



Larger images are better remembered during naturalistic encoding

Shaimaa Masarwa^{a,b}, Olga Kreichman^{a,b}, and Sharon Gilaie-Dotan^{a,b,c,1} 

^aSchool of Optometry and Vision Science, Faculty of Life Science, Bar Ilan University, Ramat Gan 5290002, Israel; ^bThe Gonda Multidisciplinary Brain Research Center, Bar Ilan University, Ramat Gan 5290002, Israel; and ^cInstitute of Cognitive Neuroscience, University College London, London WC1N 3AZ, United Kingdom

Edited by Morris Moscovitch, Department of Psychology, University of Toronto, Toronto, ON, Canada; received November 1, 2021; accepted December 3, 2021 by Editorial Board Member Michael S. Gazzaniga

We are constantly exposed to multiple visual scenes, and while freely viewing them without an intentional effort to memorize or encode them, only some are remembered. It has been suggested that image memory is influenced by multiple factors, such as depth of processing, familiarity, and visual category. However, this is typically investigated when people are instructed to perform a task (e.g., remember or make some judgment about the images), which may modulate processing at multiple levels and thus, may not generalize to naturalistic visual behavior. Visual memory is assumed to rely on high-level visual perception that shows a level of size invariance and therefore is not assumed to be highly dependent on image size. Here, we reasoned that during naturalistic vision, free of task-related modulations, bigger images stimulate more visual system processing resources (from retina to cortex) and would, therefore, be better remembered. In an extensive set of seven experiments, naïve participants ($n = 182$) were asked to freely view presented images (sized 3° to 24°) without any instructed encoding task. Afterward, they were given a surprise recognition test (midsized images, 50% already seen). Larger images were remembered better than smaller ones across all experiments ($\sim 20\%$ higher accuracy or ~ 1.5 times better). Memory was proportional to image size, faces were better remembered, and outdoors the least. Results were robust even when controlling for image set, presentation order, screen resolution, image scaling at test, or the amount of information. While multiple factors affect image memory, our results suggest that low- to high-level processes may all contribute to image memory.

vision | memory | images | size | naturalistic

We are constantly exposed to many images, and despite not making intentional efforts, some of these images are burned into memory, while others are not. It is yet unclear what determines what we remember or not under such naturalistic (unintentional, not externally dictated) conditions. Multiple factors are suggested to contribute to memory, such as 1) the “level of processing” of a stimulus (in the visual domain, for example, faces are less remembered when attending to gender [“shallower”] than when attending to honesty or likeableness [“deeper”] (1, 2)), and in the linguistic–lexical domain words are less remembered when attending to fonts [shallower] than when attending to their semantic [deeper] aspects [e.g., refs. 3 and 4]); 2) familiarity (e.g., ref. 5); and 3) visual category (e.g., ref. 6). However, most visual memory investigations involve predetermined experimental direct or indirect memory-encoding tasks (e.g., refs. 7–10). These predetermined experimental tasks may suppress or facilitate visual processing of the stimulus at multiple levels, influencing even early stages as V1 and thalamic processing (11–13), and examples of such influences include enhancement of spatial resolution or contrast sensitivity (14, 15) and even alterations of receptive field properties (16–20) at the attended (task-related) location. Since instructed task can modulate visual processing, attention, eye movements, and working memory (16, 21–24), it is unclear if one can generalize the

findings about visual memory during instructed behavior to visual memory during naturalistic everyday encoding.

A key principle guiding visual processing at early to intermediate stages is that bigger images are processed by bigger subparts of the visual system, and this is true from as early as the retina to retinotopic cortex (25, 26). Since bigger images entail more visual system processing resources at least in the early to intermediate levels of processing, here we reasoned that when no instructed task that may modulate processing is involved and despite certain levels of invariance to size at higher-level areas (27–29), the enhanced processing of bigger images by the visual system will support stronger registration into memory than that of smaller images that are processed by fewer resources (Fig. 1A).

Results

We tested this in an extensive set of seven experiments ($n = 182$) where participants naturalistically viewed images of various sizes while freely viewing them without any instructed encoding or knowledge about the memory test that would follow. Each participant underwent passive viewing of images of different sizes ($3^\circ \times 3^\circ$ to $24^\circ \times 24^\circ$, exposure phase) (Fig. 1B) while being asked to freely view the images presented without being informed of any memory-related task that would follow. This exposure phase was followed by a surprise recognition test phase where

Significance

It is unclear what makes some of the numerous visual scenes we encounter every day memorable (while others are not) when we make no intentional effort to memorize them. Here, we reasoned that although visual perception is somewhat size invariant (e.g., we can recognize a person from multiple distances), visual memory would depend on image size. Across experiments, where participants freely viewed images without any memory- or nonmemory-related task (similar to naturalistic visual behavior), larger images were remembered better than smaller ones (about 1.5 times better), and this effect was proportional to image size. Our study indicates that physical stimulus dimensions (as the size of an image) influence memory, and this may have significant implications to learning, aging, development, etc.

Author contributions: S.G.-D. conceived the study; S.M., O.K., and S.G.-D. designed research; S.M. and O.K. performed research; S.M., O.K., and S.G.-D. analyzed data; O.K. and S.G.-D. wrote the paper; and all authors approved the final version of the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. M.M. is a guest editor invited by the Editorial Board.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹To whom correspondence may be addressed. Email: shagido@gmail.com.

This article contains supporting information online at <http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2119614119/-DCSupplemental>.

Published January 19, 2022.

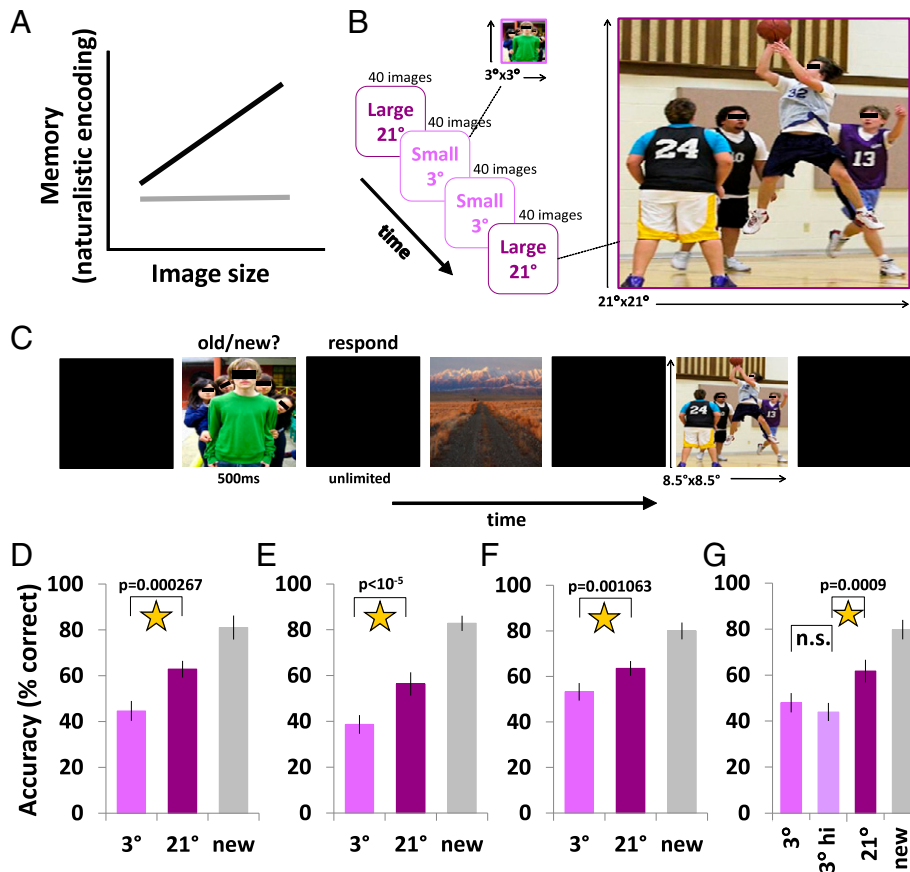


Fig. 1. Larger images were remembered better than smaller ones during naturalistic encoding (Experiments 1 to 4). (A) Hypothesized dependency of image memory during naturalistic encoding on image size: positive monotonic size dependency in black (larger images are better remembered than smaller images) or size invariance in gray (memory is not affected by image size). (B) In Experiments 1 to 4 during the exposure phase, participants viewed 80 small (3° × 3°; lighter purple) and 80 large (21° × 21°; darker purple) images (2 s each, small and large matched in contents) presented in four blocks of 40 images each. Participants were only instructed to freely view the images without being given any memory or encoding instructions (more details are in the text and in *Materials and Methods*). Note that the block order depicted here was changed in Experiment 3 to small–large–small. (C) In Experiments 1 to 4, the test phase that followed the exposure phase included 320 mid-sized images (~8° × 8°, 500 ms each, 160 presented in the exposure), and participants were given a surprise memory test where they were asked to report whether they recalled seeing each image or not. (D–G) Test phase accuracy on the surprise old/new recognition task (on the y axis) by condition (x axis). Results reveal that across all experiments, images presented in larger format (21°, darker purple) during the exposure phase were better remembered than those presented in smaller format (3°, lighter purple), and this was true in Experiment 1 (D; $n = 17$, $P = 0.000267$); when we controlled for image sets in Experiment 2 (E; $n = 16$, $P < 10^{-5}$); when we changed block order to control for recency or primacy effects (32–38) in Experiment 3 (F; $n = 17$, $P = 0.001063$); or when we controlled for reduced resolution, world size, or viewing distance in Experiment 4 (G; $n = 16$, $P = 0.0009$); n.s.: not significant. 3°hi represents the small(highRes) condition in another shade of light purple, new images condition in gray. Error bars represent SEM. * denotes $P < 0.005$.

participants were asked to judge for each image presented whether they recalled seeing it earlier or not (old/new recognition task); in all experiments, images were mid-sized (8° × 8°, except for Experiment 6) (*Materials and Methods*), and midsize was determined by taking into account the cortical magnification factor (30, 31) (Fig. 1C and *Materials and Methods*). In Experiment 1, participants ($n = 17$) were exposed to 80 small (3° × 3°) and 80 large (21° × 21°) colored images from multiple categories that appeared in four blocks (large–small–small–large; each image was presented for 2 s, small and large images were matched in contents) (Fig. 1B) and were later asked to make old/new memory judgements on a set of 320 mid-sized images (8° × 8°, 160 old; the old and new were matched in contents and randomly ordered; 500-ms exposure) (*Materials and Methods*). Accuracy for larger (21°) images was significantly higher than for smaller (3°) images [small: $44.6 \pm 4.4\%$ (SEM), large: $62.9 \pm 3.7\%$ (SEM), small vs. large $P = 0.000267$, $t(15) = 4.73$, paired two tailed] (Fig. 1D and *SI Appendix, Tables S1 and S2*). To rule out the possibility that the results may have been driven by specific images, a new group of participants ($n = 16$) ran the same experiment but with

the small and large image sets swapped (Experiment 2) (*Materials and Methods*). Accuracy for larger images was again significantly higher than for smaller ones, indicating that Experiment 1's results were not driven by specific images [small: $38.7 \pm 4.1\%$ (SEM), large: $56.4 \pm 5.1\%$ (SEM), $P < 10^{-5}$, $t(14) = 6.84$, paired two tailed] (Fig. 1E and *SI Appendix, Tables S1 and S2*). To rule out the possibility that the results of Experiments 1 and 2 were driven by recency or primacy effects (32–38), as the larger images were presented in the first and last blocks, a new group of participants ($n = 17$) underwent the same experiment but with swapping of the block order (small–large–large–small, Experiment 3) (*Materials and Methods*). Here too, we found that accuracy was higher for the larger images regardless of block order [small: $53.2 \pm 3.9\%$ (SEM), large: $63.5 \pm 3.3\%$ (SEM), $P = 0.001063$, $t(15) = 4.043$, paired two tailed] (Fig. 1F and *SI Appendix, Tables S1 and S2*). To examine whether the results were driven by image world size or retinal size and to test whether screen resolution may have degraded the information conveyed in the smaller images, in Experiment 4 ($n = 16$), on top of the large (21°) and smaller image (3°) conditions from Experiments 1 to 3 (same

viewing distance [60 cm] and spatial resolution), we added a new condition with the same world size as the large close images but the same retinal size (3°) as the small close images. This was achieved by displaying images of the same physical size as the large close images but from a different (further) viewing distance (4 m). This new condition occupied the same small visual angle as smaller close images ($3^\circ \times 3^\circ$) but had much higher ($\times 6.6^2$) spatial resolution than them (*Materials and Methods*). We found (Fig. 1G) that retinal size had influence on visual memory but that world size did not, as when retinal size changed but world size was kept constant [large 21° (from 60 cm) vs. small(highRes) 3° (from 4 m)], images with bigger retinal size (21°) were significantly better remembered [large $61.7 \pm 4.9\%$ (SEM), small(-highRes) $43.9 \pm 3.9\%$ (SEM), $P = 0.0009$] (*SI Appendix, Tables S1 and S2*). On the other hand, when world size changed (eight-fold difference in height and width) and viewing size changed (60 cm vs. 4 m) but retinal size was kept constant [small (3° from 60 cm) vs. small(highRes) (3° from 4 m)], no difference in memory was found [small $47.9 \pm 4.2\%$ (SEM), small(highRes) $43.9 \pm 3.9\%$ (SEM), $P > 0.4$] (*SI Appendix, Table S2*), which also indicated that limited screen resolution was not likely to be a main factor influencing lower memory for smaller images.

To parametrically investigate the effect of image size on memory during naturalistic encoding, we built Experiment 5, a new four sizes (height and width of 3° , 6° , 12° , and 24°) \times four visual categories (faces, people, indoors, and outdoors) experiment (Fig. 2A) that also allowed us to assess the effect of visual categories on memory. We used images with predefined memorability scores [the LaMem Dataset (6)] such that image sets of each experimental size had equal memorability scores and equal contribution of each visual category (Fig. 2A and *SI Appendix, Fig. S1A*). Furthermore, we also checked that images in the different experimental sizes were well balanced for luminance levels (*SI Appendix, Fig. S1B*). Two new groups ($n_1 = 25$, $n_2 = 26$) underwent two versions of this experiment (image sets of size conditions were swapped across versions [3° with 24° , 6° with 12°]) (*Materials and Methods*). Since there was no effect of experimental version ($P > 0.64$) and no interaction between size and version ($P > 0.9$) (*SI Appendix, Table S2*), we collapsed the data across the two versions. We found a main effect of image size on memory [$F(3,150) = 57.31$, $P < 0.0001$] with $3^\circ < 6^\circ < 12^\circ$ (post hoc P values < 0.0001) (Fig. 2B and *SI Appendix, Table S2*), indicating

that in this range of image sizes (3° to 12°), image memory is size dependent. There was also a main effect of visual category on memory [$F(3,150) = 15.76$, $P < 0.0001$], with faces most remembered and outdoor scenes remembered the least (Fig. 2C and *SI Appendix, Table S2*).

We also wanted to control for the possibility that our results were due to less time spent looking at the smaller images relative to that spent looking at the larger images. Therefore, in Experiments 6 and 7 (see below) we monitored participants' eye movements throughout the experiments and excluded participants who looked at the smaller images less than 80% of their presentation time (Experiment 6: 10 of 30 excluded, Experiment 7: 12 of 35 excluded, see *SI Appendix, Fig. S2*) (*Materials and Methods*).

In Experiment 6, we further controlled for the possibility that our results were driven by the relative size differences (larger images in the exposure were scaled down in the test, while smaller images in the exposure were enlarged in the test) and not by the initial image size at exposure. Participants freely viewed small (3°) and large (24°) images during the exposure phase and were then tested on them either as in our previous experiments (medium-sized images) or in their original presentation size (as they appeared in exposure) (*Materials and Methods*). Both smaller images' and larger images' memory did not benefit from being tested with the same size as during exposure relative to being tested with medium-sized images (as in Experiments 1 to 5; one-way ANOVA, $n = 18$, Bonferroni/Dunn post hoc: 3° Sml to 3° Mid: $P = 0.0775$ [nonsignificant], 24° Mid to 24° Lrg: $P = 0.87$) (Fig. 3A; more details in *SI Appendix, Table S2*). We also replicated our main findings across experiments that during naturalistic encoding, bigger images are better remembered than smaller ones (one-way ANOVA, $n = 18$, Bonferroni/Dunn post hoc: 24° Mid vs. 3° Sml: $P < 0.0001$) (Fig. 3A and *SI Appendix, Tables S1 and S2*).

To control for the possibility that larger images are better remembered since they convey more information than smaller images, we ran Experiment 7, where we directly compared memory for smaller and larger images that contained the same amount of information. Larger blurred images were directly enlarged from smaller sharp images and then blurred to eliminate pixelation and artificially added edges. This process created larger images with the same amount of information present in

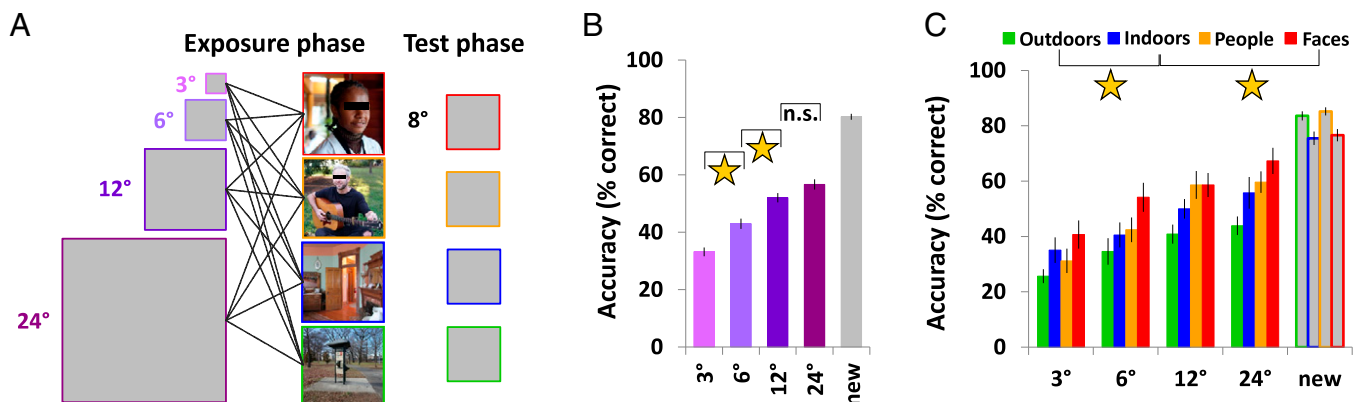


Fig. 2. Image memory during free viewing increases proportionally with image size for 3° to 12° images (Experiment 5). (A) Experimental design of Experiment 5. During exposure, participants ($n = 51$) viewed four blocks (random block and within-block image order), and each block included 40 uniformly sized images [$3^\circ \times 3^\circ$, $6^\circ \times 6^\circ$, $12^\circ \times 12^\circ$, or $24^\circ \times 24^\circ$ indicated by lighter to darker shades of purple, 2 s per image; all from the LaMem Dataset (6)] from four visual categories (faces [red], people [orange], indoors [blue], and outdoors [green]). In the test phase, 320 mid-sized images (160 old, $\sim 8^\circ \times 8^\circ$, 500 ms each) were presented. Experimental instructions were the same as in Experiments 1 to 4 (*Materials and Methods*). (B) Test phase accuracy on the surprise old/new recognition task (on the y axis) by condition (x axis). A significant effect of image size on image memory was found ($P < 0.0001$), with memory for 12° greater than memory for 6° , which was greater than memory for 3° (P values < 0.0001) (*SI Appendix, Tables S1 and S2*); n.s.: not significant. (C) The same as in B but with finer category-related results. A significant effect of visual category was found ($P < 0.0001$), with faces and people best remembered and outdoors the least remembered (*SI Appendix, Tables S1 and S2*). Error bars represent SEM. * denotes $P < 0.005$.

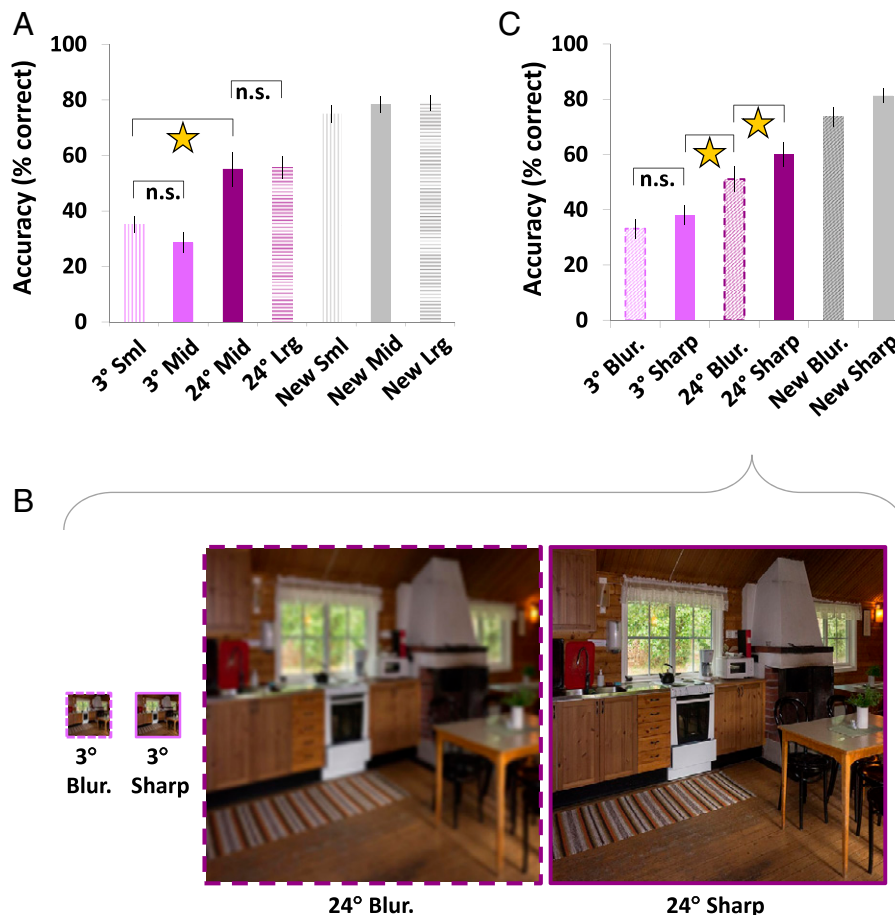


Fig. 3. Effects of test-exposure relative scaling and level of information on memory. (A) Experiment 6 ($n = 18$) results. Test phase accuracy on the surprise old/new recognition task (on the y axis) by condition (x axis) when investigating test-exposure relative image size scaling effects on memory. At exposure, images were presented in two sizes (3° , 24°). During the test phase, images were presented either at the same size as in the exposure (original size, no scaling) or mid-sized (as in Experiments 1 to 5, with scaling relative to exposure). 3° Mid and 24° Mid are the same as in Experiment 5 and are indicated by the same colors; 3° Sml and 24° Lrg are new conditions indicated by striped bars. We found no effect of relative image scaling ($P = 0.0912$), and critically smaller images at exposure were not remembered significantly better when tested with the same small image size (3° , $P = 0.0775$; n.s.: not significant; note that corrected significance threshold was at 0.0083). Smaller images that were tested as small (no scaling) were remembered significantly less than bigger images (e.g., 24° Mid, $P < 0.0001$). Replicating the results of Experiments 1 to 5, here too bigger images were better remembered ($P < 0.0001$) (Results and SI Appendix, Tables S1 and S2). (B) Demonstration of Experiment 7 exposure stimuli: four conditions (two sizes [3° , 24°] \times two sharpness levels [blurred, sharp]). Critical contrast was between memory of large blurred images relative to memory of small sharp images, differing in size but not in information content (Materials and Methods). The images (reproduced from ref. 39) are provided for demonstration purposes. (C) Experiment 7 ($n = 23$) surprise recognition test phase results. 3° Sharp and 24° Sharp are the same as in Experiment 5 and indicated by the same colors, and 3° Blur. (3° Blurred) and 24° Blur. (24° Blurred) are new conditions indicated by diagonal striped bars. We found a significant effect of image size ($P < 0.0001$) and of sharpness ($P = 0.0022$). Critically, memory of larger blurry images (24° Blur.) was better than that of sharp smaller images (3° Sharp), even though they both carried the same amount of image information ($P < 0.0001$), indicating that size, even without adding image information, contributes to image memory. Larger sharp images were also remembered better than blurry ones ($P = 0.0023$). n.s.: not significant. Error bars represent SEM. * denotes $P < 0.005$ (see Results).

them as the smaller sharp images (Materials and Methods). The experiment itself included four conditions presented at exposure (two sizes [3° , 24°] \times two sharpness levels [Blurred, Sharp]) (Fig. 3B), where the two additional conditions (smaller 3° Blurred and larger 24° Sharp images) were meant to mask the contrast of interest (smaller 3° Sharp vs. larger 24° Blurred) and enabled us to test replicability of the original finding (larger 24° Sharp vs. smaller 3° Sharp), the overall effect of size, and the effect of blurring. We hypothesized that the original finding from our earlier experiments would replicate and that we would find a main effect of size and a main effect of blurring. At test, images that were presented as blurred during exposure were presented as blurred, and images that were presented as sharp during exposure were presented as sharp. A one-way ANOVA ($n = 23$) on condition allowed us to test our main contrast of interest (larger 24° Blurred vs. smaller 3° Sharp) where we found that bigger images were remembered better than smaller ones even when they did

not contain more information (larger 24° Blurred: $51.09 \pm 4.77\%$ [SEM], smaller 3° Sharp: 37.93 ± 3.62 [SEM], one-way ANOVA Bonferroni/Dunn post hoc: 24° Blurred vs. 3° Sharp: $P < 0.0001$) (Fig. 3C and SI Appendix, Tables S1 and S2). This finding further substantiates the main finding from Experiments 1 to 6. We also found that sharp larger images were better remembered by 9% than blurred larger images (24° Blurred: $51.09 \pm 4.77\%$ [SEM], 24° Sharp: $60.11 \pm 4.5\%$ [SEM], one-way ANOVA Bonferroni/Dunn post hoc: $P < 0.0023$), which could indicate that details have an additional contribution to memory on top of image size. A two-way ANOVA allowed us to test the effect of size that was found significant ($P < 0.0001$) and the effect of sharpness that was found significant ($P = 0.0022$), with no interaction ($P = 0.28$) (details are in SI Appendix, Table S2). Furthermore, the results of Experiments 6 and 7 (given our exclusion criteria based on eye movement data) allowed us to conclude that our findings about the influence of image size on visual memory during naturalistic

encoding were not due to less time spent looking at the smaller images relative to that spent looking at the larger images.

In addition, across experiments we found that the highest accuracy in each experiment was to the new (unseen) images (Figs. 1 *D–G*, 2 *B* and *C*, and 3 *A* and *C* and *SI Appendix, Table S1*), and no priming effects [implicit or explicit (40)] were found (*SI Appendix, Tables S4* and *S5*).

Memorability Analysis. Given the data we obtained, we were able to examine for a specific image if it was better remembered when it was presented in bigger format vs. in smaller format. To that end, we first examined the images used in Experiments 1 and 2 ($n = 160$ images) where images presented in Experiment 1 (17 participants) as small (3° , $n = 80$ images) were presented in Experiment 2 (16 new participants) as big (21°) and vice versa. As can be seen in Fig. 4*A*, most images were better remembered when presented in larger format [mean accuracy difference of $18 \pm 1.5\%$ (SEM), $t(159) = 11.74$, $P < 10^{-22}$, $n = 160$ images] (Fig. 4*A*, *D*, and *E*). We then analyzed the data obtained in Experiment 5 in a similar manner (25 participants in version 1, 26 participants in version 2, swapping images across versions between 3° and 24° and between 6° and 12°). Here too (Fig. 4*B* and *C*), we found that the same image was on average better remembered when presented at 24° relative to 3° [mean accuracy

difference of $23.4 \pm 1.8\%$ (SEM), $t(79) = 13.0$, $P < 10^{-20}$, $n = 80$ images] (Fig. 4*B*, *D*, and *E*), and the same was found for 12° relative to 6° [mean accuracy difference of $9.04 \pm 1.6\%$ (SEM), $t(79) = 5.58$, $P < 10^{-6}$, $n = 80$ images] (Fig. 4*C–E* and *SI Appendix, Fig. S1*).

Discussion

Our results show that physical dimension as image size has significant effect on memory of images during naturalistic encoding. This challenges a simplistic view of visual memory inheriting properties of size-invariant high-level vision and suggests that information present at lower levels of the visual hierarchy (25) as well as within the high-level cortex itself (27–29) propagates downstream. Here, we did not employ the classic incidental memory paradigm involving incidental encoding of images (e.g., refs. 9 and 10) or any intentional encoding (e.g., refs. 7 and 8). Rather, similar to naturalistic visual behavior, there was no encoding task, and thus, our results may not be comparable with earlier incidental or intentional memory studies (e.g., refs. 8 and 41–43), even to those that have investigated the relation between image size and memory (e.g., refs. 44–48). In addition, our old/new memory task on one item at a time may have further contributed to the consistently lower memory performance we found

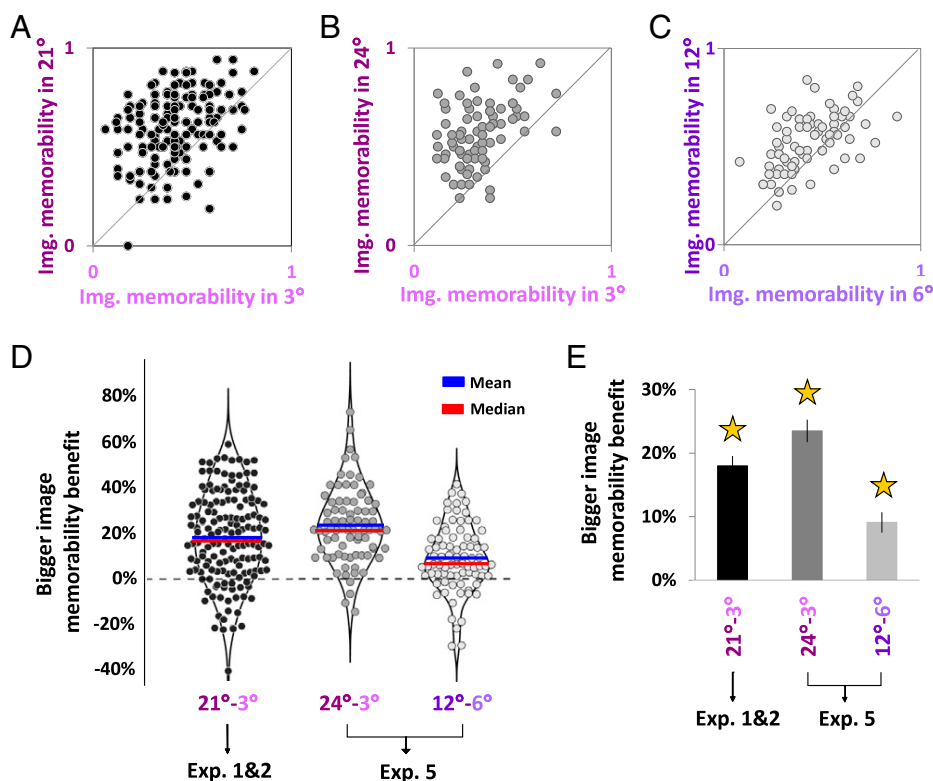


Fig. 4. Image memorability by image size reveals per-image higher image memorability when presented larger. (A) For each image ($n = 160$ images, each is indicated by a dot) its average correct recognition across participants (Experiment 1, $n = 17$; Experiment 2, $n = 16$) when it was presented in large format (21° in Experiments 1 and 2, on the y axis) relative to its average correct recognition across participants when it was presented in smaller format (3° in Experiments 2 and 1, on the x axis) is shown. Note that each image was presented in one experiment as big and in the other as small since the image sets swapped sizes between experiments. Note that most images fall above the diagonal, indicating that they were better remembered (higher memorability) when presented as larger (21°). (B) The same analysis as in A for images used in Experiment 5 in versions 1 and 2 ($n_1 = 25$, $n_2 = 26$). For each image ($n = 80$ images), its average correct recognition across participants (memorability) when it was presented as big (24° , on the y axis) is plotted relative to when it was presented as small (3° , on the x axis). Here again, most images fall above the diagonal. (C) The same analysis as in B for the images when they were presented as 12° (y axis) relative to when they were presented as 6° (x axis; $n = 80$ images). (D) For each image, its memorability benefit is the difference in its memorability when presented in larger format (21° , 24° , or 12°) relative to when presented in smaller size (3° or 6°); each dot represents the memorability benefit of larger presentation for each image. *D, Left* matches data from A; *D, Center* matches data from B; and *D, Right* matches data from C. For each plot, the mean is indicated in blue, and the median is in red. (E) Average memorability benefit of larger presentation size across images in each analysis (data matching the blue mean lines in D; error bars, SEM). *Significant differences in memorability of the larger format relative to the small format (all $P < 10^{-6}$, paired *t* tests) (*Materials and Methods* and *Results*).

across our experiments relative to earlier encoding-based findings (e.g., refs. 7, 8, and 49). In line with earlier studies (6, 50, 51), we found that face images were best remembered, and this was true across image sizes (limited to the categories and sizes we used), which could possibly be attributed to their social significance (52, 53). The size–memory apparently linear relation we found in the size range that we investigated (3° to 12°) (Fig. 2 B and C) may not generalize to bigger ($>24^\circ$) or smaller ($<3^\circ$) images. In fact, we hypothesize that there would be a sharp drop in performance for images smaller than 3° as well as for images that are “too big” to be perceived, as when viewing a movie from the first row in the cinema. Furthermore, in this study, we have investigated two-dimensional (2D) images that are present in a three-dimensional (3D) world in printed (e.g., journals, billboards) or electronic form. Our results (Experiment 4) indicate that the retinal size rather than the world size of the 2D images plays an important role in affecting visual memory. It is unclear if our results will generalize to real-world 3D objects that are different from 2D images in multiple dimensions (e.g., tangible, provide stereoscopic 3D and oculomotor information). Additionally, in a 3D environment, factors such as familiar size and size constancy may influence perception, perceptual judgements (54–56), and even visual memory in a different manner.

The image size effect on visual memory during naturalistic encoding may be attributed to multiple factors that may be modulated by image size as 1) larger expanse of visual system resources such as retinotopic cortex responding to and processing bigger images (57, 58), 2) possible different eye movement patterns for different image sizes, 3) different spatial frequency contents across the different sizes, 4) different spatial integration, and 5) differences in attention or saliency. While we found that the higher levels of detail present in larger images affected visual memory (Experiment 7) (Fig. 3 B and C and *SI Appendix, Fig. S3B*), we also found that size itself, even when not providing more information, was a contributing factor to memory. Boundary extension or contraction (59–62) that has been shown to be modulated by image properties (59) may also be modulated by image size and thus, may partially contribute to the size effect on visual memory that we found. While our results may not generalize to active intentional or incidental memory paradigms, our study does demonstrate that physical image dimension can affect image memory under conditions that closely mimic naturalistic daily visual behavior. However, it does not imply that other physical dimensions would have similar effects or that other nonphysical factors (e.g., cognitive) do not play a role in image memory.

Materials and Methods

Participants. A group of 182 participants took part in the study; each participated in only one experiment. Seventeen participated in Experiment 1 (12 women, aged 28.1 ± 6.9 y, 15 right handers), 16 participated in Experiment 2 (10 women, aged 25.4 ± 6.0 y, 15 right handers), 17 participated in Experiment 3 (11 women, aged 25.1 ± 5.7 y, all right handers), 16 participated in Experiment 4 (10 women, aged 25.7 ± 6.3 y, 15 right handers), 51 participated in Experiment 5 (25 in version 1, 26 in version 2, 31 women, aged 25.6 ± 6.6 y, 46 right handers), 30 participated in Experiment 6 (21 women, aged 22.8 ± 4.24 y, 28 right handers; 10 were excluded by the eye movement analysis criteria [*Eye Tracking Analyses*], 2 additional participants were excluded since they reported pressing the keys inconsistently), and 35 participated in Experiment 7 (28 women, aged 24 ± 4 y, 30 right handers; 12 were excluded by the eye movement analysis criteria) (*Eye Tracking Analyses*).

The experimental protocol was approved by the Bar Ilan University Ethics Committee. All the participants signed written informed consent before their participation. All participants had normal or corrected to normal far and near vision (all were checked for near and far visual acuity before the experiment began).

General Procedures. Experiments 1 to 3 and 5 to 7 were conducted on an Eizo FG2421 24" high-definition liquid crystal display monitor (HD LCD) with resolution of $1,920 \times 1,080$ pixels running at 100 Hz, and Experiment 4 was

conducted on an Asus VG248QE 24" monitor with $1,920 \times 1,080$ -pixel resolution running at 144 Hz. All experiments were run using an in-house developed platform for psychophysical and eye-tracking experiments [PSY (63–67)] developed by Yoram S. Bonneh (68) running on a Windows personal computer. All experiments were performed in a dark room. Viewing distance was 60 cm from the screen in all experiments and conditions except for the small–highRes condition in Experiment 4 (viewed from 4 m; details are in Experiment 4). Experiments 1 to 5 each took ~ 25 min, and Experiments 6 and 7 took ~ 30 min due to the eye tracking overhead (see below). ANOVA and post hoc statistical analyses were performed with StatView software 5.0, and Bonferroni/Dunn post hoc analyses were used. Eye movements were recorded during Experiments 6 and 7 (exposure and test) with an EyeLink infrared system (SR Research) equipped with a 35-mm lens. Head movements were limited by a chin rest, and a standard five-point calibration was performed before the exposure phase. Eye movements were recorded binocularly; only left eye data were analyzed. Analyses were based on data sampled at 250 Hz.

All seven experiments included two consecutive parts: 1) the exposure phase: passive viewing of images of different sizes (participants were asked to freely view and attend the images presented without being informed of any memory-related task that would follow) and 2) the test phase: an old/new surprise recognition task (participants were asked to report if they recalled seeing each image [old] or not [new] with no time limitation; no feedback was given). The background color of the screen across all experiments was always black.

Experiment 1. At the exposure phase, 160 images (colored photographs) of the study set were presented in four blocks (each block of 40 images) in the order [large–small–large] (Fig. 1B) (image order within each block was randomized). Small images subtended a visual angle of $3.15^\circ \times 3.15^\circ$, and large images subtended a visual angle of $20.78^\circ \times 20.78^\circ$. Images were presented in a sequence; each image was displayed for 2 s followed by a 500-ms black screen interstimulus interval, and no response was required. Participants were asked to freely view and attend the images (fixation was not superimposed on the images; no fixation was required). All images were taken from the internet and resized in MATLAB by the *imresize* function (bicubic interpolation) to equal width and height (800×800 pixels) to avoid within- and across-condition size differences. These uniformly sized images were then scaled to be displayed according to the experimental condition (small $\sim 3^\circ$ or large $\sim 21^\circ$; scaling details are given below). The images included different visual categories (faces, people, hands, animals, food, flowers, indoor places, outdoor places, and vehicles), and the images of each visual category were distributed equally between the small and large image sets. Image scaling (reduction in size) from the large source images to smaller sizes was done by the in-house software for psychophysical experiments (63–67). For the scaling factor less than one, the software took one representative pixel from the larger “source” image area to be placed in the new location in the smaller “target” image. The pixel to be written in the target location was the last in a top left to bottom right scanning of the region in the source image that was about to be scaled (approximately every six to seven \times six to seven pixels in the source image were converted to one pixel in the target smaller image in this experiment).

A test phase followed the exposure phase, where participants were required to perform an old/new surprise recognition memory task on 320 mid-sized images (visual angle of $8.39^\circ \times 8.39^\circ$) (information is in *Test Phase Image Size Reasoning* and scaling information in this section): 160 “old” (previously seen in the exposure phase) and 160 “new” images that were presented sequentially in random order. Each image appeared for 500 ms followed by a black screen until a response was provided; participants were required to report if they recalled seeing each image (old) or not (new) without time limitations, and no feedback was given.

Experiment 2: Controlling for Image Set. In this experiment, the experimental protocol was precisely the same as that of Experiment 1 except that the image sets of the small and large conditions from Experiment 1 were switched (images displayed as small in Experiment 1 were now displayed as big and vice versa).

Experiment 3: Controlling for Primacy or Recency Effects. The experimental protocol was precisely the same as in Experiment 1 with the only difference between the experiments being the block order in the exposure that was changed to small–large–large–small. This also meant that the same set of images that appeared in a small block in Experiment 1 appeared here in a small block and the same for the larger images.

Experiment 4: World Size vs. Retinal Size. In this experiment in the exposure phase, we included, as in Experiments 1 to 3, the small and large conditions

from a 60-cm viewing distance and in addition, a new condition of small images with higher resolution [“small(highRes)” presented large images but viewed from 4 m to occupy the same visual angle as the original small images condition but providing much higher ($\times 6.6^2$) spatial resolution].

Stimuli. The exposure phase included three conditions: 1) large (visual angle of $20.78^\circ \times 20.78^\circ$, viewing distance of 60 cm as in Experiments 1–3), 2) small (visual angle of $3.15^\circ \times 3.15^\circ$, viewing distance of 60 cm as in Experiments 1–3), and 3) small, higher resolution [small(highRes)], where large images were viewed from 4 m so that they occupied the same visual angle (same retinal image size) as the small condition ($3.15^\circ \times 3.15^\circ$). We divided the original image set used in the exposure phase (as in Experiments 1 to 3, 159 of the 160 images) to six subsets of similar size (three with 26 images and three with 27 images), where each of these subsets had similar types of images (faces, people, places, food, etc.) and was used in one of the experimental blocks in the exposure phase.

Procedure. The exposure phase took approximately 1 min longer than in Experiments 1 to 3 since participants were required to move to a distant seating position for the small(highRes) condition. Of the 16 participants who performed this experiment, in the exposure phase six participants underwent the small–small(highRes)–large–small–small(highRes)–large block order, six participants underwent the small(highRes)–large–small–small(highRes)–large–small block order, and four performed the large–small–small(highRes)–large–small–small(highRes) block order to counterbalance condition order and so that each condition type would appear in a first block and in a last block of one of the versions. The small and large conditions were viewed from 60 cm as in Experiments 1 to 3, and the small(highRes) was viewed from 4 m. In the test phase, 318 images (159 previously seen, 159 new matching in content to the older ones, all randomly ordered, each image presented for 500 ms as in Experiments 1 to 3) were presented from 60 cm at a visual angle of $8.39^\circ \times 8.39^\circ$ as in Experiments 1 to 3.

Experiment 5: Parametric Investigation of Image Size on Memory.

Stimuli. Images were taken from the “LaMem” Dataset with memorability scores for each image (6) and were all resized to 900×900 pixels to avoid within- and across-condition size differences. These uniformly sized images were then scaled to be displayed according to the experimental condition (3° , 6° , 12° , 24°). We made sure that for each visual category (faces, people, indoors, and outdoors), the memorability scores for that category across size conditions would be uniform, and this assured that memorability scores across sizes were comparable (Fig. S1A).

Procedure. The exposure phase included 160 images presented in four blocks, each of a fixed specific size (3° , 6° , 12° , or 24°) that included 40 images (10 from each visual category); block order and image order within blocks were random. There were two versions of the exposure session; image sets assigned to the 3° condition in version 1 were assigned to the 24° condition in version 2 and vice versa, and image sets assigned to the 6° condition in version 1 were assigned to the 12° condition in version 2 and vice versa. As in Experiments 1 to 4, each image was presented for 2 s followed by a black screen of 500 ms. The task was to view and attend the images. No responses or fixations were required.

The test phase that followed included 320 midsized images (160 old, 160 new, visual angle of $8^\circ \times 8^\circ$) (*Test Phase Image Size Reasoning*) that were presented sequentially in random order, and participants were required to report for each image if they recalled seeing it earlier (old) or not (new). Each image was presented for 500 ms, after which a black screen appeared until a response was given (there was no time limit). No feedback was given.

Experiment 6: Investigating Exposure–Test Relative Size Effects. The experimental design was similar to the previous experiments, but in the test, we tested some of the images in their original presentation size (size at the test was the same size as at the exposure), while some were shown as midsized to allow us to compare with results of the earlier experiments.

Stimuli. The images that were used in this experiment were precisely the ones used in Experiment 5 (see above). In the exposure, the images were presented (as before) at either 3° or 24° . During the test, images were presented either at the same size as in the exposure or as midsized ($8^\circ \times 8^\circ$) as in Experiment 5.

Procedure. Since eye movements were recorded during the experiment, participants were seated while supported by a chin rest to minimize head movements. Eye tracker calibration was performed before the main experiment began. After calibration, at the exposure phase, participants viewed 160 images (80 small [$3^\circ \times 3^\circ$], 80 large [$24^\circ \times 24^\circ$]) presented in eight blocks (20 images per block); there were four blocks per size, with each block containing five images from each visual category (faces, people, indoors, and outdoors) and randomized image order within a block. Importantly, for each category we made sure that all five-image sets of that category had comparable memorability scores across all the blocks [LaMem (6)]. There were two

versions of the exposure session (exposure 1, exposure 2), and each version contained a different block order (exposure 1 block order: large–small–small–large–small–small–large, exposure 2 block order: small–large–large–small–small–large–small). The test phase included 320 images (160 old, 160 new); 120 of the old images (60 small and 60 large) were presented with the same size as in the exposure phase, and the additional 40 old images were presented as midsized (as in Experiments 1 to 5). The 160 new images were also presented in the same size proportions (60 as small, 60 as large, 40 as midsized). Test images were presented in 16 blocks of 20 images per block (10 old, 10 new), with the same presentation size across the block and 5 images from each visual category per block. There were two versions of the test phase (test 1, test 2) varying in block order (number of participants who performed the different versions of exposure and test were counterbalanced across participants: seven participants performed exposure 1 with test 1, seven participants performed exposure 2 with test 1, eight participants performed exposure 1 with test 2, and eight participants performed exposure 2 with test 2). Additional experimental details and setup are the same as in Experiment 5. The whole experiment (including eye tracking calibration) lasted approximately 30 min.

Experiment 7: Investigating the Effect of the Amount of Information on Memory.

Stimuli. All images in this experiment are the ones used in Experiment 5 (apart from the blurring procedure; see below) and taken from the LaMem Dataset (6). The exposure phase included smaller (3°) and larger (24°) images in blurred or sharp presentations. The smaller sharp images (112×112 pixels, $3^\circ \times 3^\circ$) were based on the original 900×900 -pixel images (Experiment 5) that were presented during the experiment by subsampling from each 900×900 image a smaller 112×112 image using in-house software (63–67) (scaling details in Experiment 1). The larger blurred images (900×900 pixels, $24^\circ \times 24^\circ$) were created from the original images (900×900 pixels) in the following manner. 1) Each image was reduced in size to 112×112 pixels using the MATLAB 2018b (MathWorks Inc.) *imresize* function (with the default bicubic interpolation). 2) Each image was then enlarged to 900×900 pixels by the *imresize* function in MATLAB (bicubic interpolation). 3) To eliminate apparent pixelation and artificially added edges caused by the enlargement, each image was then blurred in Adobe Photoshop CS6 version 13.0 x64 (Adobe Systems Inc.) with a Gaussian blur of 3.5 pixels (minimal radius that successfully eliminated the pixelation and apparent added edges). During exposure, these images were presented without any scaling. Importantly, this enlargement ensured that the larger images had the same or a smaller amount of information (as the blurring caused by the Gaussian blur may reduce the amount of information present in the outcome larger images relative to the smaller ones). The images of the two additional secondary conditions (smaller blurred, larger sharp conditions) that were meant to mask the contrast of interest in our experiment were created as follows. The smaller blurred images (112×112 pixels, $3^\circ \times 3^\circ$) were also based on 900×900 images (but blurred images; see below) that were presented during the experiment by scaling each 900×900 image to a smaller 112×112 image using in-house software (as in the smaller sharp images; see above). Here, to create a perception of comparable blurriness of these smaller images (as the blurriness in the larger presented images), the 900×900 images were blurred in Adobe Photoshop (see above) with a Gaussian blur of 8 pixels (compare with 3.5 pixels for larger blurred) so that when it was scaled by the in-house software, it created a comparable perception of the same blurriness as that of the larger blurred images. The larger sharp images (900×900 pixels, $24^\circ \times 24^\circ$) were the original 900×900 images as in Experiment 5 presented with no scaling. The test phase included sharp and blurred images all presented as 320×320 pixels ($\sim 8^\circ \times 8^\circ$). Old (already seen) images that were presented as sharp in the exposure phase were also presented here as sharp, and images that were presented in the exposure as blurred were also presented here as blurred. The old blurred images had the same amount of information as the smaller sharp (or larger blurred) images (from exposure) and were created in the following manner. 1) Each image was reduced in size to 112×112 pixels using the MATLAB 2018b *imresize* function (with the default bicubic interpolation). 2) Each image was then enlarged to 320×320 pixels by the *imresize* function in MATLAB (bicubic interpolation). 3) To eliminate apparent pixelation and artificially added edges caused by the enlargement, each image was then blurred in Adobe Photoshop (see above) with a Gaussian blur of two pixels (sufficient to eliminate the pixelation and apparent added edges). These images were presented with no scaling or subsampling. The old sharp images were based on the original sharp 900×900 images that were scaled to 320×320 pixels using an in-house software (as described in Experiment 1). As for the new images, the new blurred images were created and presented in the same manner as

the old blurred images, and the new sharp images were created and presented in the same manner as the old sharp images.

Procedure. Eye movements were recorded during both phases of Experiment 7. Participants underwent eye tracking calibration before the exposure phase. At exposure phase, participants viewed 160 images (80 small [$3^\circ \times 3^\circ$] and 80 large [$24^\circ \times 24^\circ$]). Half of the images of each size were blurred (40 small blurred, 40 small sharp, 40 large blurred, and 40 large sharp) and presented in eight blocks (20 images per block), with four blocks for each sharpness level (sharp or blur) and each block containing five images from each visual category (faces, people, indoors, and outdoors). To counterbalance condition order and to control for the possibility that certain images are more distinct and may be remembered better even after blurring, we created four versions of the exposure with two block order and also swapped the images that were sharp and those that were blurred between them (altogether, 2×2 versions; exposures 1 to 4). As in Experiments 1 to 5, each image was presented for 2 s followed by a black screen of 500 ms; the participant's task was to view the images, and no responses or fixations were required.

The test phase included 320 images (160 old, 160 new). Eighty of the old images were presented blurred (the same as in the exposure phase: 40 small blurred [$3^\circ \times 3^\circ$] and 40 large blurred [$24^\circ \times 24^\circ$]), and the other 80 old images were presented as sharp. Half of the 160 new images were also presented blurred, and the other half were sharp. All images were presented midsized (visual angle of intermediate size $8^\circ \times 8^\circ$, as in Experiments 1 to 5). The images were presented in 16 blocks (20 images per block, 10 old and 10 new images within each block, same level of sharpness/amount of information across the block, 5 images from each visual category). Participants were required to report for each image if they recalled seeing it earlier (old) or not (new). Each image was presented for 500 ms, after which a black screen appeared until a response was given (there was no time limit). There were two versions of the test phase (test 1, test 2), with each of them matching the exposure versions with a specific set of blurred images and a specific set of sharp images. Ten participants performed exposure 1 with test 1, eight participants performed exposure 2 with test 1, nine participants performed exposure 3 with test 2, and eight participants performed exposure 4 with test 2.

Test Phase Image Size Reasoning. In order for the test phase images to be unbiased toward the small or large images, we aimed to present them at an intermediate size that would be midway (size wise) between the small and large images but in cortical rather than in retinal space. Taking into account the cortical magnification factor that enhances foveal representations while reducing peripheral representations, we chose the midsize to be of a factor a bigger than the smaller images and a factor a smaller than the larger images such that $3^\circ \times a^2 = 21^\circ$ for Experiments 1 to 4 (resulting in test phase image

size of 8°) and $3^\circ \times a^2 = 24^\circ$ for Experiments 5 to 7. Based on the estimates for M given in earlier studies (30, 31), we used two equations that estimate the magnification factor in human's V1 [$M_{\text{linear}} = 17.3/(E + 0.75)$ by ref. 30 and $M_{\text{linear}} = 29.2/(E + 3.67)$ by ref. 31], both yielding similar results for the test phase image size: 8.5° (for Experiments 1 to 4) and 9.5° (for Experiments 5 to 7). Thus, the choices of $8.39^\circ \times 8.39^\circ$ as the midsize for the test images in Experiments 1 to 4 and $8^\circ \times 8^\circ$ in Experiments 5 to 7 were slightly closer (in cortical space) to the size of the smaller images and thus, unlikely to bias the results in favor of the bigger images.

Eye Tracking Analyses. Eye movements during Experiments 6 and 7 were recorded to verify that participants were viewing the images during exposure. To this end, we computed image dwell time during exposure for each participant (SI Appendix, Fig. S2). Dwell time was calculated as the percentage of time a participant spent within an image area. Image area was defined as image size and an additional 1° margin on each side of the image. For each participant, this was calculated per image and averaged across all images of the same image size. Participants with less than 80% dwell time on the 3° images were excluded from the analyses of Experiments 6 and 7 (participants excluded: Experiment 6: $n = 10$ of 30, Experiment 7: $n = 12$ of 35), and this is presented in SI Appendix, Fig. S3 below the threshold line.

Memorability Analysis: Per-Image Analysis. We examined for each image if size had any effect on its memorability by comparing performance when it was presented in bigger format relative to when presented as a smaller image. We performed this per-image analysis on the images used in Experiments 1 and 2 (same images were used, and images sets were switched between experiments; image sizes we compared for each image were those used in Experiments 1 and 2: 3° and 21°) and on the images used in Experiment 5 comparing per-image performance between version 1 and version 2 (since each image appeared in one version as bigger and in the other as smaller, image sizes were compared for each image as used in Experiments 5 version 1 and version 2: 3° and 24° or 6° and 12°).

Data Availability. Anonymized psychophysics, images, and scripts have been deposited in a publicly available database at the Center for Open Science website (<https://osf.io/7sr3cd>) (69).

ACKNOWLEDGMENTS. We thank Yulia Golland, Nurit Gronau, Daniel Levy, Yoram Bonne, Rafi Malach, and Ifat Levy for discussions and suggestions and Yuri Maximov for technical assistance. This work was supported by Israel Science Foundation Grant 1458/18 (to S.G.-D.).

1. G. H. Bower, M. B. Karlin, Depth of processing pictures of faces and recognition memory. *J. Exp. Psychol.* **103**, 751–757 (1974).
2. W. A. Bainbridge, The resiliency of image memorability: A predictor of memory separate from attention and priming. *Neuropsychologia* **141**, 107408 (2020).
3. F. I. M. Craik, E. Tulving, Depth of processing and the retention of words in episodic memory. *J. Exp. Psychol. Gen.* **104**, 268–294 (1975).
4. M. Moscovitch, F. I. M. Craik, Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. *J. Verbal Learn. Verbal Behav.* **15**, 447–458 (1976).
5. A. D. Wagner, J. D. E. Gabrieli, M. Verfaellie, Dissociations between familiarity processes in explicit recognition and implicit perceptual memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **23**, 305–323 (1997).
6. A. Khosla, A. S. Raju, A. Torralba, A. Oliva, "Understanding and predicting image memorability at a large scale" in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (IEEE, 2015), pp. 2390–2398.
7. T. F. Brady, T. Konkle, G. A. Alvarez, A. Oliva, Visual long-term memory has a massive storage capacity for object details. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 14325–14329 (2008).
8. L. Standing, J. Conezio, R. N. Haber, Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychon. Sci.* **19**, 73–74 (1970).
9. C. C. Williams, J. M. Henderson, R. T. Zacks, Incidental visual memory for targets and distractors in visual search. *Percept. Psychophys.* **67**, 816–827 (2005).
10. I. S. Utchkin, J. M. Wolfe, Visual search for changes in scenes creates long-term, incidental memory traces. *Atten. Percept. Psychophys.* **80**, 829–843 (2018).
11. S. A. Hillyard, E. K. Vogel, S. J. Luck, Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **353**, 1257–1270 (1998).
12. A. C. Huk, D. J. Heeger, Task-related modulation of visual cortex. *J. Neurophysiol.* **83**, 3525–3536 (2000).
13. Y. B. Saalman, S. Kastner, Gain control in the visual thalamus during perception and cognition. *Curr. Opin. Neurobiol.* **19**, 408–414 (2009).
14. M. Carrasco, Visual attention: The past 25 years. *Vision Res.* **51**, 1484–1525 (2011).
15. F. Pestilli, M. Carrasco, Attention enhances contrast sensitivity at cued and impairs it at uncued locations. *Vision Res.* **45**, 1867–1875 (2005).
16. J. Moran, R. Desimone, Selective attention gates visual processing in the extrastriate cortex. *Science* **229**, 782–784 (1985).
17. K. Anton-Erxleben, V. M. Stephan, S. Truee, Attention reshapes center-surround receptive field structure in macaque cortical area MT. *Cereb. Cortex* **19**, 2466–2478 (2009).
18. N. M. Weinberger, Dynamic regulation of receptive fields and maps in the adult sensory cortex. *Annu. Rev. Neurosci.* **18**, 129–158 (1995).
19. T. Womelsdorf, K. Anton-Erxleben, S. Truee, Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation. *J. Neurosci.* **28**, 8934–8944 (2008).
20. C. D. Gilbert, W. Li, Top-down influences on visual processing. *Nat. Rev. Neurosci.* **14**, 350–363 (2013).
21. A. L. Yarbus, *Eye Movements and Vision* (Springer US, 1967).
22. J. H. Reynolds, L. Chelazzi, Attentional modulation of visual processing. *Annu. Rev. Neurosci.* **27**, 611–647 (2004).
23. A. Gazzaley, A. C. Nobre, Top-down modulation: Bridging selective attention and working memory. *Trends Cogn. Sci.* **16**, 129–135 (2012).
24. W. Zhang, S. J. Luck, Feature-based attention modulates feedforward visual processing. *Nat. Neurosci.* **12**, 24–25 (2009).
25. R. B. H. Tootell et al., Functional analysis of V3A and related areas in human visual cortex. *J. Neurosci.* **17**, 7060–7078 (1997).
26. K. Grill-Spector, R. Malach, The human visual cortex. *Annu. Rev. Neurosci.* **27**, 649–677 (2004).
27. K. Grill-Spector et al., Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* **24**, 187–203 (1999).
28. E. T. Rolls, G. C. Baylis, Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Exp. Brain Res.* **65**, 38–48 (1986).
29. Y. Han, G. Roig, G. Geiger, T. Poggio, Scale and translation-invariance for novel objects in human vision. *Sci. Rep.* **10**, 1411 (2020).

30. J. C. Horton, W. F. Hoyt, The representation of the visual field in human striate cortex. A revision of the classic Holmes map. *Arch. Ophthalmol.* **109**, 816–824 (1991).
31. P. Kovács, B. Knakker, P. Hermann, G. Kovács, Z. Vidnyánszky, Face inversion reveals holistic processing of peripheral faces. *Cortex* **97**, 81–95 (2017).
32. R. A. Bjork, W. B. Whitten, Recency-sensitive retrieval processes in long-term free recall. *Cognit. Psychol.* **6**, 173–189 (1974).
33. M. W. Howard, V. Venkatadass, K. A. Norman, M. J. Kahana, Associative processes in immediate recency. *Mem. Cognit.* **35**, 1700–1711 (2007).
34. D. Rundus, Analysis of rehearsal processes in free recall. *J. Exp. Psychol.* **89**, 63–77 (1971).
35. H. Shteingart, T. Neiman, Y. Loewenstein, The role of first impression in operant learning. *J. Exp. Psychol. Gen.* **142**, 476–488 (2013).
36. W. A. McKenzie, M. S. Humphreys, Recency effects in direct and indirect memory tasks. *Mem. Cognit.* **19**, 321–331 (1991).
37. E. Capitani, S. Della Sala, R. H. Logie, H. Spinnler, Recency, primacy, and memory: Reappraising and standardising the serial position curve. *Cortex* **28**, 315–342 (1992).
38. I. Fischler, D. Rundus, R. C. Atkinson, Effects of overt rehearsal procedures on free recall. *Psychon. Sci.* **19**, 249–250 (1970).
39. W. Carter, Kitchen area and fireplace in Fjällastugan in Gullmarskogen nature reserve, Lysekil Municipality, Sweden. *Wikimedia* (2017). https://commons.wikimedia.org/wiki/File:Fj%C3%A4llastugan_kitchen_area.jpg. Accessed 29 September 2021.
40. E. Tulving, D. L. Schacter, Priming and human memory systems. *Science* **247**, 301–306 (1990).
41. C. C. Williams, Incidental and intentional visual memory: What memories are and are not affected by encoding tasks? *Vis. Cogn.* **18**, 1348–1367 (2010).
42. T. F. Brady, T. Konkle, G. A. Alvarez, A review of visual memory capacity: Beyond individual items and toward structured representations. *J. Vis.* **11**, 4 (2011).
43. J. M. Wolfe, Y. I. Kuzmova, How many pixels make a memory? Picture memory for small pictures. *Psychon. Bull. Rev.* **18**, 469–475 (2011).
44. P. A. Kollers, R. L. Duchnicky, G. Sundstroem, Size in the visual processing of faces and words. *J. Exp. Psychol. Hum. Percept. Perform.* **11**, 726–751 (1985).
45. K. M. Robinson, L. Standing, Effects of size changes on memory for words and pictures. *Percept. Mot. Skills* **71**, 919–922 (1990).
46. B. Uttl, P. Graf, A. L. Siegenthaler, The influence of object relative size on priming and explicit memory. *PLoS One* **3**, e3109 (2008).
47. I. Biederman, E. E. Cooper, Size invariance in visual object priming. *J. Exp. Psychol. Hum. Percept. Perform.* **18**, 121–133 (1992).
48. H. D. Zimmer, Size and orientation of objects in explicit and implicit memory: A reversal of the dissociation between perceptual similarity and type of test. *Psychol. Res.* **57**, 260–273 (1995).
49. R. N. Shepard, Recognition memory for words, sentences, and pictures. *J. Verbal Learn. Verbal Behav.* **6**, 156–163 (1967).
50. W. Sato, S. Yoshikawa, Recognition memory for faces and scenes. *J. Gen. Psychol.* **140**, 1–15 (2013).
51. P. Isola, Jianxiang Xiao, D. Parikh, A. Torralba, A. Oliva, What makes a photograph memorable? *IEEE Trans. Pattern Anal. Mach. Intell.* **36**, 1469–1482 (2014).
52. A. Todorov, C. P. Said, A. D. Engell, N. N. Oosterhof, Understanding evaluation of faces on social dimensions. *Trends Cogn. Sci.* **12**, 455–460 (2008).
53. A. Todorov, N. Oosterhof, Modeling social perception of faces [social sciences]. *IEEE Signal Process. Mag.* **28**, 117–122 (2011).
54. T. Konkle, A. Oliva, A familiar-size Stroop effect: Real-world size is an automatic property of object representation. *J. Exp. Psychol. Hum. Percept. Perform.* **38**, 561–569 (2012).
55. M. V. Maltz et al., Familiar size affects the perceived size and distance of real objects even with binocular vision. *J. Vis.* **21**, 21 (2021).
56. R. Sousa, J. B. J. Smeets, E. Brenner, Does size matter? *Perception* **41**, 1532–1534 (2012).
57. I. Levy, U. Hasson, G. Avidan, T. Hendler, R. Malach, Center-periphery organization of human object areas. *Nat. Neurosci.* **4**, 533–539 (2001).
58. R. B. Tootell et al., The retinotopy of visual spatial attention. *Neuron* **21**, 1409–1422 (1998).
59. W. A. Bainbridge, C. I. Baker, Boundaries extend and contract in scene memory depending on image properties. *Curr. Biol.* **30**, 537–543.e3 (2020).
60. H. Intraub, Rethinking visual scene perception. *Wiley Interdiscip. Rev. Cogn. Sci.* **3**, 117–127 (2012).
61. E. A. Maguire, S. L. Mullanly, The hippocampus: A manifesto for change. *J. Exp. Psychol. Gen.* **142**, 1180–1189 (2013).
62. H. Intraub, M. Richardson, Wide-angle memories of close-up scenes. *J. Exp. Psychol. Learn. Mem. Cogn.* **15**, 179–187 (1989).
63. O. Kreichman, Y. S. Bonne, S. Gilaie-Dotan, Investigating face and house discrimination at foveal to parafoveal locations reveals category-specific characteristics. *Sci. Rep.* **10**, 8306 (2020).
64. G. Rosenzweig, Y. S. Bonne, Familiarity revealed by involuntary eye movements on the fringe of awareness. *Sci. Rep.* **9**, 3029 (2019).
65. H. Harris et al., Perceptual learning in autism: Over-specificity and possible remedies. *Nat. Neurosci.* **18**, 1574–1576 (2015).
66. Y. S. Bonne, A. Cooperman, D. Sagi, Motion-induced blindness in normal observers. *Nature* **411**, 798–801 (2001).
67. Y. S. Bonne, D. Sagi, U. Polat, Spatial and temporal crowding in amblyopia. *Vision Res.* **47**, 1950–1962 (2007).
68. Y. S. Bonne, Y. Adini, U. Polat, Contrast sensitivity revealed by microsaccades. *J. Vis.* **15**, 11 (2015).
69. S. Gilaie-Dotan, S. Masarwa, Larger images are better remembered during naturalistic encoding. Open Science Framework. <https://osf.io/7sr3d/>. Deposited 19 December 2021.