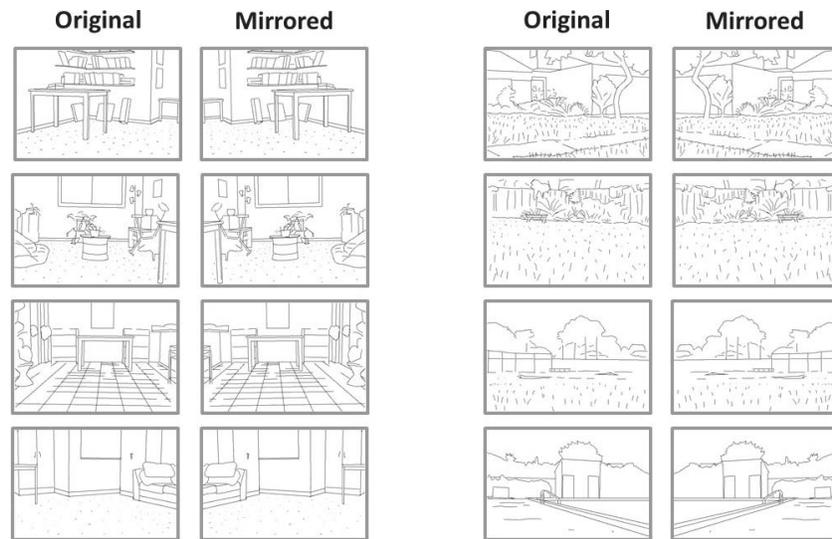
Figure A7. Sanocki and Epstein's (1997) original line drawing stimuli, left columns; mirror-reversed versions of their stimuli, right columns. The images are largely mirror-symmetric, which makes the mirror-reversed line drawing condition in Experiment 3 uninformative.

Figure A8. Line drawings and corresponding photographs from Experiment 1—ordered by accuracy at depth discrimination from the line drawings alone (which is indicated next to each line drawing).

Figure A8. Continued

Figure A8. Continued

Figure A8. Continued

AsPredicted
Pre-Registration made easy
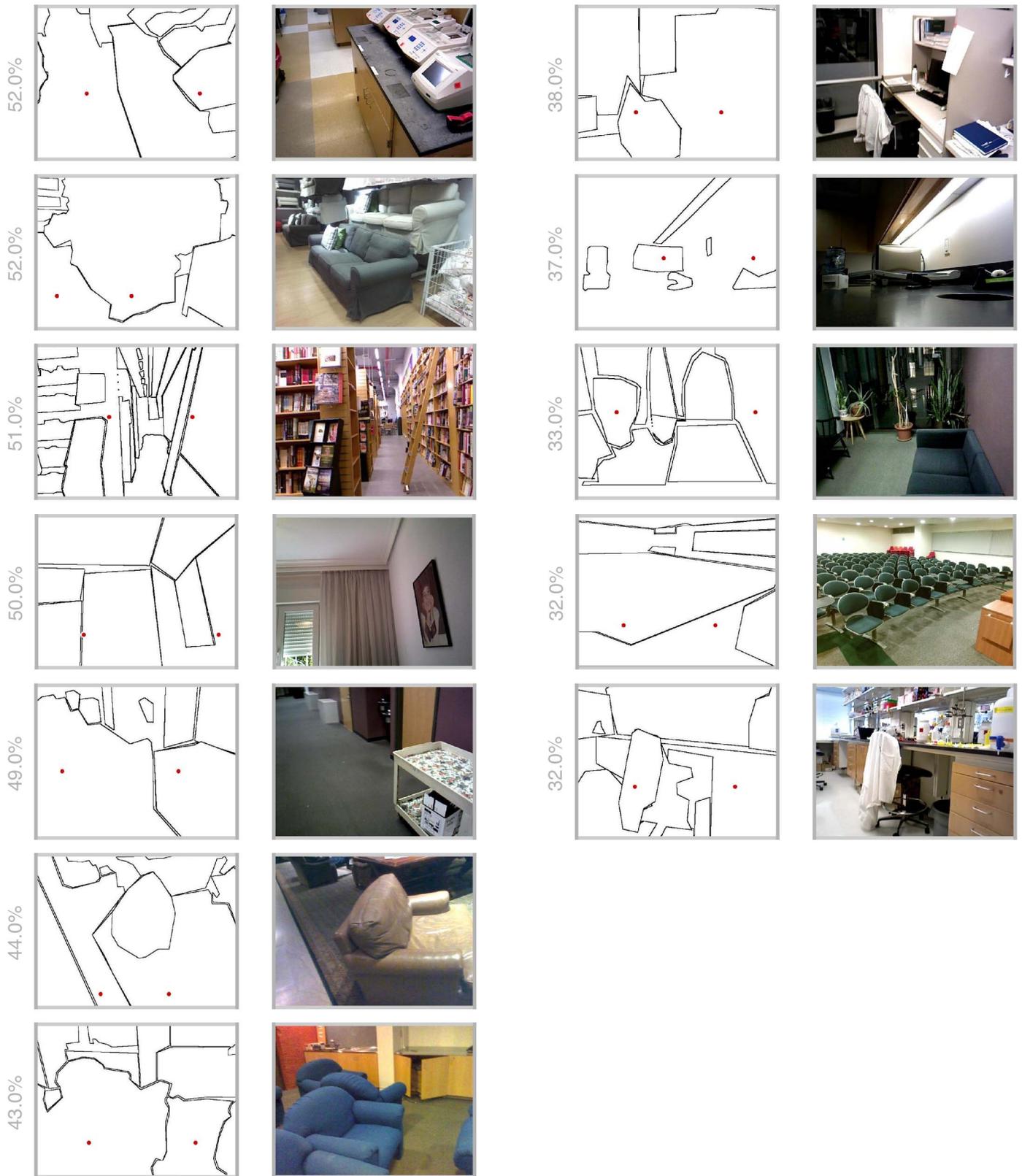
Experiment 1 (#8513)

Author(s)
Anna Shafer-Skelton (University of California, San Diego) - ashafers@ucsd.edu
Tim Brady (University of California, San Diego) - timbrady@ucsd.edu

Created: 02/17/2018 04:35 PM (PT)
Public: 06/27/2018 12:39 PM (PT)

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
Participants will judge which region of a scene image would be closer to them in real life. They will be previewed either using a line drawing, a photograph of the same scene as the target image, or an uninformative rectangle preview. Our question of interest is whether participants' depth judgments will be facilitated by the photo preview and/or the line drawing preview, relative to the uninformative rectangle preview.

3) Describe the key dependent variable(s) specifying how they will be measured.
reaction time (milliseconds) and accuracy (proportion correct answers)

4) How many and which conditions will participants be assigned to?
Three conditions within participants: line drawing preview, photograph preview, and rectangle preview.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will analyze reaction time and accuracy data for the different preview conditions. For reaction times, our statistics will be based on each subject's median reaction time for each condition, and our primary analysis will two paired t-tests: one between median reaction times in the photograph preview condition vs. the rectangle preview condition, and the other between line drawing and rectangle preview conditions.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
Individual trials will be excluded if the participant responded sooner than 150ms after the target image appeared, and only correct trials will be used for reaction time analyses.

Participants will be excluded and replaced with a participant from the same counterbalance order if any of the following apply:

- Their overall accuracy is greater than 3 standard deviations below the mean accuracy.
- Their overall accuracy is less than 55%.
- They responded with the same key on more than 80% of trials.
- Fewer than 50% of their trials can be included in the main analysis, either because they were excluded for having an RT below 150ms, or because they were incorrect.
- Their median RT in any of the preview conditions is above 2 seconds.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
We will collect data from 102 participants (17 in each of 6 counterbalance conditions). If a subject is excluded, they will be replaced with another subject from the same counterbalance condition.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

Version of AsPredicted Questions: 2.00

Available at http://aspredicted.org/yw5bg.pdf
(Permanently archived at http://web.archive.org/web/*/http://aspredicted.org/yw5bg.pdf)

Figure A9. Preregistration for Experiment 1.

AsPredicted
Pre-Registration made easy

Experiment 2 (#8331)

Author(s)
Anna Shafer-Skelton (University of California, San Diego) - ashafers@ucsd.edu
Tim Brady (University of California, San Diego) - timbrady@ucsd.edu

Created: 02/09/2018 01:51 PM (PT)
Public: 06/27/2018 03:09 PM (PT)

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
We want to test whether a preview of a scene photograph (compared to a rectangle) facilitates participants' performance for a judgment unrelated to depth--in this study, whether a square is the left-most or right-most of two shapes presented on top of an image. Target images will be the same as mirroredScenePrime 1-3, except that target locations will be marked by squares and diamonds rather than by dots.

Also in this study, photo and rectangle previews will be intermixed to avoid order effects present in the blocked design in the first version of this experiment.

3) Describe the key dependent variable(s) specifying how they will be measured.
reaction time (milliseconds) and accuracy (proportion correct answers)

4) How many and which conditions will participants be assigned to?
Two conditions within participants: photograph preview image and rectangle preview image. Preview conditions are intermixed to minimize order effects.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will analyze reaction time and accuracy data for the different preview conditions. For reaction times, our statistics will be based on each subject's median reaction time for each condition, and our primary analysis will be a paired t-test between median reaction times in the photograph preview condition vs. the rectangle preview condition.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
Individual trials will be excluded if the participant responded sooner than 150ms after the target image appeared, and only correct trials will be used for reaction time analyses.

Participants will be excluded and replaced with a participant from the same counterbalance order if any of the following apply:

- Their overall accuracy is greater than 3 standard deviations below the mean accuracy.
- Their overall accuracy is less than 55%.
- They responded with the same key on more than 80% of trials.
- Fewer than 50% of their trials can be included in the main analysis, either because they were excluded for having an RT below 150ms, or because they were incorrect.
- Their median RT in either of the two preview conditions is above 2 seconds.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
We will collect data from 100 participants (50 for each of two sets of images). If a subject is excluded, they will be replaced with another subject from the same image list.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

Version of AsPredicted Questions: 2.00

Available at http://aspredicted.org/8g5v2.pdf
(Permanently archived at http://web.archive.org/web/*/http://aspredicted.org/8g5v2.pdf)

Figure A10. Preregistration for Experiment 2.

AsPredicted
Pre-Registration made easy

Experiment 3 (#7008)

Author(s)
Anna Shafer-Skelton (University of California, San Diego) - ashafers@ucsd.edu
Tim Brady (University of California, San Diego) - timbrady@ucsd.edu

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
Participants will judge which of two chairs in a scene image would be closer to them in real life.
Before viewing the target image, they will be given a preview image that is either a line drawing of
the photograph, a mirror-reversed line drawing of the photograph, a scene photograph identical to
the target image (minus the chairs), or a blank image with a rectangle in the center.

One question is whether we can replicate Sanocki & Epstein's (1997) finding that
depth judgments after a line drawing preview are faster than after an uninformative rectangle
preview. We also to plan to examine the extent to which this benefit – as well as any preview benefit from real photographs -- is robust to the presence of
dynamic visual masks during the delay between the preview and the target image.

This version is a higher-powered and more-compellingly-masked version of AsPredicted #6444, mirroredScenePrime_sanockiRepMask, which did not find a
significant benefit of line drawing previews compared to uninformative rectangles.

3) Describe the key dependent variable(s) specifying how they will be measured.
reaction time (milliseconds) and accuracy (proportion correct answers)

4) How many and which conditions will participants be assigned to?
Four preview image conditions within participants: line drawing preview, mirrored line drawing
preview, scene photo, and rectangle prime. Two mask conditions within participants: masked (200ms delay) or un-
masked (84ms delay). In this study, conditions are intermixed, and each combination of condition and scene
image is equally likely across participants.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will analyze reaction time and accuracy data for each condition. For reaction times, our statistics
will be based on each subject's median reaction time for each condition.

We are interested in whether the photograph and line drawings preview conditions provide a
reaction time benefit over the mirrored line drawing and rectangle line drawing conditions (in
unmasked conditions), and whether these benefits persist or are abolished with masks. To assess
this, we will subtract the baseline conditions from their respective matched conditions (e.g.,
unmasked photograph RT – unmasked rectangle RT, unmasked un-mirrored RT – unmasked
mirrored RT) to compute preview benefits and compare these to 0 via t-tests. We will then use
paired t-tests to compare these preview benefits between masked and unmasked conditions.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
Individual trials will be excluded if the participant responded sooner than 150ms after the target
image appeared, and only correct trials will be used for reaction time analyses.

Participants will be excluded and replaced with a participant from the same counterbalance order if
any of the following apply:
- Their overall accuracy is greater than 3 standard deviations below the mean accuracy.
- Their overall accuracy is less than 55%.
- They responded with the same key on more than 80% of trials.
- Fewer than 50% of their trials can be included in the main analysis, either because they were
excluded for having an RT below 150ms, or because they were incorrect.
- Their median RT in any of the prime conditions is above two seconds.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the

Figure A11. Preregistration for Experiment 3.

AsPredicted
Pre-Registration made easy

number will be determined.

306 subjects will be collected for this version of the experiment (102 participants in each of the 3 possible counterbalance conditions). If a subject is excluded, they will be replaced with another subject from the same counterbalance condition.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
In order to test how long any effects may take to develop, a secondary analysis will separately analyze the data (across participants) for the first half and second half of the experiment. In order to be included in the secondary analysis, participants must meet all main-experiment qualifications, in addition to having median RTs in all of the *combinations* of experiment-half and condition that are less than or equal to 2 seconds.

Version of AsPredicted Questions: 2.00

Available at http://aspredicted.org/rk6f6.pdf
(Permanently archived at http://web.archive.org/web/*/http://aspredicted.org/rk6f6.pdf)

Figure A11. Continued

AsPredicted
Pre-Registration made easy

Experiment A1: lineDrawingDepthJudge - San Diego, January 2018 (#7716)

Author(s)
Tim Brady (University of California, San Diego) - timbrady@ucsd.edu
Anna Shafer-Skelton (University of California, San Diego) - ashafers@ucsd.edu

Created: 01/05/2018 12:25 PM (PT)
Public: 01/09/2019 10:43 PM (PT)

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
We want to know whether there is sufficient scene layout information in a set of line drawings for participants to judge which of two dots would be closer to the viewer in three dimensions. The line drawings were created from a database of scene images that includes ground-truth depth measurements.

3) Describe the key dependent variable(s) specifying how they will be measured.
The dependent variable is accuracy (proportion correct answers).

4) How many and which conditions will participants be assigned to?
This experiment does not include different conditions.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
Our primary analysis is a t-test to determine whether participants are above chance at identifying which dot is on the closer part of the line drawings.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.
Individual trials will be excluded if the participant responded sooner than 150ms after the target image appeared.

Participants will be excluded and replaced if:
- Their overall accuracy is greater than 3 standard deviations below the mean accuracy
- They responded with the same key on more than 80% of trials
- If more than 50% of their trials are excluded due to responding too soon.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
We will collect 100 participants' worth of data.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)
We may also use this data set to divide line drawings into "informative" (above-chance responses) and "uninformative" (not-above-chance responses) groups, for performing post-hoc analyses on mirroredScenePrime1, 2, and/or 3.

Figure A12. Preregistration for Experiment A1.

Experiment A2: mirroredScenePrime1 - San Diego, spring 2017 (#3881)

Author(s)
Anna Shafer-Skelton (University of California, San Diego) - ashafers@ucsd.edu
Tim Brady (University of California, San Diego) - timbrady@ucsd.edu

1) Have any data been collected for this study already?
No, no data have been collected for this study yet

2) What's the main question being asked or hypothesis being tested in this study?
Participants will judge which region of a scene image would be closer to them in real life. They will be primed either using a line drawing that has consistent global features (e.g., perspective, openness) and scene layout information to the target image (un-mirrored line drawing prime) or a line drawing with the same global features but different scene layout information (mirrored line drawing prime). Our hypothesis is that there will be a benefit for the un-mirrored line drawing prime.

3) Describe the key dependent variable(s) specifying how they will be measured.
reaction time (milliseconds) and accuracy (proportion correct answers)

4) How many and which conditions will participants be assigned to?
Four conditions within participants: un-mirrored line drawing prime, mirrored line drawing prime, photograph prime, and rectangle prime.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will analyze reaction time and accuracy data for the different prime conditions. For reaction times, our statistics will be based on each subject's median reaction time for each condition, and our primary analysis will be a paired t-test between median reaction times in the mirrored line drawing condition vs. the un-mirrored line drawing condition.

Individual trials will be excluded if the participant responded sooner than 150ms after the target image appeared, and only correct trials will be used for reaction time analyses.

Participants will be excluded and replaced with a participant from the same counterbalance order if any of the following apply:

- Their overall accuracy is greater than 3 standard deviations below the mean accuracy.
- Their overall accuracy is less than 55%.
- They responded with the same key on more than 80% of trials.
- Fewer than 50% of their trials can be included in the main analysis, either because they were excluded for having an RT below 150ms, or because they were incorrect.

6) Any secondary analyses?

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
We will collect data from 100 participants (25 in each counterbalance condition). If a subject is excluded, they will be replaced with another subject from the same counterbalance condition.

8) Anything else you would like to pre-register? (e.g., data exclusions, variables collected for exploratory purposes, unusual analyses planned?)

Figure A13. Preregistration for Experiment A2.

AsPredicted
Pre-Registration made easy

Experiment A3: mirroredScenePrime2 - San Diego, May 2017 (#3972)

Author(s)
Anna Shafer-Skelton (University of California, San Diego) - ashafers@ucsd.edu
Tim Brady (University of California, San Diego) - timbrady@ucsd.edu

Created: 05/04/2017 10:28 AM (PT)
Public: 01/09/2019 10:49 PM (PT)

1) Have any data been collected for this study already?
No, no data have been collected for this study yet

2) What's the main question being asked or hypothesis being tested in this study?
Participants will judge which region of a scene image would be closer to them in real life. They will be primed either using a line drawing that has consistent global features (e.g., perspective, openness) and scene layout information to the target image (un-mirrored line drawing prime) or a line drawing with the same global features but different scene layout information (mirrored line drawing prime). Our hypothesis is that there will be a benefit for the un-mirrored line drawing prime.

3) Describe the key dependent variable(s) specifying how they will be measured.
reaction time (milliseconds) and accuracy (proportion correct answers)

4) How many and which conditions will participants be assigned to?
Four conditions within participants: un-mirrored line drawing prime, mirrored line drawing prime, photograph prime, and rectangle prime. In this study, prime conditions are blocked, and order is counterbalanced across participants.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will analyze reaction time and accuracy data for the different prime conditions. For reaction times, our statistics will be based on each subject's median reaction time for each condition, and our primary analysis will be a paired t-test between median reaction times in the mirrored line drawing condition vs. the un-mirrored line drawing condition.

Individual trials will be excluded if the participant responded sooner than 150ms after the target image appeared, and only correct trials will be used for reaction time analyses.

Participants will be excluded and replaced with a participant from the same counterbalance order if any of the following apply:

- Their overall accuracy is greater than 3 standard deviations below the mean accuracy.
- Their overall accuracy is less than 55%.
- They responded with the same key on more than 80% of trials.
- Fewer than 50% of their trials can be included in the main analysis, either because they were excluded for having an RT below 150ms, or because they were incorrect.

6) Any secondary analyses?

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
We will collect data from 100 participants (25 in each counterbalance condition). If a subject is excluded, they will be replaced with another subject from the same counterbalance condition.

8) Anything else you would like to pre-register? (e.g., data exclusions, variables collected for exploratory purposes, unusual analyses planned?)

Figure A14. Preregistration for Experiment A3.

Experiment A4: MirroredScenePrime3 - San Diego, May 2017 (#4039)

Author(s)
Anna Shafer-Skelton (University of California, San Diego) - ashafers@ucsd.edu
Tim Brady (University of California, San Diego) - timbrady@ucsd.edu

1) Have any data been collected for this study already?
No, no data have been collected for this study yet

2) What's the main question being asked or hypothesis being tested in this study?
Participants will judge which region of a scene image would be closer to them in real life. They will be primed either using a scene photograph identical to the target image (minus the dots indicating which locations to judge) or a mirror-reversed scene photograph. Our main hypothesis is that there will be a benefit for the un-mirrored compared to the mirrored prime.

3) Describe the key dependent variable(s) specifying how they will be measured.
reaction time (milliseconds) and accuracy (proportion correct answers)

4) How many and which conditions will participants be assigned to?
Three conditions within participants: un-mirrored scene photo, mirrored scene photo, and rectangle prime. In this study, prime conditions are blocked, and every potential order of blocks is equally likely across participants.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.
We will analyze reaction time and accuracy data for the different prime conditions. For reaction times, our statistics will be based on each subject's median reaction time for each condition, and our primary analysis will be a paired t-test between median reaction times in the mirrored line drawing condition vs. the un-mirrored line drawing condition.

Individual trials will be excluded if the participant responded sooner than 150ms after the target image appeared, and only correct trials will be used for reaction time analyses.

Participants will be excluded and replaced with a participant from the same counterbalance order if any of the following apply:

- Their overall accuracy is greater than 3 standard deviations below the mean accuracy.
- Their overall accuracy is less than 55%.
- They responded with the same key on more than 80% of trials.
- Fewer than 50% of their trials can be included in the main analysis, either because they were excluded for having an RT below 150ms, or because they were incorrect.
- Their median RT in any of the three prime conditions is above two seconds.

6) Any secondary analyses?
In order to test how long any effects may take to develop, a secondary analysis will separately analyze the data (across participants) for each block of the experiment (e.g., 1st block, 2nd block, etc.). In order to be included in the secondary analysis, participants must meet all main-experiment qualifications, in addition to having median RTs in all of the *combinations* of block and condition that are less than or equal to 2 seconds.

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.
102 subjects will be collected for this version of the experiment (17 participants in each of the 6 possible counterbalance conditions). If a subject is excluded, they will be replaced with another subject from the same counterbalance condition.

8) Anything else you would like to pre-register? (e.g., data exclusions, variables collected for exploratory purposes, unusual analyses planned?)

Figure A15. Preregistration for Experiment A4.