Report

Current Biology

Boundaries Extend and Contract in Scene Memory Depending on Image Properties

Highlights

- Boundary contraction is as common as boundary extension for naturalistic scene images
- Boundary transformation direction is highly predictable from basic image properties
- Boundary transformations occur during recognition, recall, and minimal memory load
- These results dispute boundary extension as a universal phenomenon of scene memory

Authors

Wilma A. Bainbridge, Chris I. Baker

Correspondence

wilma@uchicago.edu

In Brief

Boundary extension has been regarded as a universal phenomenon in scene memory, reflecting scene extrapolation beyond image edges. Bainbridge and Baker examine 1,000 naturalistic scene images (n = 2,000) and discover boundary contraction as equally common. Distortions are predictable from image properties and occur even during minimal memory tasks.

Bainbridge & Baker, 2020, Current Biology *30*, 1–7
February 3, 2020 Published by Elsevier Ltd.
https://doi.org/10.1016/j.cub.2019.12.004

Please cite this article in press as: Bainbridge and Baker, Boundaries Extend and Contract in Scene Memory Depending on Image Properties, Current Biology (2019), https://doi.org/10.1016/j.cub.2019.12.004

Current Biology
Report

CellPress

Boundaries Extend and Contract in Scene Memory Depending on Image Properties

Wilma A. Bainbridge[1,2,3,4,*] and Chris I. Baker[1]
[1]Laboratory of Brain and Cognition, National Institute of Mental Health, Bethesda, MD 20814, USA
[2]Twitter: @WilmaBainbridge
[3]Present address: Department of Psychology, University of Chicago, Chicago, IL 60637, USA
[4]Lead Contact
*Correspondence: wilma@uchicago.edu
https://doi.org/10.1016/j.cub.2019.12.004

SUMMARY

Boundary extension, a memory distortion in which observers consistently recall a scene with visual information beyond its boundaries, is widely accepted across the psychological sciences as a phenomenon revealing fundamental insight into memory representations [1–3], robust across paradigms [1, 4] and age groups [5–7]. This phenomenon has been taken to suggest that the mental representation of a scene consists of an intermingling of sensory information and a schema that extrapolates the views of a presented scene [8], and it has been used to provide evidence for the role of the neocortex [9] and hippocampus [10, 11] in the schematization of scenes during memory. However, the study of boundary extension has typically focused on object-oriented images that are not representative of our visuospatial world. Here, using a broad set of 1,000 images tested on 2,000 participants in a rapid recognition task, we discover "boundary contraction" as an equally robust phenomenon. Further, image composition largely drives whether extension or contraction is observed—although object-oriented images cause more boundary extension, scene-oriented images cause more boundary contraction. Finally, these effects also occur during drawing tasks, including a task with minimal memory load—when participants copy an image during viewing. Collectively, these results show that boundary extension is not a universal phenomenon and put into question the assumption that scene memory automatically combines visual information with additional context derived from internal schema. Instead, our memory for a scene may be largely driven by its visual composition, with a tendency to extend or contract the boundaries equally likely.

RESULTS

Boundary extension is a widely accepted psychological phenomenon in which observers will recall an image as containing visual information beyond the boundaries of the originally studied image. For example, when drawing a house from memory (Figure 1A), observers will recall the house surrounded by space, even though it is cropped by the image boundaries. This effect has been taken to reveal deep insight into the nature of memory representations, specifically how the brain automatically schematizes visual information for scene memory [8–11], and is considered so intrinsic to human cognition that it is included in psychology textbooks [12–14] and in questions in the US Graduate Record Examination (GRE) [15]. It is thought that during initial viewing, the brain encodes a scene memory as an intermingling of visual information from the image and a top-down scene schema that extrapolates information beyond the visual boundaries. Then, when recalling the image, observers are unable to differentiate sensory and schematic information in the memory and recall a memory that contains both, extending the boundaries of the original image into the schema. However, despite the emphasis on boundary extension as a scene-specific phenomenon [8, 16], research has almost exclusively tested images of one or a few central, close-up objects against a generic or repeating background [2]. This is a narrow definition of a scene and does not reflect our natural visual experience of scenes, which are typically made of dozens of objects dispersed across a spatial layout. In a prior study designed to measure the visual information in memory (n = 30), we collected 407 drawings of a variety of 60 scenes from memory [17] and obtained online crowd-sourced scores of boundary transformation for each drawing (Figure 1B). Surprisingly, only 55.0% of the images showed boundary extension, although 43.3% showed the opposite effect of boundary contraction. This led us to wonder whether boundary extension is indeed ubiquitous across visual scenes. Here, we test the extent of boundary transformations: first, with a replication of [17] using a completely different paradigm and, second, with two separate experiments testing a new set of 1,000 photographs across 2,000 participants and two judgment tasks.

First, we wanted to replicate the findings from [17] using a different paradigm. Although boundary extension was originally studied using drawings [1], recent studies have employed a rapid serial visual presentation (RSVP) paradigm [4, 10]. One hundred participants took part in an online experiment on Amazon Mechanical Turk (AMT), in which they viewed an image for 250 ms, followed by a dynamic scrambled mask (250 ms), and then viewed the original image a second time (Figure 1B). Importantly, participants did not know the two images were identical

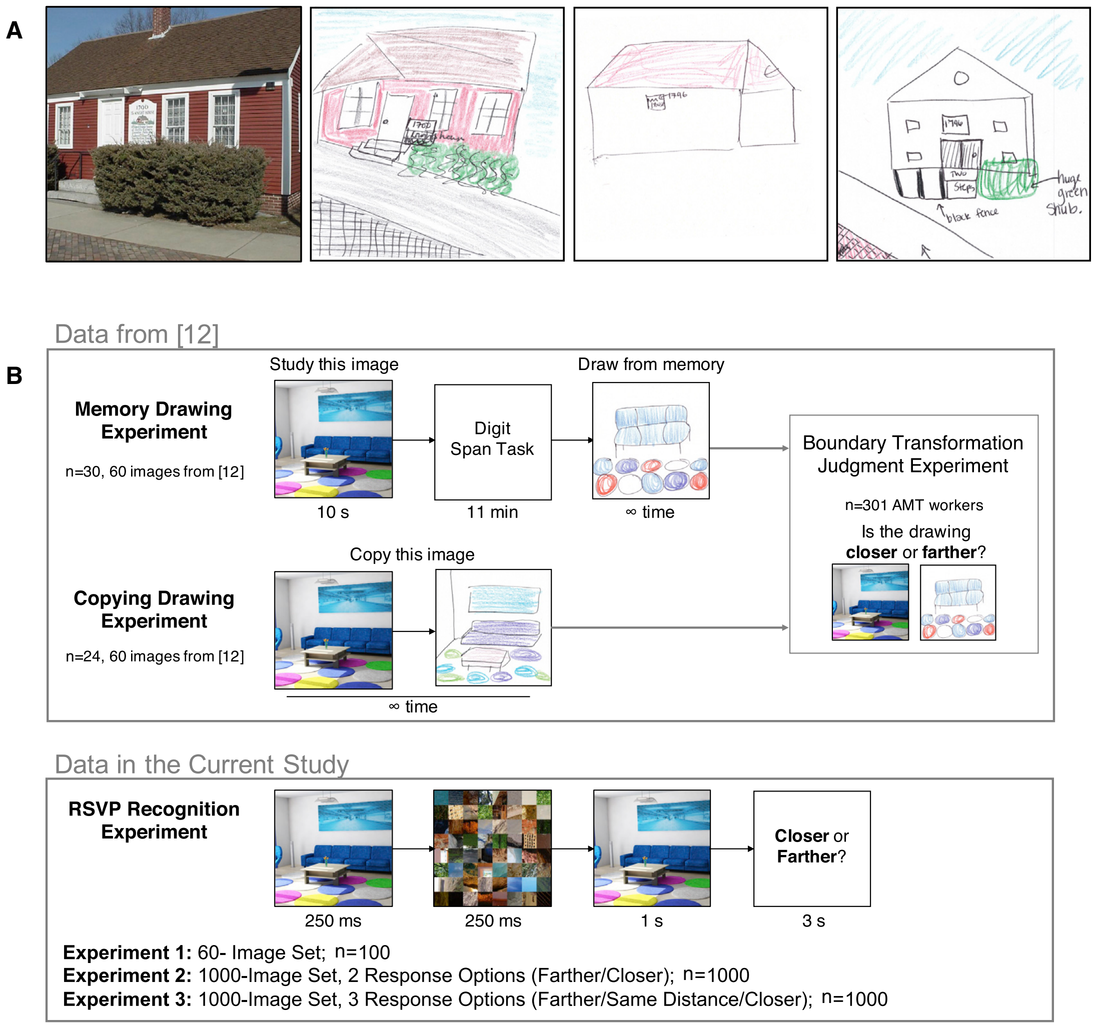

Figure 1. An Example of Boundary Extension and the Experimental Paradigms

(A) Three separate participants drew the same image of a house from memory (from [17]). All three participants extended the boundaries around the house, drawing empty space above it and to the right and left, even though the house in the original image is truncated by the boundaries of the image.

(B) Depictions of the methods of the experiments described in the current study (see STAR Methods). Data from three experiments from [17] are discussed: (1) the Memory Drawing Experiment, in which participants drew scene images from memory; (2) the Copying Drawing Experiment, in which participants copied images while viewing them; and (3) the Boundary Transformation Judgment Experiment, where separate online workers judged the level of boundary transformation for the drawings. In the current study, three experiments were conducted using the RSVP Recognition paradigm: (1) an experiment with the 60-image set; (2) an experiment with the 1,000-image set, with 2 response options; and (3) a second experiment with the 1,000-image set, with 3 response options.

See Figure S1 for performance comparison across the Memory Drawing Experiment and RSVP Recognition paradigm.

and reported whether the second image was closer or farther than the first. Using this test, 51.7% of images showed boundary extension, although 48.3% showed boundary contraction, with no significant difference in the proportion (Pearson chi-square test: $p = 0.715$). Further, there was a significant correlation in boundary transformation score between the memory drawing experiment and the RSVP recognition experiment (Spearman rank correlation: $\rho = 0.39$; $p = 0.0018$), demonstrating that the existence of both boundary extension and contraction replicates at the image level across paradigms (Figure S1).

Although examples of both boundary extension and contraction were observed in this set of 60 images, what patterns emerge if we measure a large, diverse set of real-world images? We assembled a set of 1,000 images (Figure 2A), with 500 each randomly and representatively sampled from two image databases: Google Open Images (GOI) [18] and the Scene Understanding Database (SUN) [19]. We chose these two databases given differences in the nature of the images; GOI contains 1.7 million object-oriented scenes, categorized by an object label (e.g., flower and pillow) and often containing a single, central object, similar to those used in prior boundary extension research. Meanwhile, SUN is scene oriented, containing 131,000 images representing 908 scene categories (e.g., castle and zoo). Using these images, we conducted two large-scale RSVP recognition experiments, the first employing a 2-option response (closer/farther) and the second giving participants a third “same distance” option, to see whether boundary transformations emerged even if participants could report no change between

Please cite this article in press as: Bainbridge and Baker, Boundaries Extend and Contract in Scene Memory Depending on Image Properties, Current Biology (2019), https://doi.org/10.1016/j.cub.2019.12.004

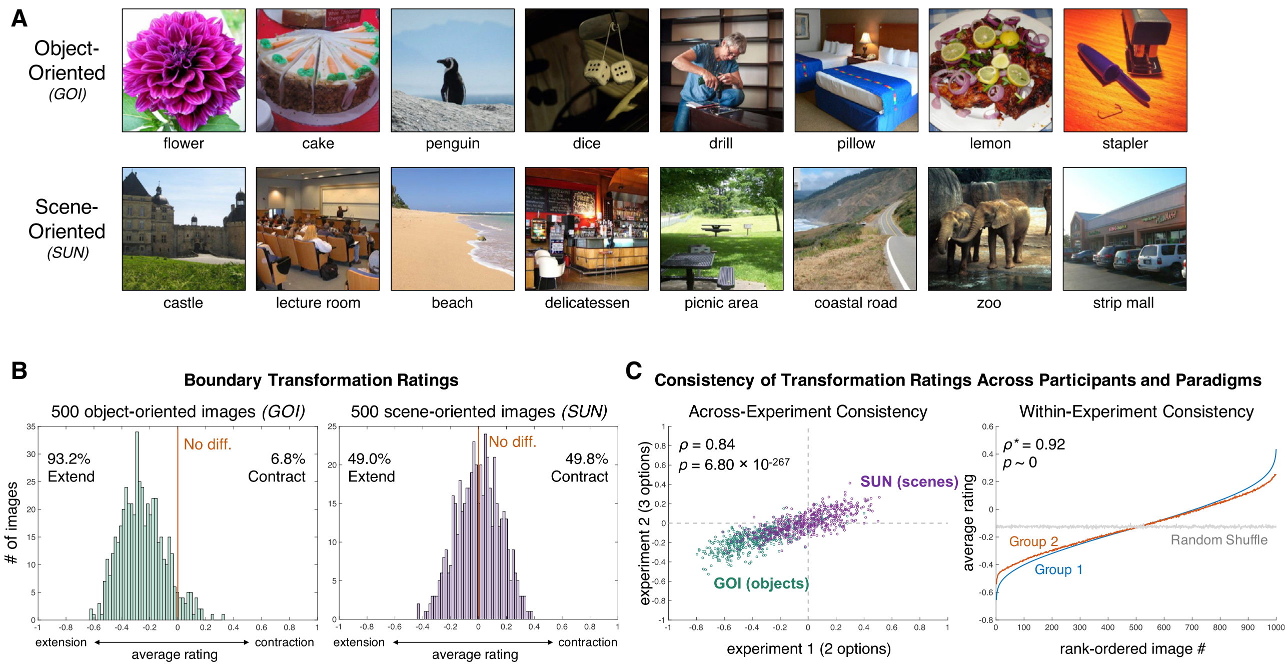

Figure 2. Example Images and Boundary Transformation Distribution

(A) Examples from the 1,000 images and their labels used in the current study, originating from the object-oriented GOI database and the scene-oriented SUN database. The object-oriented images tend to focus on the labeled object but sometimes contain other objects or perspectives. In contrast, the scene-oriented images tend to be broad perspectives of a scene but sometimes contain central objects.

(B) Histograms of average boundary transformation rating for each image, averaged across two experiments (2-option and 3-option RSVP experiments). Although object-oriented images, similar to those in previous boundary-extension experiments, find a high rate of boundary extension (93.2%), scene-oriented images show equal rates of boundary extension and contraction (49.0% versus 49.8%, respectively). Histograms separated by experiment are shown in Figure S2.

(C) Results of consistency analyses on the boundary transformation scores across all images. The left shows a scatterplot of boundary transformation scores across the two 1,000-image sets experiments (2-option and 3-option experiments), with the GOI images in green and the SUN images in purple. The boundary transformation scores for the two experiments are significantly correlated (Spearman rank correlation). The right shows results of a split-half consistency analysis on boundary transformation scores across all images and both experiments. The blue line indicates average boundary transformation scores determined by a random half of the participants across 1,000 iterations, and the orange line indicates the average boundary scores from the other half of participants, sorted in the same order. The gray line indicates the other half of participants sorted randomly. Group 1 and Group 2 are highly similar and significantly correlated, demonstrating that participants are highly consistent in the boundary transformation ratings they make for a given image. Split-half consistency analyses separated by experiment are shown in Figure S2.

images. Each experiment had 1,000 participants, with 2,000 participants in total making judgments on the 1,000 images (n = 200 per image). Given this broader and more representative sampling of visual scenes, is boundary extension the primary phenomenon?

Striking differences emerge for the two image sets (Figure 2B). In each experiment, the 500 object-oriented GOI images show a strong tendency for boundary extension, with 93.2% of images on average showing extension and 6.8% showing contraction (Pearson chi-square test: $\chi^2 = 746.5$; $p \sim 0$). In contrast, the 500 scene-oriented SUN images show an even split, with 49.0% of images on average displaying extension and 49.8% displaying contraction ($\chi^2 = 0.06$; $p = 0.800$). Although this may imply that boundary contraction is just as common as boundary extension in these more naturalistic scenes, an alternate possibility is that participants could not perceive a change and responded at chance or it could reflect across-participant variability. To address these possibilities, we conducted split-half consistency analyses across 1,000 iterations to see whether random halves of the participant pool showed agreement (see STAR Methods) and applied the Spearman-Brown prediction formula to quantify reliability of the entire image set. When participants were forced to judge an image change (2 options), both the object-oriented GOI images (Spearman-Brown reliability: $\rho^* = 0.80$; $p < 0.0001$) and scene-oriented SUN images ($\rho^* = 0.79$; $p < 0.0001$) showed significant consistency in participant responses. When participants could indicate no change (3 options), there was still significant consistency in participant responses for both the GOI images ($\rho^* = 0.76$; $p < 0.0001$) and SUN images ($\rho^* = 0.71$; $p < 0.0001$). There was also a significant correlation across experiments for all images (Spearman rank correlation: $\rho = 0.84$; $p = 6.80 \times 10^{-267}$), as well as when split into GOI images ($\rho = 0.75$; $p = 4.38 \times 10^{-92}$) and SUN images ($\rho = 0.72$; $p = 1.36 \times 10^{-81}$). These results indicate that whether an image extends or contracts is truly intrinsic to an image, consistent across observers, experiments, and not due to random responding (Figures 2C and S2). The data from these two experiments were thus combined for all other analyses (internal split-half consistency when combined: $\rho^* = 0.92$; $p \sim 0$). We also examined consistency based on the proportion of times

Please cite this article in press as: Bainbridge and Baker, Boundaries Extend and Contract in Scene Memory Depending on Image Properties, Current Biology (2019), https://doi.org/10.1016/j.cub.2019.12.004

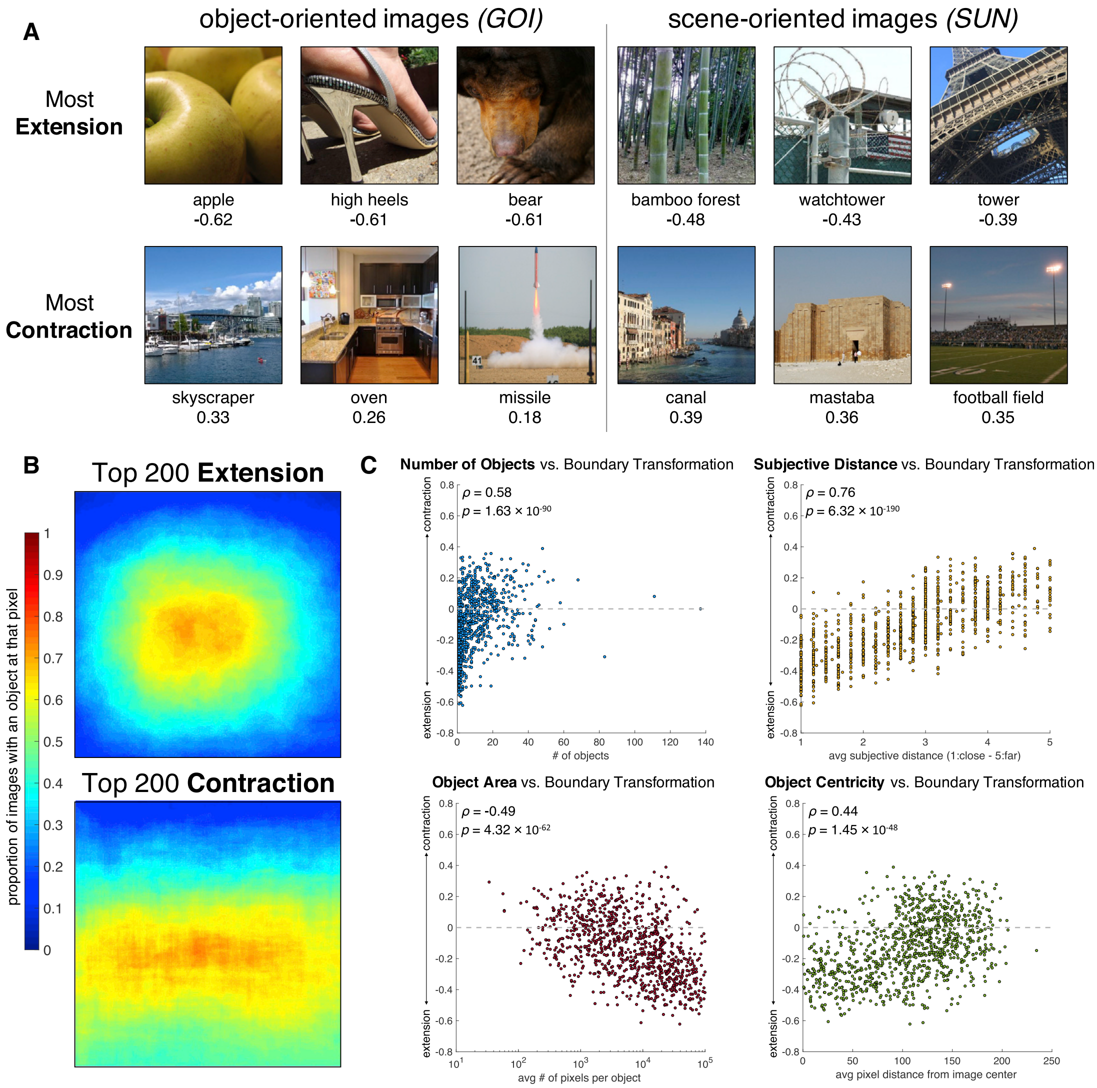

Figure 3. The Influence of Image Composition on Boundary Transformation

(A) The images from the 1,000 image set showing the highest boundary extension and contraction, split by database. Sores range from −1 (extension) to 1 (contraction). Across databases, the images eliciting the most extension are images with few objects at a very close subjective distance from the observer, while those eliciting the most contraction tend to be wide images of scenes.

(B) The location of objects within the top 200 images that show the most boundary extension and contraction. Each pixel is colored by the proportion of images with an object at that pixel. Extending images tend to have a centrally located object, while contracting images tend to have a spread of objects along the lower visual field.

(C) Scatterplots showing the relationship between average boundary transformation rating and four other metrics: number of objects in the image; subjective ratings of distance to the main object (1, close to 5, far); average object area (total number of pixels) per image; and average object distance from the image center (pixels). Statistics indicate results of Spearman's rank correlations. Images that elicit more boundary extension have fewer, larger, centrally located, subjectively close objects, while images that elicit more boundary contraction have several, smaller, dispersed, far objects. These results show that direction of boundary transformation is highly related to image composition and that more traditional scenes cause more boundary contraction.

images were placed into the same category (extension or contraction) by random participant splits (across 1,000 iterations). With the SUN images (which contain equal examples of extension and contraction), both boundary extension and boundary contraction show significant sorting agreement (permuted estimates of chance; extension: M = 87.7%, $p < 0.0001$; contraction: M = 69.6%, $p < 0.0001$). In all, these results indicate that although object-oriented images will tend to cause boundary extension, scene-oriented images will have an equal tendency to elicit boundary extension or contraction, with high agreement across observers. Thus, prior studies may have primarily observed a boundary extension effect because they sampled object-oriented images.

What is it about the images that makes some elicit boundary extension and others contraction? Qualitatively, one can observe (Figure 3A) that images that cause extension tend to have few objects at a close-up perspective, although images that cause contraction tend to have several objects at a farther

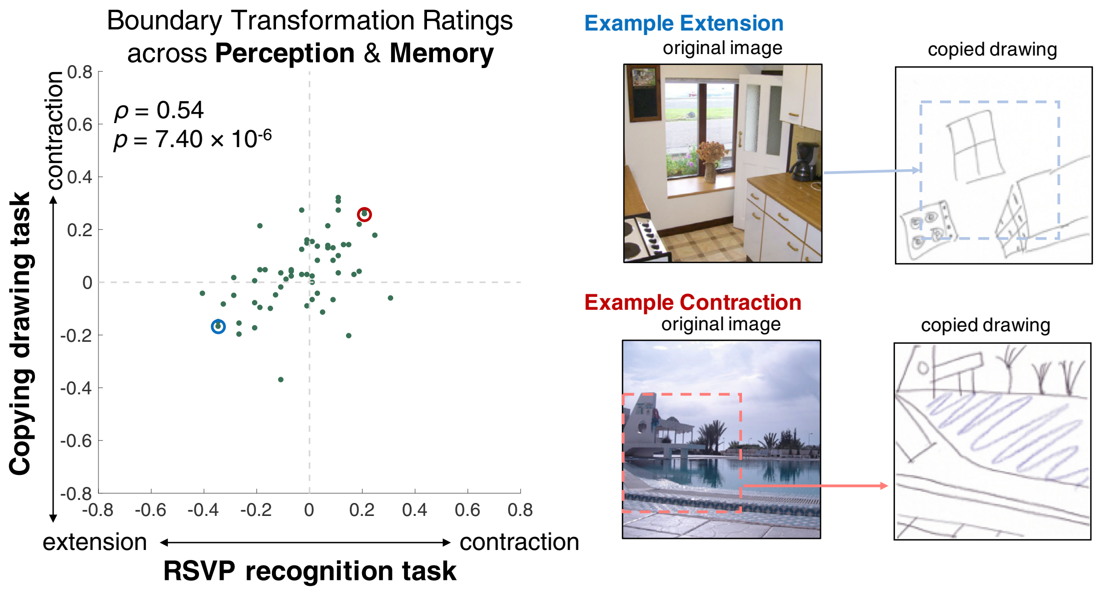

Figure 4. Boundary Transformation Scores Replicate across Memory and Image-Copying Paradigms

(Left) A scatterplot comparing boundary transformation scores in the RSVP recognition task and a copying drawing task where participants copied an image while viewing it. Boundary transformation scores correlate significantly between tasks, with many images showing similar boundary transformations in both tasks. (Right) Example drawings exhibiting boundary extension and contraction (circled in the scatterplot) by participants instructed to copy a photograph while viewing it are shown.

distance. We investigated the locations and sizes of the objects in the images, using human-labeled annotations of object outlines from the GOI and SUN databases [18–20]. Average heatmaps of object locations in the top 200 boundary-extending and top 200 boundary-contracting images show a clear difference (Figure 3B): top-extending images contain central objects, although top-contracting images have a greater spread of objects below the horizon. Further, the top boundary-extending images have few objects (M = 2.4 objects; SD = 4.2), in contrast with boundary-contracting images (M = 15.4, SD = 13.7; independent samples t test: $t(397) = 12.78$, $p = 1.43 \times 10^{-31}$). To objectively quantify these patterns, we compared the boundary transformation ratings across all 1,000 images to several image metrics: number of objects; average object area; average object distance from center; and subjective distance from the observer to the most salient object (Figure 3C). There was a significant positive correlation between boundary transformation rating and number of objects (Spearman rank correlation: $\rho = 0.58$; $p = 1.63 \times 10^{-90}$) and average object centricity ($\rho = 0.44$; $p = 1.45 \times 10^{-48}$) and a significant negative correlation with average object area ($\rho = -0.49$; $p = 4.32 \times 10^{-62}$). To investigate subjective distance, participants in an online AMT experiment rated ‘‘how far away is the main object’’ in each image (n = 5/image), on a scale from 1 (extreme close up) to 5 (far away) [21]. Subjective distance from the main object was significantly correlated with boundary transformation rating ($\rho = 0.76$; $p = 6.32 \times 10^{-190}$). Thus, images with several, smaller, dispersed, and distant objects are more likely to elicit boundary contraction, while images with fewer, larger, more central, and close objects are more likely to elicit boundary extension.

Finally, are these boundary transformations limited to memory or do they occur even during a largely perceptual task? We revisited the 60 scene images tested at the beginning of this study, reanalyzing data from a copying drawing experiment [17]. In this experiment, 24 participants were instructed to produce drawings of 30–35 scene images while viewing them, with no time limit. This copying experiment had minimal memory load, as participants had continuous access to the image and had unlimited ability to correct their drawings. The resulting 728 drawings were scored on AMT (n = 7/drawing) for whether they were closer or farther than the original image. Of these 60 images, significantly more showed contraction (63.3%) than extension (35.0%; Pearson chi-square test: $\chi^2 = 9.64$; $p = 0.002$; Figure 4). Importantly, boundary transformation scores for copied drawings and memory drawings of an image were significantly correlated (Spearman rank correlation: $\rho = 0.64$; $p = 3.81 \times 10^{-8}$) as well as with its RSVP recognition boundary transformation score ($\rho = 0.54$; $p = 7.40 \times 10^{-6}$). This shows that boundary transformation score is consistent to an image, whether measured with a memory task or a copying task with minimal memory load.

DISCUSSION

In this study, we show that although boundary extension has been considered a universal phenomenon of scene memory, the opposite effect of boundary contraction is just as common for naturalistic scene images. We also observed a strong relationship between boundary transformation and an image's visual composition, suggesting the effect may be largely driven by the images themselves rather than scene category schemas of the images. Further, we observe boundary distortions even during minimal memory load (i.e., copying an image), suggesting that this effect may not be memory specific.

These results serve as clear evidence against the multisource memory model originally used to describe boundary extension [8]. This model suggests that a scene memory comprises an intermingling of sensory information and a top-down extrapolated scene schema. Accordingly, only boundary extension is expected because it reflects a source monitoring error, in which an observer cannot disentangle information from these multiple sources and thus recalls more than observed. The current study shows an equally strong phenomenon of boundary contraction, which should not occur within this multisource memory framework.

An alternate explanation could be a normalization process, in which images are normalized toward a spatial scene schema [16, 22]. In the domain of object perception and memory, a consistent bias has been observed toward the normative size for real-world objects [23]. In the domain of scene memory, a regression to the typical scene-viewing distance has only been reported after long delays (48 h), but not at shorter timescales [24]. In the current study, boundary transformation scores show the highest correlation with subjective image distance, with the scenes rated an average subjective distance of "nearby/same room" (4–20 ft). Subjectively, far scenes thus tend to contract and subjectively close scenes tend to extend, possibly reflecting a transformation to the normative distance at which we experience scenes. We observe boundary transformations even during minimal memory load, providing evidence that any normalization may occur even during perception. However, future work is needed to precisely quantify the typical viewing distances of scenes and systematically manipulate a range of scene properties to test the normalization hypothesis. It is also possible that processes during perception, such as attentional capture by salient aspects of a scene, may contribute to these boundary transformations.

Although boundary extension has been described as a phenomenon innate to scene memory, we find that more typically scene-like images result in boundary contraction. Images focused on few, central, close objects elicit boundary extension, although images containing several, dispersed, distant objects cause contraction. These high correlations between image-based metrics and boundary transformation imply that previous neuropsychological studies of boundary extension identifying a role for scene-selective cortex [9] and hippocampus [10, 11, 25] in scene memory extrapolation may instead be measuring a metric correlated with boundary transformations, like scene distance [21], size, or clutter [26]. Further, these correlations imply that boundary transformation is highly predictable. Rather than boundary transformations solely reflecting the outputs of an observer's cognitive processes, much of the effect may be determined by the inputs or the stimuli. Other work has shown that memory is highly influenced by the stimulus; e.g., images have an intrinsic memorability generalizable across observers [27, 28]. In the current study, we controlled for memorability of the 60-image set [17] and found no difference in boundary transformation score between memorable and forgettable scene exemplars (independent samples t test: $p = 0.359$), suggesting that the factors that influence boundary transformations and memorability may differ. Future work will need to account for these stimulus effects (memorability and tendency to distort boundaries) to fully understand the cognitive processes of the observer.

Although some work has reported boundary restriction (similar to the boundary contraction that we report), this was only during highly specific circumstances. For example, Intraub and colleagues [29] found that observers contracted their memory for wide-angled, object-oriented scenes after a 48-h delay but experienced only boundary extension at shorter timescales. In some other studies, boundary contraction has been observed, but not focused on; for example, Chadwick and colleagues [11] reported a significant proportion of trials showed boundary extension, but close to 15% also displayed boundary contraction. Otherwise, boundary restriction has only been reported in memory for emotional images [30, 31] and auditory stimuli [32]. The current study is the first work to show boundary contraction is as immediate, automatic, and widespread as boundary extension.

Finally, these results highlight the importance of continuously revisiting assumptions in the psychological sciences and underscore the necessity to always consider broad, representative samples when making global inferences about the brain. The key example images of boundary extension continue to be used in the literature 30 years from discovery [1, 24], and the numerous replications of the effect could in part be attributed to the recycling of narrow stimulus sets. In fact, when we conducted our original drawing study, we were surprised to observe only a meager tendency toward boundary extension [17], given the prominence of the effect in the literature. Importantly, it was the diversity of the images in this study that allowed us to realize that boundary extension only works with a limited view of what constitutes a "scene image." Sampling a spread of images that cause extension, contraction, and no boundary transformation will be necessary to understanding how we form mental representations of naturalistic images. As technology continues to provide access to larger image sets and broader populations, it is essential that we regularly re-examine psychological phenomena taken as fact.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - 60-Image Set Experiments: Drawing and RSVP
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Split-Half Consistency Analyses
  - Object Map Analyses
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cub.2019.12.004.

ACKNOWLEDGMENTS

We would like to thank Susan Wardle, Edward Silson, and Adrian Gilmore for their advice on the project and comments on the manuscript. This research was supported by the Intramural Research Program of the NIH (ZIA-MH-002909), under National Institute of Mental Health Clinical Study Protocol 93-M-1070 (NCT00001360).

AUTHOR CONTRIBUTIONS

W.A.B. and C.I.B. conceived and planned the experiments. W.A.B. conducted the experiments and analyzed the data. W.A.B. and C.I.B. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 12, 2019
Revised: November 5, 2019
Accepted: December 2, 2019
Published: January 23, 2020

REFERENCES

1. Intraub, H., and Richardson, M. (1989). Wide-angle memories of close-up scenes. J. Exp. Psychol. Learn. Mem. Cogn. *15*, 179–187.
2. Hubbard, T.L., Hutchison, J.L., and Courtney, J.R. (2010). Boundary extension: findings and theories. Q J Exp Psychol (Hove) *63*, 1467–1494.
3. Gagnier, K.M., and Intraub, H. (2012). When less is more: line-drawings lead to greater boundary extension than color photographs. Vis. Cogn. *20*, 815–824.
4. Intraub, H., and Dickinson, C.A. (2008). False memory 1/20th of a second later: what the early onset of boundary extension reveals about perception. Psychol. Sci. *19*, 1007–1014.
5. Candel, I., Merckelbach, H., Houben, K., and Vandyck, I. (2004). How children remember neutral and emotional pictures: boundary extension in children's scene memories. Am. J. Psychol. *117*, 249–257.
6. Chapman, P., Ropar, D., Mitchell, P., and Ackroyd, K. (2005). Understanding boundary extension: normalization and extension errors in picture memory among adults and boys with and without Asperger's syndrome. Vis. Cogn. *12*, 1265–1290.
7. Seamon, J.G., Schlegel, S.E., Hiester, P.M., Landau, S.M., and Blumenthal, B.F. (2002). Misremembering pictured objects: people of all ages demonstrate the boundary extension illusion. Am. J. Psychol. *115*, 151–167.
8. Intraub, H. (2012). Rethinking visual scene perception. Wiley Interdiscip. Rev. Cogn. Sci. *3*, 117–127.
9. Park, S., Intraub, H., Yi, D.-J., Widders, D., and Chun, M.M. (2007). Beyond the edges of a view: boundary extension in human scene-selective visual cortex. Neuron *54*, 335–342.
10. Mullally, S.L., Intraub, H., and Maguire, E.A. (2012). Attenuated boundary extension produces a paradoxical memory advantage in amnesic patients. Curr. Biol. *22*, 261–268.
11. Chadwick, M.J., Mullally, S.L., and Maguire, E.A. (2013). The hippocampus extrapolates beyond the view in scenes: an fMRI study of boundary extension. Cortex *49*, 2067–2079.
12. Goldstein, E.B. (2009). Sensation and Perception, Eighth Edition (Wadsworth CENGAGE Learning), p. 129.
13. Matlin, M.W., and Farmer, T.A. (2019). Cognition, Tenth Edition (John Wiley & Sons, Inc), pp. 176–178.
14. Eysenck, M.W., and Brysbaert, M. (2018). Fundamentals of Cognition, Third Edition (Routledge).
15. Educational Testing Service (2017). GRE Psychology Test Practice Book (Educational Testing Service).
16. Intraub, H., and Bodamer, J.L. (1993). Boundary extension: fundamental aspect of pictorial representation or encoding artifact? J. Exp. Psychol. Learn. Mem. Cogn. *19*, 1387–1397.
17. Bainbridge, W.A., Hall, E.H., and Baker, C.I. (2019). Drawings of real-world scenes during free recall reveal detailed object and spatial information in memory. Nat. Commun. *10*, 5.
18. Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J., Krasin, I., Pont-Tuset, J., Kamali, S., Popov, S., Malloci, M., Duerig, T., et al. (2018). The Open Images Dataset V4: unified image classification, object detection, and visual relationship detection at scale. arXiv,. arXiv:1811.00982v1. https://arxiv.org/abs/1811.00982.
19. Xiao, J., Hays, J., Ehinger, K.A., Oliva, A., and Torralba, A. (2010). SUN database: large-scale scene recognition from abbey to zoo. (IEEE) 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.
20. Benenson, R., Popov, S., and Ferrari, V. (2019). Large-scale interactive object segmentation with human annotators. arXiv. arXiv:1903.10830v2. https://arxiv.org/abs/1903.10830.
21. Lescroart, M.D., Stansbury, D.E., and Gallant, J.L. (2015). Fourier power, subjective distance, and object categories all provide plausible models of BOLD responses in scene-selective visual areas. Front. Comput. Neurosci. *9*, 135.
22. Bartlett, F. (1932). Remembering: A Study of Experimental and Social Psychology (Cambridge University Press).
23. Konkle, T., and Oliva, A. (2007). Normative representation of objects: evidence for an ecological bias in object perception and memory. Proc. Ann. Meet. Cogn. Sci. Soc. *29*, 407–412.
24. Intraub, H., Gottesman, C.V., Willey, E.W., and Zuk, I.J. (1996). Boundary extension for briefly glimpsed photographs: do common perceptual processes result in unexpected memory distortions? J. Mem. Lang. *35*, 118–134.
25. Maguire, E.A., Intraub, H., and Mullally, S.L. (2016). Scenes, spaces, and memory traces: what does the hippocampus do? Neuroscientist *22*, 432–439.
26. Park, S., Konkle, T., and Oliva, A. (2015). Parametric coding of the size and clutter of natural scenes in the human brain. Cereb. Cortex *25*, 1792–1805.
27. Bainbridge, W.A., Isola, P., and Oliva, A. (2013). The intrinsic memorability of face photographs. J. Exp. Psychol. Gen. *142*, 1323–1334.
28. Bainbridge, W.A. (2019). Memorability: how what we see influences what we remember. In Psychology of Learning and Motivation: Knowledge and Vision, K.D. Federmeier, and D.M. Beck, eds. (Elsevier Academic), pp. 1–24.
29. Intraub, H., Bender, R.S., and Mangels, J.A. (1992). Looking at pictures but remembering scenes. J. Exp. Psychol. Learn. Mem. Cogn. *18*, 180–191.
30. Safer, M.A., Christianson, S.-Å., Autry, M.W., and Österlund, K. (1998). Tunnel memory for traumatic events. Appl. Cogn. Psychol. *12*, 99–117.
31. Takarangi, M.K.T., Oulton, J.M., Green, D.M., and Strange, D. (2015). Boundary restriction for negative emotional images is an example of memory amplification. Clin. Psychol. Sci. *4*, 82–95.
32. Hutchison, J.L., Hubbard, T.L., Ferrandino, B., Brigante, R., Wright, J.M., and Rypma, B. (2012). Auditory memory distortion for spoken prose. J. Exp. Psychol. Learn. Mem. Cogn. *38*, 1469–1489.
33. Stoet, G. (2010). PsyToolkit: a software package for programming psychological experiments using Linux. Behav. Res. Methods *42*, 1096–1104.
34. Isola, P., Xiao, J., Torralba, A., and Oliva, A. (2011). What makes an image memorable? CVPR 2011 (IEEE), pp. 145–152.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Google Open Images Database V5	[18]	https://opensource.google.com/projects/open-images-dataset
Scene Understanding (SUN) Database	[19]	https://groups.csail.mit.edu/vision/SUN/
Drawings of real-world scenes during free recall reveal object and spatial information in memory	[17]	https://dataverse.harvard.edu/dataverse/drawingrecall
Boundary transformation scores for 1,000 images	This study	https://osf.io/28nzt/
Software and Algorithms		
MATLAB 2016a	MathWorks	RRID:SCR_001622
Amazon Mechanical Turk	Amazon	RRID:SCR_012854
Psytoolkit	[33]	https://www.psytoolkit.org/

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Wilma Bainbridge (Current e-mail: wilma@uchicago.edu). This study did not generate new unique reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Thirty healthy adults participated in the Memory Drawing Experiment, (21 female, average age 23.6 years, SD = 2.77) and 24 separate adults participated in the Perceptual Drawing Experiment (14 female, average age 24 years, SD = 4.14). In-lab participants were consented following the rules of the National Institutes of Health (NIH) Institutional Review Board (NCT00001360). 301 Amazon Mechanical Turk (AMT) workers participated in the Boundary Extension Scoring, 100 workers participated in the 60-Image RSVP Experiment, 2,000 workers participated in the 1000-Image RSVP Experiments, and 152 workers participated in the Distance Rating Experiment. Workers did not indicate age/gender. AMT participants acknowledged participation following guidelines set by the NIH Office of Human Subjects Research Protections (OHSRP). Twenty-three participants from the 2,000 workers in the 1000-Image RSVP Experiments were removed for not responding, or responding too quickly on all trials (< 200 ms). All participants were compensated for their time.

METHOD DETAILS

The experiments reported in the current study are a mix of re-analyses of findings from [17] and new experiments. 60-Image Set drawing data (for both memory drawings and copying drawings) and boundary transformation ratings were collected in [17]. RSVP scores for the 60-Image Set and all 1000-Image Set data (2 experiments with independent sets of 1,000 participants) were collected in the current study.

60-Image Set Experiments: Drawing and RSVP

Stimuli

60 images from 30 scene categories (2 images each) were taken from the SUN Database [19]. Images were controlled to be counterbalanced for image memorability [26], with one image per category of high memorability and one of low memorability, gathered from the image memorability dataset in [34]. High and low memorability were defined as images at least 0.5 SD away from the mean in the overall image distribution. Scene categories were chosen to be as maximally different as possible, and were chosen with a mix of indoor (e.g., kitchen, conference room), natural outdoor (e.g., mountain, garden), and manmade outdoor categories (e.g., city street, house). Images were resized to 512 pixels and shown at approximately 14 degrees of visual angle for the drawing tasks. Images were presented at 350 pixels for the online RSVP task.

Memory Drawing Experiment

This experiment was originally conducted in [17]. Thirty in-lab participants studied 30 images on a computer screen (1 per scene category) for 10 s each with an interstimulus interval (ISI) of 500 ms. The 30 images were pseudo-randomly selected to be half memorable and half forgettable images, counterbalanced across participants. Participants were instructed to study the images in as much detail as possible, but did not know how they would be tested. Participants then completed a 11-minute difficult digit span memory task where they had to study and then verbally recall 42 digit series ranging from 3-9 digits in length. Finally, participants were given 30 blank squares on paper at the same sizes as the originally presented images and were instructed to draw the studied images in as much detail as possible, in any order. Participants were given a black pen and colored pencils, and were allowed to label items that were difficult to draw. Participants were given as much time as needed to complete their drawings. Finally, participants completed an object-based cued recall task and a recognition task for these images; these results are not used in the current study. In spite of the 11-minute delay and 30 studied images, participants' memory drawings contained large amounts of object and spatial detail [17].

Copying Drawing Experiment

This experiment was originally conducted in [17]. Twenty-four in-lab participants (separate from the Memory Drawing Experiment participants) viewed 30 images (1 per scene category) presented on a computer screen and were told to draw each image on a blank square (at the same size as the original image) on a sheet of paper in as much detail as possible. Images were pseudo-randomly selected to be half memorable and half forgettable images, counterbalanced across participants. Participants drew one image at a time, and when they said they were finished with a drawing, they continued onto the next image. They were given as much time as needed to complete their drawings, and the image remained visible on the computer screen for the duration of their drawing time. Participants were given the option to draw additional images (up to 35 total) if there was additional time.

Boundary Transformation Scoring

This experiment was originally conducted in [17]. 301 AMT participants judged the level of boundary transformation of the drawings resulting from the Memory Drawing Experiment and the Copying Drawing Experiment. Participants saw a drawing and its corresponding image and were asked to decide whether "the drawing is closer, the same, or farther than the original photograph." They were told to ignore any extra or missing objects in the drawing. Participants responded on a 5-item scale (as in [10]) consisting of: much closer, slightly closer, the same distance, slightly farther, and much farther. Participants could also indicate "can't tell" if they weren't able to judge (e.g., if the drawing was incomprehensible). Seven AMT participants responded for each drawing, and mean boundary extension rating was calculated for each drawing. Scores were then transformed to a scale of -1 (much farther = extension) to +1 (much closer = contraction).

60-Image RSVP Experiment

This experiment was conducted for the current study to see whether separate, recognition-based metrics of boundary transformation replicate those of the drawing experiments. 100 AMT workers participated in an experiment using a RSVP paradigm commonly used to replicate boundary extension effects [4, 10] on online psychophysics experiment platform Psytoolkit [33]. For each trial, workers saw an image presented for 250 ms, followed by a 250 ms dynamic mask (5 mosaic-scrambled images unrelated to the original image presented for 50 ms each, at the same size as the original image), and then the original image was shown again. After 1 s, a response screen came on asking if the second image was closer (scored as -1, or boundary extension) or farther (scored as +1, or boundary contraction) than the first image. While participants believed the first and second images would have different levels of zoom, in reality the first image was always the same as the second image. Participants were given up to 3 s to respond, and then the next trial started after a 1 s fixation. Each participant saw all 60 images, and the experiment took approximately 4-5 minutes overall. Each image thus ultimately received 100 ratings.

1000-Image RSVP Experiments

The two experiments reported here are specific to the current study (i.e., separate from the data from [17]).

Stimuli

A set of 1000 images were assembled to reflect a representative sampling of naturalistic photographs. 500 of the images were downloaded from object-oriented image set Google Open Images (GOI) [18], a dataset containing 1.7 million images from Flickr of 600 different object classes (e.g., apple, toy). The other 500 images were downloaded from scene-oriented image set Scene Understanding (SUN) Database [19], a dataset containing 131,000 images from the Internet of 908 different scene categories (e.g., amusement park, restaurant). Images were selected that were at least 350 × 350 pixels in size, and only images provided with human-annotated object outlines [20] were used. Low quality images, blurry images, artwork, clearly altered images, advertisements, watermarked images, and images with graphic material were all removed. Of the remaining images, one image was randomly sampled per category (or two if all categories were exhausted). Note that some images also included people. This results in an image set that should be generally representative of our experiences with images, including both images focused on objects, as well as images focused on scenes. All images were center-cropped and resized to 350 × 350 pixels.

1000-Image RSVP Experiment 1

An almost identical experimental paradigm for the 1000-Image Set was used to the 60-Image RSVP Experiment and to previous boundary extension experiments [4, 10]. Timing was identical, but each participant made judgments of whether images were closer or farther (2 response options) for a random 100 images rather than 60 images, resulting in a total experimental time of 5-7 minutes. 1,000 people participated in this experiment, and each image thus received 100 ratings.

1000-Image RSVP Experiment 2

A second experiment was conducted identical to 1000-Image RSVP Experiment 1, except that participants had three response options: closer, same distance, or farther. This allowed participants to report that they experienced no difference between presentations of the identical images. Each image received an additional 100 ratings from a new set of 1,000 participants.

Distance Rating Experiment

An AMT experiment was conducted to collect subjective ratings of the distance from the observer to the most salient object in each image, to approximate the depth of the image. Five ratings were collected per image (n = 152 workers total), and workers were allowed to rate as many images as they liked. For each image, they were asked to judge "How far away is the main object/thing in this picture''? on a 5-point scale of: 1) extreme closeup, ~0-2 ft; 2) arm's length, ~3-4 ft; 3) nearby / same room, < 20 ft; 4) short walk, < 100ft; 5) far away, > 100 ft). This scale was adapted from [21]. Distance per image was taken as the average subjective rating across participants.

QUANTIFICATION AND STATISTICAL ANALYSIS

Split-Half Consistency Analyses

All analyses were conducted using MATLAB R2016a. Split-half consistency analyses were used to examine whether boundary transformation ratings were consistent across participants for individual images. Across 1000 iterations, participants were split into two random halves and sorted in the same image order. Mean boundary transformation score was then calculated for each image across each participant half. The two halves were then Spearman rank correlated to estimate the degree to which participant ratings agreed. Chance is calculated as the Spearman rank correlation between one half and the other half with a shuffled image order. Consistency ρ^* is measured as the average rank correlation across the 1000 iterations corrected with the Spearman-Brown prediction formula, to give an estimate of consistency across the whole set of participants. P value is calculated as proportion of trials in which the chance correlation was higher than the split-half correlation.

A split-half binning consistency analysis was also conducted to see whether images were consistently binned as boundary extending or contracting. Across 1000 iterations, participants were split into two random halves. Across one participant half, the images were sorted into boundary contracting (average boundary transformation score < 0) and boundary extending images (average boundary transformation score > 0). The percentage categorization agreement was then calculated with the other half of participants; resulting in a measure of what percentage of images are sorted into the same category by the other half.

Object Map Analyses

All images used in the 1000-Image RSVP Experiments include human-made annotations of the outlines of the individual objects in the images [19, 20]. To create average object maps, we assembled the top 200 boundary extending (or boundary contracting) images, and then calculated at each pixel the proportion of those images that contained an object at that pixel. ''Objects'' named ground, sky, ceiling, wall, floor, background, and grass were not included in the analysis. Using these object outlines, we were also able to count number of objects per image, and calculate other image properties. We computed average object area by taking the average number of pixels contained within each object outline for a given image. Object centricity was calculated as average 2-dimensional Euclidean distance between each object center of mass and the center of the image. Finally, average subjective distance from the observer to the most salient object in the image was collected from the Distance Rating Experiment.

DATA AND CODE AVAILABILITY

The data generated during this study (images, drawings, boundary transformation scores, and image properties) will be made publicly available with a link from the Lead Contact's website as well as a repository on the Open Science Framework (https://osf.io/28nzt/).