Downloaded from clal learnmem.cshlp.org on November 5, 2026 - Published by Cold Spring Harbor Laboratory Press

Research

Properties of visual episodic memory following repeated encounters with objects

Mark W. Schurgin[1] and Jonathan I. Flombaum[2]

[1]Department of Psychology, University of California, San Diego, California 92093, USA; [2]Department of Psychological & Brain Sciences, Johns Hopkins University, Baltimore, Maryland 21218, USA

A person sees an object once, and then seconds, minutes, hours, days, or weeks later, she sees it again. How is the person's visual memory for that object changed, improved, or degraded by the second encounter, compared to a situation in which she will have only seen the object once? The answer is unknown, a glaring lacuna in the current understanding of visual episodic memory. The overwhelming majority of research considers recognition following a single exposure to a set of objects, whereas objects reoccur regularly in lived experience. We therefore sought to address some of the more basic and salient questions that are unanswered with respect to how repetition affects visual episodic memory. In particular, we investigated how spacing between repeated encounters affects memory, as well as variable input quality across encounters and changes in viewed orientation. Memory was better when the spacing between encounters was larger, and when a first encounter with an object supplied high quality input (compared to low quality input first, followed later by higher quality input). These experiments lay a foundation for further understanding how memory changes, improves, and degrades over the course of experience.

Human visual episodic memory is remarkable, variously described as massive, invariant, and explicit: respectively, storing a large number of objects, able to recognize an object despite changes in appearance and viewing conditions, and able to discriminate between objects that are different but share visual properties (Quiroga et al. 2005; DiCarlo and Cox 2007; Brady et al. 2008; Rust and Stocker 2010; Schurgin 2018). Stated simply: people remember a lot about the things that they see over a lifetime, and they remember with a level of precision that remains out of reach for artificial systems (Pinto et al. 2008; DiCarlo et al. 2012; Andreopoulos and Tsotsos 2013). The impressive nature of visual episodic memory is salient in everyday experience. Most people effortlessly recognize a great deal of their past visual experiences, a fact that is well documented experimentally (Shepard 1967; Standing 1973; Cansino et al. 2002; Brady et al. 2008, 2009, 2013; Konkle et al. 2010; Yonelinas et al. 2010; Squire and Wixted 2011; Yu et al. 2012; Guerin et al. 2012). An important difference between typical laboratory experiments and everyday experience, however, is that experiments usually involve a single exposure to each of the images a subject will eventually be tested on later, whereas everyday experience usually involves reoccurrences.

How then do our memories change as we reencounter previously seen visual stimuli? Seeing an object more than once does improve later memory performance with respect to that object (Hintzman 1976; Wiggs et al. 2006; Reagh and Yassa 2014). Beyond this, surprisingly little is currently known.

The purpose of the current study is by no means to fill this void entirely. Instead, we sought to ask three questions that appear, to us, as intuitively and empirically motivated places to begin. The project is largely descriptive at this time, with the expectation that establishing a few basic effects will eventually help researchers to characterize the mechanisms that produce any observed changes. The three questions of interest currently are: (1) Does memory respond to changes in the quality of an input over reoccurrences? (2) Do viewpoint (orientation) changes over reoccurrences affect memory? (3) And (how) does the spacing or time between reoccurrences affect memory?

Corresponding author: maschurgin@ucsd.edu
Article is online at http://www.learnmem.org/cgi/doi/10.1101/lm.047167.117.

Experiment I: variable quality across encounters

To begin, we considered the fact that experiences with objects are likely to vary in quality. Therefore, we asked whether memory benefits when a lower quality encounter either precedes or follows a higher quality one? This is an important first question because all three potential outcomes seem plausible. It may be that the last encounter an observer has with an object has the greatest bearing on subsequent long-term memory performance: a recency-effect. In contrast, it may be that when observers are first given a high-quality encounter this fosters better conditions for the assimilation of information in the future, resulting in more robust memories. And of course, it is possible that memory is commutative (like addition), with a poorer input and a better input always assimilating to the same endpoint, regardless of the sequence. While not explicitly investigating repeated encounters, previous research has shown there is a performance benefit when the presentation of images during study and test conditions are congruent (Ray and Reingold 2003). Thus, it is possible that memory mechanisms operate ideally when there is no variability across inputs, regardless of whether those encounters are higher or lower quality. It is important to note that each of these possibilities may provide a more nuanced explanation of the reconsolidation literature, which suggests that when new information about an item is introduced the memory for that item is reactivated, and the new information subsequently alters the original memory through reconsolidation

© 2018 Schurgin and Flombaum This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first 12 months after the full-issue publication date (see http://learnmem.cshlp.org/site/misc/terms.xhtml). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

Downloaded from clal learnmem.cshlp.org on November 5, 2026 - Published by Cold Spring Harbor Laboratory Press

and reencoding mechanisms (Nader and Hardt 2009; Alberini and LeDoux 2013).

To answer this question, we presented participants with a stream of real-world objects during an encoding phase. Each image in the stream appeared twice, and each image was embedded in variable noise. To create noise, we randomly scrambled a percentage of pixels in each image, a technique that was successful in previous experiments (Schurgin and Flombaum 2015, 2018). The logic is that randomly scrambling pixels is a way to continuously degrade the ability to recognize an object, hopefully approximating the effects of a variety of real world factors that would do the same, such as dim lighting, viewing from a distance, or viewing through a partially transparent surface. By varying the amount of noise, we sought specifically to control the quality of input during first and second encounters. In addition to presenting some objects in high-noise first, and some in low-noise first, we also presented some objects with low-noise (i.e., better input) on both occasions, some with more noise on both occasions, and we also included images that appeared only once, with either high or low noise; these four extra conditions were intended to supply baselines for comparison.

Thus, the experiment included six conditions: an image could appear twice with both presentations in low-noise, both presentations in high-noise, an initial presentation in high and then low-noise (high-low), and an initial presentation in low and then high-noise (low-high). Additionally, 1/3 of the images only appeared once, either with higher or lower noise (see Fig. 1). Encoding of images took place in an incidental encoding paradigm, and recognition memory was later tested using a two-alternative forced-choice procedure (2AFC) wherein a previously shown image was paired with an unshown foil, and participants were asked to identify which of the two images they previously saw. We chose to use a 2AFC test rather than individual object report (i.e., Old/New), as performance is generally better in 2AFC tasks (given the same underlying memory strength, it is always easier to pick the maximum likelihood of two things; see Macmillan and Creelman 2004), and we wanted to maximize our ability to observe differences across conditions. In half of the 2AFC tests, the "new" image was categorically different from the image paired with it. We call these "New" test trials. For a more challenging test of recognition, half the tests involved a categorically similar object as the paired foil, trials which we will call "Similar" tests.

For all analyses, we converted percent correct into d′, which is a signal detection measure that offers an unbiased measure of a memory match signal strength (Green and Swets 1966). We entered the data into a within-subjects ANOVA ($n = 29$) with test type and exposure type as a factor. We observed a main effect of test type, $F_{(1,28)} = 438.63$, $P < 0.01$, $\eta_p^2 = 0.94$. There was a main effect of exposure type, $F_{(5,140)} = 72.07$, $P < 0.01$, $\eta_p^2 = 0.72$. And we observed a trending interaction between test and exposure type, $F_{(5,140)} = 2.24$, $P = 0.054$, $\eta_p^2 = 0.074$

We followed up with planned comparisons intended to address two specific questions. First, was our noise manipulation effective in altering memory performance? For the New test we found significantly better performance when a low-noise encounter followed a high-noise encounter (d′ = 2.93) compared to when both encounters were high-noise (d′ = 2.34), $t_{(28)} = 4.53$, $P < 0.01$, Cohen's dz = 0.84. This comparison was also significant in the case of the Similar test (d′ = 1.45 versus 1.18), $t_{(28)} = 2.36$, $P = 0.03$, Cohen's dz = 0.47. Thus, noise had an impact on performance in the minimal sense that low-noise did improve performance, compared to encountering only high-noise. Also indicating that the noise manipulation was effective, two low-noise encounters (d′ = 2.02) produced significantly better performance than low-followed by high-noise (d′ = 1.67) for the difficult Similar test, $t_{(28)} = 3.52$, $P < 0.01$, Cohen's dz = 0.66.

We also investigated performance for the dual encounter relative to single encounter conditions. Across both New and Similar tests, a single high noise encounter resulted in significantly worse performance relative to all the dual exposure conditions (all P's < 0.01). However, there were some interesting differences relative to a single low noise encounter. For the New test, performance for a single low encounter (d′ = 2.65) was significantly better than the High-High condition (d′ = 2.34), $t_{(28)} = 2.63$, $P = 0.01$, Cohen's dz = 0.49. And for the Similar test, there was no significant difference across the single low encounter (d′ = 0.99) and High-High condition (d′ = 1.18), $t_{(28)} = 1.47$, $P = 0.15$, Cohen's dz = 0.27. This suggests that a single low noise encounter contained just as much, if not slightly more information, than two high noise encounters. For all other comparisons, across both New and Similar tests performance was significantly better for the dual encounter conditions versus a single low noise encounter (all P's < 0.03).

The main question of interest was whether order matters when encounters have mixed quality. We found that order does matter: a low-noise encounter first was better than having it come second. For the New test, planned follow-up analyses found a significant difference between High-Low (d′ = 2.93) and Low-High (d′ = 3.14) conditions, $t_{(28)} = 2.06$, $P = 0.048$, Cohen's dz = 0.38. For the Similar test, the comparison produced a similar effect size, but it was only trending in terms of significance (d′ = 1.45 versus 1.67), $t_{(28)} = 1.76$, $P = 0.09$, Cohen's dz = 0.33 (Fig. 2).

To summarize, the interesting result here was that a low-noise encounter followed by a high-noise encounter supported better recognition performance than the reverse (i.e., high-noise first). This is consistent with an integration mechanism in which a high-fidelity exposure forms a basis that allows more information to be gleaned from a later low-fidelity exposure. This could be realized, for example, if pattern completion (see Yassa and Stark 2011) depends more on the strength of the stored information that triggers it, as opposed to the quality of the incoming information that becomes assimilated. Regardless of specific mechanism, the result is descriptively interesting, perhaps

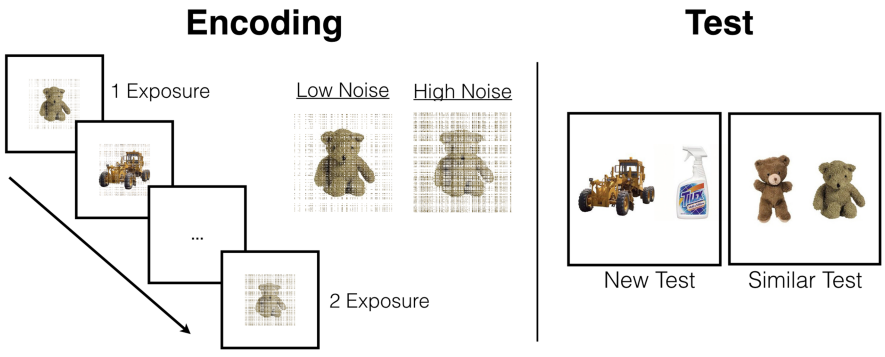


Figure 1. Illustration of the methods used in Experiment 1. During incidental encoding, participants saw images of real-world objects twice, embedded in either low (40%) or high (70%) noise across encounters (in addition to single encounter control conditions). At subsequent test, participants saw two images on screen. One image had been previously encoded, whereas the other was a novel foil. For New tests, the foil was a completely new, categorically distinct image. For Similar tests, the foil was a similar-looking object.

Downloaded from clal learnmem.cshlp.org on November 5, 2026 - Published by Cold Spring Harbor Laboratory Press

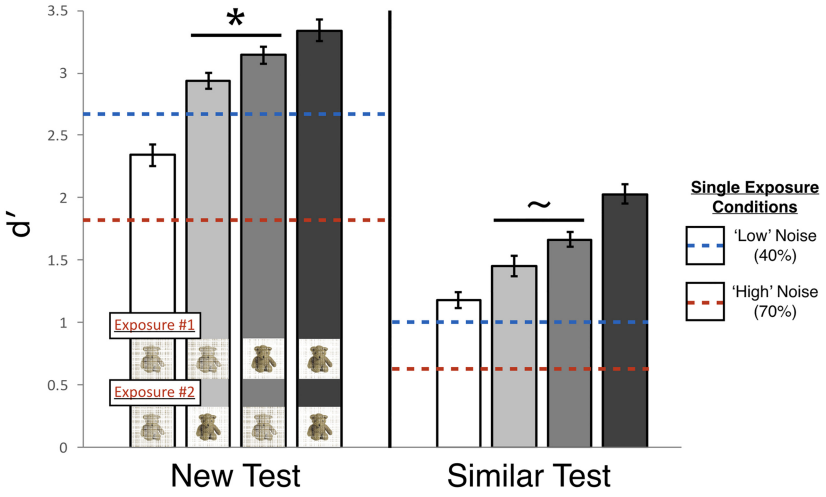


Figure 2. Results of Experiment 1. The primary finding was that performance was better when observers first encountered a low-noise image followed by a high-noise image compared to the reverse (i.e., high-noise first). * designates $P < 0.05$, ~ designates $P = 0.09$. Error bars represent within-subject error (to remove between-subjects variability, see Cousineau 2005).

even surprising. Recognition performance was "worse" when the exemplar that observers saw most "recently" was the "better fidelity" one.

Experiment 2: variable orientation across encounters

Orientation or viewpoint changes are a typical type of everyday variability; the same object will rarely (if ever) be seen from the exact same orientation as before, introducing dramatic changes to surface feature information across encounters (DiCarlo and Cox 2007; Rust and Stocker 2010).

We sought to investigate what happens when observers see an object twice from the same compared with two different orientations, conditions we will call Consistent Encounters and Variable Encounters. We also manipulated orientation at test, sometimes showing an object at one of the orientations in which it was shown before, and sometimes showing an object at a third, never-before-shown orientation. During test, the object was always paired with a new foil, a never-before seen and categorically different object. Therefore, the test manipulation in this experiment was whether the target object was shown in an orientation it had been shown in before (Original Orientation Test), or in a new orientation (New Orientation Test; see Fig. 3). Although long-term memory performance is sometimes impaired when an object needs to be recognized at a new orientation (from the one when it was encoded; Srinivas 1995; Schurgin and Flombaum 2018), certain task demands have been shown to produce viewpoint-invariant recognition (Tarr and Hayward 2017), and it has also been theorized that viewpoint invariant representations may be best acquired from multiple, variable encounters with the same object (Srinivas 1995).

In other words, tolerance is possibly bred from familiarity since a relevant object can look different from various viewpoints. Accordingly, we predicted that in the New Orientation Test, memory performance would be better in the Variable Encounter condition compared with the Consistent Encounter condition, since having already seen an object from two orientations, as opposed to one, would make it easier to recognize that same object from a never-before-seen orientation.

But that is not what we found. There was no significant difference in New Orientation Test performance as a function of exposure condition: Consistent Encounter ($d' = 2.53$) versus Variable Encounter ($d' = 2.65$) conditions, $t_{(29)} = 0.71$, $P = 0.48$. Nor was there a difference in the Original Orientation Test condition: Consistent Encounter ($d' = 2.91$) versus Variable Encounter ($d' = 2.72$) conditions, $t_{(29)} = 1.21$, $P = 0.24$ (see Fig. 4).

Given the overall high performance observed in d′, we sought another measure that would be perhaps more sensitive to condition, latency to respond.[1] Previous research has found this to be a sensitive measure, particularly when examining differences in object recognition by orientation (Srinivas 1995; see also Srinivas and Verfaellie 2000). But here, we failed to find significant differences for Variable versus Consistent conditions: for the New Orientation Test, Consistent Encounter (latency to respond = 1.71 sec) versus Variable Encounter (1.70 sec) conditions, $t_{(29)} = 0.21$, $P = 0.83$. For the Original Orientation Test condition, Consistent Encounter (1.57 sec) versus Variable Encounter (1.58 sec) conditions, $t_{(29)} = 0.17$, $P = 0.86$.

Encountering an object twice from two different orientations did not produce more tolerant memories. But since there was no difference between the encounter conditions with any test, it remains difficult to draw a strong conclusion. The task used here was perhaps too easy, in general. The high performance observed is likely due in part to all images being presented at encoding without noise and that fewer items were encoded (due to stimuli constraints) relative to Experiment 1. Future research should further investigate how repeated exposures with orientation changes do or do not impact subsequent recognition.

When comparing different test performance across each condition, we did find a significant effect such that in the Consistent Encounter condition performance was significantly better for the Original Orientation than the Different Orientation test, $t_{(29)} = 2.75$, $P = 0.01$. We failed to observe this effect in the Variable Encounter condition, $t_{(29)} = 0.55$, $P = 0.59$. This suggests that overall, performance was slightly better in the Consistent Encounter condition.

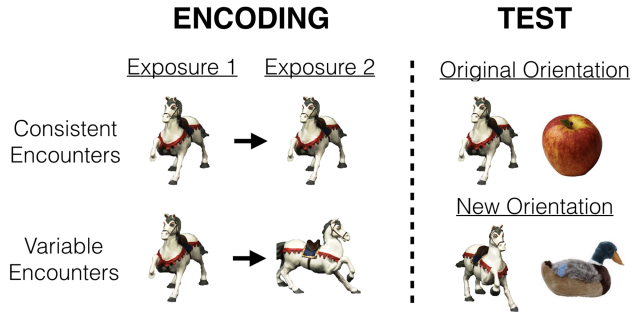


Figure 3. Illustration of the design of Experiment 2. During encoding, objects were encountered twice either from the same orientation (Consistent Encounters) or two different orientations (Variable Encounters). At subsequent test, the target was shown alongside a foil in the orientation it had been shown in before (Original Orientation Test) or a new orientation (New Orientation Test).

[1]We thank an anonymous reviewer for making this suggestion.

Downloaded from clal learnmem.cshlp.org on November 5, 2026 - Published by Cold Spring Harbor Laboratory Press

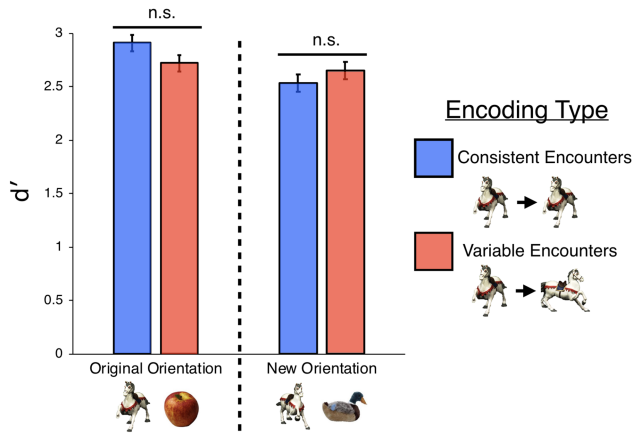


Figure 4. Results of Experiment 2. Across both Original Orientation and New Orientation tests, we failed to find a significant difference whether an object had been encountered during encoding twice from the same orientation (Consistent Encounters) or two different orientations (Variable Encounters). Error bars represent within-subject error.

Experiment 3: variable distance across encounters

What about the amount of time that elapses between two encounters with the same stimulus? Spacing effects have been explored in the context of verbal and semantic episodic memory (Melton 1970; Smolen et al. 2016), but little is known about how variable distance across encounters may affect "visual" episodic memory. Previous research has found that when participants are shown a continuous stream of images and tasked with detecting repeated objects, there is no significant difference in performance for items repeated with a short lag (3–14 items) or a long lag (19–34 items) (Suzuki et al. 2011). But detecting repeated targets in a stream of images may involve working memory and other executive functions in ways that later recognition tests do not. The question remains: Is it better for a second encounter with an object to take place immediately, or with some time in-between, in terms of later recognition of that object?

We explored this question by showing objects twice during an incidental encoding paradigm. (This experiment also included objects that were shown only once, to supply a performance baseline). Here, however, we varied the distances between repetitions in terms of the number of intervening objects shown before a second encounter. Objects were either repeated after one intervening item (i.e., 2-back, which amounted to 3 sec between encounters), or they repeated following ten intervening items (i.e., 11-back, amounting to 25.5 sec between encounters; see Fig. 5). At subsequent test, we used the same procedure as in Experiment 1. Participants were given a 2AFC test, where half the trials were "New" test trials (where the foil was categorically different from the image paired with it) and the other half of trials were "Similar" test trials involving a categorically similar object as the paired foil.

We chose these two spacing intervals because they were the shortest and longest spacing available using the incidental encoding approach of the previous experiments. The 2-back spacing involved nearly immediate repetition, but with enough time to preclude any chance of repetition blindness (Kanwisher, 1987). And 11-back was the longest spacing available to ensure that half the objects could repeat at this duration over the course of a 25-min exposure session. Contrasting the longest and shortest spacing available allowed us to compare two reasonable options for the interaction between timing and repetition. One possibility is that repeated exposure reinforces memory best when there has been little opportunity for intervening decay and/or interference. This hypothesis would predict that front-loading repetition would benefit memory performance, compared to repetition over a longer gap. On the other hand, a longer gap may supply a better basis for reinforcement through repetition by supplying a base memory that has had more time to enter long-term memory and benefit from potential reactivation during the second encounter.

We entered the data into a within-subjects ANOVA ($n = 28$) with test type and distance as a factor. We observed a main effect of test type, $F_{(1,27)} = 138.6$, $P < 0.01$, $\eta_p^2 = 0.93$. We also observed a main effect of distance, $F_{(2,54)} = 17.04$, $P < 0.01$, $\eta_p^2 = 0.39$. Finally, we observed a trending interaction between test type and distance, $F_{(2,54)} = 2.98$, $P = 0.059$, $\eta_p^2 = 0.10$.

For the New test, follow-up analyses revealed that 11-back performance (d′ = 3.74) was significantly better than 2-back performance (d′ = 3.38), $t_{(27)} = 2.75$, $P < 0.01$, Cohen's dz = 0.54. In addition, 2-back performance was significantly better than a single exposure (d′ = 3.08), $t_{(27)} = 2.46$, $P = 0.02$, Cohen's dz = 0.46 (see Fig. 6).

For the Similar test, follow-up analyses failed to find a significant difference between 11-back (d′ = 1.71) and 2-back (d′ = 1.63) performance, $t_{(27)} = 0.92$, $P = 0.37$, Cohen's dz = 0.17. 2-back performance compared with a single exposure trended toward significance with a reasonable effect size, (d′ = 1.41), $t_{(27)} = 1.89$, $P = 0.07$, Cohen's dz = 0.36.

Overall, performance was better for objects encountered twice (2-back or 11-back) compared to once (single exposure). More importantly, for the New test, performance was significantly better for objects encountered in the 11-back compared to the 2-back condition. When additional time between encounters was introduced subsequent recognition performance benefited.

Why would a longer delay supply a better basis upon which to assimilate information during the repeated exposure? A possibility has to do with transfer from short-term or working memory into long-term memory. Specifically, a short delay might mean that the assimilation of the two encounters happens in a short-term memory store, and that the single assimilated memory moves into long-term memory. That long-term memory then lays dormant, never becoming reactivated during the experiment, until the recognition test, that is. In the longer delay, however, the trace from the first encounter may have already moved into long-term memory when the repetition arrives. The second encounter may now reactivate a long-term memory. The key difference between the delay conditions may therefore be whether or not a long-term memory for a given stimulus becomes reactivated prior to the

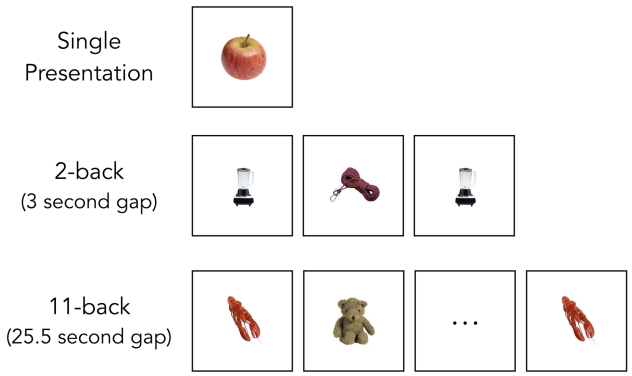


Figure 5. Encoding conditions of Experiment 3. Objects were encountered either once (Single Presentation), or twice with one intervening item (2-back, 3 second gap) or ten intervening items (11-back, 25.5 second gap).

Downloaded from clal learnmem.cshlp.org on November 5, 2026 - Published by Cold Spring Harbor Laboratory Press

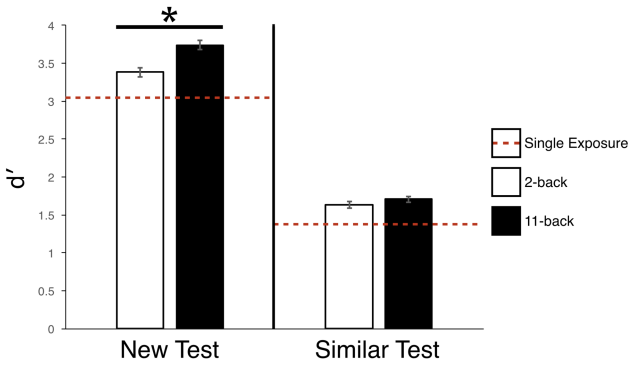


Figure 6. Results of Experiment 3. Across both New and Similar tests, performance increased from the single encounter condition (red dashed-line) to the 2-back condition (white), and was best at the 11-back condition (black). * designates $P < 0.01$. Error bars represent within-subject error.

recognition test. Imaging and other approaches may be useful for testing this idea, or otherwise explaining the causes of the delay differences observed.

Discussion

While the majority of research investigating visual episodic memory has been concerned with performance following one exposure, the formation of our visual memories in the real-world develops over time and over repeated encounters with stimuli. We sought to expand our understanding of the dynamics of memory changes across time through examining how memories change after reexposure. What are the variables that improve our memory strength or tolerance over time?

We observed that both the quality of an observer's initial encounter with an object (Experiment 1) as well as the time between encounters (Experiment 3) affect memory performance. However, we failed to find that encountering an object multiple times at different orientations creates more tolerant memories (Experiment 2). Altogether, these findings identify several factors demonstrating how memory changes, improves, and degrades over the course of multiple experiences.

Clarifying integration mechanisms in memory

The present results clarify how integration mechanisms operate over multiple prolonged encounters with objects. Specifically, in Experiment 1 we observed that when given multiple encounters with an object, even if one of those encounters was embedded in high noise, recognition performance was better compared to a single low noise exposure. For example, High-Low noise performance was always significantly better than a single Low noise exposure, suggesting that even when observers were initially given a degraded input, this still allowed them to integrate subsequent information to construct a robust long-term memory. This is an important point, as it not only precludes the possibility that observers primarily rely on their last encounter with an object to inform subsequent memory tests (i.e., a recency-effect), but it also demonstrates that integration over long durations does occur, even when one is supplied an initially degraded representation.

This suggests information is being indexed multiple times in order to construct a more diagnostic representation. Considering this can occur even when initially supplied a degraded (high noise) encounter, it is interesting to consider what mechanisms may be facilitating this indexing process. One possibility is that participants are relying on some sort of schema or categorical knowledge template (Konkle et al. 2010), which facilitates integrating even noisy encounters into a representation that can be subsequently indexed with new information. However, if this is the case, would these same effects occur for previously unseen objects, such as greebles (artificially created objects observers would have no previous experience with)? Previous research suggests that integration mechanisms would likely be less efficient, as recognition performance for greebles relies on training and repeated experience (Gauthier et al. 1998), and short-term memory for greebles is affected by experience (Wagar and Dixon 2005). This question remains far from resolved, and future research should investigate the potential representational structure that allows for the assimilation and indexing of multiple experiences. The methods developed here could be used productively to explore questions about unfamiliar objects.

It is also important to note that recognition was better when a low-noise encounter was followed by a high-noise encounter than the reverse (i.e., high-noise first). This is surprisingly novel, as performance was worse when the last encounter an observer had with an image was "better." This supports that integration mechanisms in memory operate best when a high-fidelity exposure forms a basis that allows more information to be gleaned from a later low-fidelity exposure. In the reconsolidation literature, a memory is reactivated when new information about that item is reencountered, and that new information subsequently alters (or is integrated) with the original memory through reconsolidation and reencoding (Nader and Hardt 2009; Alberini and LeDoux 2013). Our results suggest that reconsolidation processes may operate best when an observer has a stronger original memory to rely on, as it is better able to assimilate newly encountered information. This is consistent with previous research, which has shown that after an initial exposure with a line drawing of a fragmented object, less information is needed to correctly identify that same object during a future test (Kennedy et al. 2007).

Our results also have several possible implications for informing the theoretical understanding of pattern completion. Pattern completion is defined as a neurocomputational process by which incomplete or degraded signals are filled-in based on previously stored representations, and it is theorized to assist in our long-term memory abilities (Yassa and Stark 2011). In the context of our Low-High effect, it could be that the noisier second encounter overlaps sufficiently with the previously stored representation to enact pattern completion. As a result, the subsequent exposure only strengthens the existing representation (this could be one potential mechanism through which the reconsolidation processes described above could occur). In contrast, a first high-noise encounter may not supply a sufficient basis for engaging pattern completion. Put another way: pattern completion may be a process that allows the past to fill in the present (so to speak), but not the present to fill in the past.

Potential implications for memory consolidation

There remains much that is unknown about the consolidation process of representations into visual episodic memory. To be clear, by consolidation we mean the representation had more time to enter long-term memory systems and amalgamate into a stable, diagnostic representation (which may be distinct from biological, molecular, or neurological definitions of consolidation processes). Previous research has shown that working memory maintenance appears to be a critical step to subsequent long-term consolidation (Ranganath et al. 2005). However, other factors such as the total number of items to be remembered (whether 20 or 360 items) or increasing exposure past 1 sec (to either 3 or 5 sec) appear to have no effect (positive or negative) on subsequent long-term memory performance (Brady et al., 2013).

Downloaded from clal learnmem.cshlp.org on November 5, 2026 - Published by Cold Spring Harbor Laboratory Press

In Experiment 3, we demonstrated that when given two encounters of the same object, memory performance is better when additional time is introduced between encounters. This is largely consistent with the learning literature in other domains, which has shown that additional consolidation time generally improves performance (Melton 1970; Smolen et al. 2016). A limitation of the present research is that only two variable distances were explored (3 versus 25.5 sec). The current results could reflect that in the 2-back condition information from the first presentation was still being processed in long-term memory, whereas in the 11-back condition the additional delay allowed a more fully formed representation to be constructed that could better assimilate information from the second encounter. It remains unknown whether additional time between encounters improves performance further, or at what time point memory performance plateaus.

It is also important to note that these effects may not be due to consolidation processes, but rather reflect subsequent encounters being more distinctive. In the 2-back condition, there is only one intervening item, and the previously held encounter might still be held to a certain extent in short-term memory. However, in the 11-back condition there are ten intervening items before the second presentation, which may in turn make that subsequent encounter more distinct. As a result, memory systems may see this as a richer learning opportunity (relative to the 2-back condition), resulting in the better memory performance observed at test.

Recently, research has shown that increasing distance between encounters in a visual memory task may "decrease" memory performance. Singh et al. (2017) observed that when observers are tasked with detecting repeats in a long sequence of object images, repetition detection decreases as the distance between repeated items increases. Initially, these results may seem incompatible with the present findings, but the confines and computational challenges in this task are quite different from visual episodic memory. Generally, in long-term memory (or object recognition) observers are not consciously attending to a stream of stimuli presented in a specific order, with the sole task of identifying variably spaced repeats. To accomplish this task observers need to utilize additional cognitive components beyond memory, including working memory and other executive functions. Considering visual episodic memory is typically described as a passive system (Squire 2004; Cowan 2008; Brady et al. 2011; Schurgin 2018), these results are likely specific to the constraints associated with recognizing repeated images within a continuous timeline (see also Suzuki et al. 2011).

Insight into tolerance

Recent theories suggest part of the encoding process supporting object recognition over the long-term involves integrating information about an object over brief encounters (Wallis and Bülthoff 2001; Cox et al. 2005; Cox and DiCarlo 2008; Schurgin et al. 2013; Schurgin and Flombaum 2017). Humans appear to temporally associate information to build object representations through saccades (Cox et al. 2005; Li and DiCarlo 2008; Isik et al. 2012; Poth et al. 2015; Poth and Schneider 2016), as well as through exploiting expectations about object physics (Schurgin et al. 2013; Schurgin and Flombaum 2017). This suggests that tolerance in the long-term may ultimately depend on experience.

While tolerance may be learned through integrating encounters in the short-term, the present results suggest that building tolerant representations does not depend on multiple encounters with an object over longer periods of time. In Experiment 2, observers constructed robustly tolerant memories regardless of whether they saw an object multiple times at the same or different orientations. This demonstrates that tolerant memory representations can be rapidly constructed without much exposure to variability across multiple encounters.

Initially this may appear surprising, but it is ultimately consistent with previous research investigating tolerance in human object recognition. When briefly flashed a photo, observers can rapidly (<300 msec) recognize objects despite never having viewed the photograph before (Potter 1976). This result is consistent whether observers are shown color photographs or simplified line drawings of objects, suggesting tolerance for recognizing objects can be rapidly extracted for even abstract stimuli (Biederman 1987; Biederman and Ju 1988). Considering how quickly humans can recognize objects despite considerable variability in inputs (including line drawings devoid of any other surface feature information), it follows that tolerance may not depend on learning through multiple prolonged encounters with the same object.

Conclusion

The present results have broad implications for researchers studying memory, learning, and object recognition. They demonstrate the importance of moving past singular experiences to investigate the integration of multiple encounters, and they begin to elucidate the nature of the mechanisms that integrate new input into long-term memories. Specifically, they suggest that these mechanisms function more effectively given an opportunity to consolidate initial encounters, consistent with many other types of nonvisual learning (Melton 1970; Smolen et al. 2016). Additionally, our findings suggest that integration works best when an initial encounter supplies a high-quality basis to assimilate with subsequent experiences. It is important to note the present results do not clarify the exact nature of the specific mechanisms producing the observed effects, and alternative explanations may exist. However, we have identified several factors that alter memory quality over repeated encounters, and hope this foundation will assist future researchers in identifying the specific mechanisms that support or degrade visual episodic memory over repeated experiences.

Materials and Methods

Experiment 1

A group of 32 Johns Hopkins University undergraduates participated in Experiment 1. The results of three participants were excluded due to responding randomly (i.e., chance performance). All participants reported normal or corrected-to-normal visual acuity. Participation was voluntary, and in exchange for extra credit in related courses. The experimental protocol was approved by the Johns Hopkins University IRB. Stimuli were presented using MATLAB and the Psychophysics toolbox (Brainard 1997; Pelli 1997). All stimuli were presented within the full display frame of 39.43° × 24.76°. During the first phase of the experiment (incidental encoding), participants were shown 300 color images of real-world objects on a computer screen (taken from Schurgin and Flombaum 2017, 2018—full database publicly available at http://markschurgin.com), with each image shown twice as already described. The cover task was to indicate whether an item onscreen was an "indoor" or "outdoor" item. They indicated their responses using the computer keyboard. Each image was on-screen for 2000 msec, with a 500 msec inter-stimulus interval (ISI). Responses were only recorded if the stimulus was on-screen.

All images of objects were shown once or twice across the entire task (organized into two randomized blocks with repeated images shown once in each block). And the images were embedded in noise by randomly scrambling a percentage of their pixels. We varied the amount of noise embedded in each image to be low (40%) or high (70%). This created six memory quality conditions: a single presentation in low-noise (Single Low), a single presentation in high-noise (Single High), when both presentations of an object were embedded in high-noise (High-High), when both

Downloaded from clal learnmem.cshlp.org on November 5, 2026 - Published by Cold Spring Harbor Laboratory Press

presentations were embedded in low-noise (Low-Low), when the first presentation was embedded in low-noise and the second presentation was in high-noise (Low-High), and when the first presentation was in high-noise and the subsequent presentation was in low-noise (High-Low). Participants completed a total of 600 trials (60 each for images shown once, 120 each for images shown twice). There were no breaks during encoding, for a total encoding session time of ~25 min. Figure 1 provides a schematic of the stimuli and procedure.

During each trial of the second phase (surprise retrieval), a previously viewed object from encoding was paired with a new object, and participants reported the one that was "old," that is, the one that appeared at some point in the encoding phase. For half of the trials the new image was a completely new, categorically distinct object (New test) and for the other half of trials the new images was a categorically similar-looking object (Similar test). Images were presented on-screen until participants made a response.

Experiment 2

A new group of 30 Johns Hopkins University undergraduates participated in Experiment 2. All participants reported normal or corrected-to-normal visual acuity. Participation was voluntary, and in exchange for extra credit in related courses. The experimental protocol was approved by the Johns Hopkins University IRB.

All methods were identical to those in Experiment 1, with the following exceptions: At encoding, participants judged whether the object was more "square" or "round." This change was made because the stimulus set used included an overwhelming number of objects that would have been rated as "indoor" objects. Participants viewed a total of 152 color images of objects without noise (taken from Geusebroek et al. 2005). Each object was shown twice (again organized into two randomized blocks, unbeknownst to participants), but either from the same orientation (Consistent Encounters condition) or two different orientations (Variable Encounters condition). There were no breaks during encoding, for a total encoding session time of ~12.7 min.

During the surprise test, participants were only given New tests. However, the image of the old object was either the same as encountered during encoding (Original Orientation Test), or the object shown from a completely new orientation (New Orientation Test).

Experiment 3

A new group of 31 Johns Hopkins University undergraduates participated in Experiment 3. The results of three participants were excluded due to responding randomly (i.e., chance performance). All participants reported normal or corrected-to-normal visual acuity. Participation was voluntary, and in exchange for extra credit in related courses. The experimental protocol was approved by the Johns Hopkins University IRB.

All methods were identical to those in Experiment 1, with the following exceptions: At encoding, participants viewed a total of 360 color images of objects without noise. Each image was shown either once (120 images), twice after a brief delay (120 images, 2-back), or twice after a longer delay (120 images, 11-back), for a total of 600 trials during encoding. All stimuli were presented continuously, in a pseudo-random order (to ensure images were presented in either 2-back, 11-back, or single presentation conditions equally). There were no breaks during encoding, for a total encoding session time of 25 min.

Acknowledgments

This research was funded by NSF BCS-1534568 to J.I.F. Experiment 1 was preregistered with aspredicted.org, available permanently at https://aspredicted.org/mw3tx.pdf. Experiment 2 was preregistered with aspredicted.org, available permanently at https://aspredicted.org/5j7s4.pdf. Experiment 3 was preregistered with aspredicted.org, available permanently at https://aspredicted.org/zbbkd.pdf - the 10-back was coded as 10 intervening items, which is why it is referred to as an 11-back in the present manuscript. A special thanks to Dr. Jack Schurgin, whose memory continues to inspire our research.

References

Alberini CM, LeDoux JE. 2013. Memory reconsolidation. *Curr Biol* **23:** R746–R750.
Andreopoulos A, Tsotsos JK. 2013. 50 years of object recognition: directions forward. *Comput Vis Image Underst* **117:** 827–891.
Biederman I. 1987. Recognition-by-components: a theory of human image understanding. *Psychol Rev* **94:** 115.
Biederman I, Ju G. 1988. Surface versus edge-based determinants of visual recognition. *Cogn Psychol* **20:** 38–64.
Brady TF, Konkle T, Alvarez GA, Oliva A. 2008. Visual long-term memory has a massive storage capacity for object details. *Proc Natl Acad Sci* **105:** 14325–14329.
Brady TF, Konkle T, Oliva A, Alvarez GA. 2009. Detecting changes in real-world objects: the relationship between visual long-term memory and change blindness. *Commun Integr Biol* **2:** 1–3.
Brady TF, Konkle T, Alvarez GA. 2011. A review of visual memory capacity: beyond individual items and toward structured representations. *J Vis* **11:** 4–4.
Brady TF, Konkle T, Gill J, Oliva A, Alvarez GA. 2013. Visual long-term memory has the same limit on fidelity as visual working memory. *Psychol Sci* **24:** 981–990.
Brainard DH. 1997. The psychophysics toolbox. *Spatial Vis* **10:** 433–436.
Cansino S, Maquet P, Dolan RJ, Rugg MD. 2002. Brain activity underlying encoding and retrieval of source memory. *Cereb Cortex* **12:** 1048–1056.
Cousineau D. 2005. Confidence intervals in within-subject designs: a simpler solution to Loftus and Masson's method. *Tutor Quant Methods Psychol* **1:** 42–45.
Cowan N. 2008. What are the differences between long-term, short-term, and working memory? *Prog Brain Res* **169:** 323–338.
Cox DD, DiCarlo JJ. 2008. Does learned shape selectivity in inferior temporal cortex automatically generalize across retinal position? *J Neurosci* **28:** 10045–10055.
Cox DD, Meier P, Oertelt N, DiCarli JJ. 2005. 'Breaking' position-invariant object recognition. *Nat Neurosci* **8:** 1145–1147.
DiCarlo JJ, Cox DD. 2007. Untangling invariant object recognition. *Trends Cogn Sci* **11:** 333–341.
DiCarlo JJ, Zoccolan D, Rust NC. 2012. How does the brain solve visual object recognition? *Neuron* **73:** 415–434.
Gauthier I, Williams P, Tarr MJ, Tanaka J. 1998. Training 'greeble'experts: a framework for studying expert object recognition processes. *Vis Res* **38:** 2401–2428.
Geusebroek JM, Burghouts GJ, Smeulders AW. 2005. The Amsterdam library of object images. *Int J Comput Vis* **61:** 103–112.
Green DM, Swets JA. 1966. *Signal detection theory and psychophysics, Vol.* 1. Wiley, New York, NY.
Guerin SA, Robbins CA, Gilmore AW, Schacter DL. 2012. Retrieval failure contributes to gist-based false recognition. *J Mem Lang* **66:** 68–78.
Hintzman DL. 1976. Repetition and memory. *Psychol Learn Motiv* **10:** 47–91.
Isik L, Leibo JZ, Poggio T. 2012. Learning and disrupting invariance in visual recognition with a temporal association rule. *Front Comput Neurosci* **6:** 37.
Kanwisher NG. 1987. Repetition blindness: type recognition without token individuation. *Cognition* **27:** 117–143.
Kennedy KM, Rodrigue KM, Raz N. 2007. Fragmented pictures revisited: long-term changes in repetition priming, relation to skill learning, and the role of cognitive resources. *Gerontology* **53:** 148–158.
Konkle T, Brady TF, Alvarez GA, Oliva A. 2010. Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *J Exp Psychol Gen* **139:** 558.
Li N, DiCarlo JJ. 2008. Unsupervised natural experience rapidly alters invariant object representation in visual cortex. *Science* **321:** 1502–1507.
Macmillan NA, Creelman CD. 2004. *Detection theory: A user's guide*. Psychology Press, New York.
Melton AW. 1970. The situation with respect to the spacing of repetitions and memory. *J Verbal Learn Verbal Behav* **9:** 596–606.
Nader K, Hardt O. 2009. A single standard for memory: the case for reconsolidation. *Nat Rev Neurosci* **10:** 224.
Pelli DG. 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vis* **10:** 437–442.
Pinto N, Cox DD, DiCarlo JJ. 2008. Why is real-world visual object recognition hard? *PLoS Comput Biol* **4:** e27.
Poth CH, Schneider WX. 2016. Breaking object correspondence across saccades impairs object recognition: the role of color and luminance. *J Vis* **16:** 1.
Poth CH, Herwig A, Schneider WX. 2015. Breaking object correspondence across saccadic eye movements deteriorates object recognition. *Front Syst Neurosci* **9:** 176.

Downloaded from clal learnmem.cshlp.org on November 5, 2026 - Published by Cold Spring Harbor Laboratory Press

Potter MC. 1976. Short-term conceptual memory for pictures. *J Exp Psychol Hum Learn Mem* **2:** 509–522.
Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I. 2005. Invariant visual representation by single neurons in the human brain. *Nature* **435:** 1102–1107.
Ranganath C, Cohen MX, Brozinsky CJ. 2005. Working memory maintenance contributes to long-term memory formation: neural and behavioral evidenc. *J Cogn Neurosci* **17:** 994–1010.
Ray CA, Reingold EM. 2003. Long-term perceptual specificity effects in recognition memory: the transformed pictures paradigm. *Can J Exp Psychol* **57:** 131.
Reagh ZM, Yassa MA. 2014. Repetition strengthens target recognition but impairs similar lure discrimination: evidence for trace competition. *Learn Mem* **21:** 342–346.
Rust NC, Stocker AA. 2010. Ambiguity and invariance: two fundamental challenges for visual processing. *Curr Opin Neurobiol* **20:** 382–388.
Schurgin MW. 2018. Visual memory, the long and the short of it: a review of visual working memory and long-term memory. *Atten Percept Psychophys*. doi: 10.3758/s13414-018-1522-y
Schurgin MW, Flombaum JI. 2015. Visual long-term memory has weaker fidelity than working memory. *Vis Cogn* **23:** 859–862.
Schurgin MW, Flombaum JI. 2017. Exploiting core knowledge for visual object recognition. *J Exp Psychol Gen* **146:** 362–375.
Schurgin MW, Flombaum JI. 2018. Visual working memory is more tolerant than visual long-term memory. *J Exp Psychol Hum Percept Perform*. doi: 10.1037/xhp0000528.
Schurgin MW, Reagh ZM, Yassa MA, Flombaum JI. 2013. Spatiotemporal continuity alters long-term memory representation of objects. *Vis Cogn* **21:** 715–718.
Shepard RN. 1967. Recognition memory for words, sentences, and pictures. *J Verbal Learn Verbal Behav* **6:** 156–163.
Singh I, Oliva A, Howard M. 2017. Visual memories are stored along a compressed timeline. *bioRxiv*: 101295.
Smolen P, Zhang Y Byrne JH. 2016. The right time to learn: mechanisms and optimization of spaced learning. *Nat Rev Neurosci* **17:** 77–88.
Squire LR. 2004. Memory systems of the brain: a brief history and current perspective. *Neurobiol Learn Mem* **82:** 171–177.
Squire LR, Wixted JT. 2011. The cognitive neuroscience of human memory since HM. *Annu Rev Neurosci* **34:** 259–288.
Srinivas K. 1995. Representation of rotated objects in explicit and implicit memory. *J Exp Psychol Learn Mem Cogn* **21:** 1019.
Srinivas K, Verfaellie M. 2000. Orientation effects in amnesics' recognition memory: familiarity-based access to object attributes. *J Mem Lang* **43:** 274–290.
Standing L. 1973. Learning 10000 pictures. *Q J Exp Psychol* **25:** 207–222.
Suzuki M, Johnson JD, Rugg MD. 2011. Decrements in hippocampal activity with item repetition during continuous recognition: an fMRI study. *J Cogn Neurosci* **23:** 1522–1532.
Tarr MJ, Hayward WG. 2017. The concurrent encoding of viewpoint-invariant and viewpoint-dependent information in visual object recognition. *Vis Cogn* **25:** 100–121.
Wagar BM, Dixon MJ. 2005. Past experience influences object representation in working memory. *Brain Cogn* **57:** 248–256.
Wallis G, Bülthoff HH. 2001. Effects of temporal association on recognition memory. *Proc Natl Acad Sci* **98:** 4800–4804.
Wiggs CL, Weisberg J, Martin A. 2006. Repetition priming across the adult lifespan—the long and short of it. *Aging Neuropsychol Cogn* **13:** 308–325.
Yassa MA, Stark CE. 2011. Pattern separation in the hippocampus. *Trends Neurosci* **34:** 515–525.
Yonelinas AP, Aly M, Wang WC, Koen JD. 2010. Recollection and familiarity: examining controversial assumptions and new directions. *Hippocampus* **20:** 1178–1194.
Yu SS, Johnson JD, Rugg MD. 2012. Hippocampal activity during recognition memory co-varies with the accuracy and confidence of source memory judgments. *Hippocampus* **22:** 1429–1437.

Received November 30, 2017; accepted in revised form April 27, 2018.

Downloaded from clal learnmem.cshlp.org on November 5, 2026 - Published by Cold Spring Harbor Laboratory Press



Properties of visual episodic memory following repeated encounters with objects

Mark W. Schurgin and Jonathan I. Flombaum

Learn. Mem. 2018, **25:**
Access the most recent version at doi:10.1101/lm.047167.117

References This article cites 57 articles, 6 of which can be accessed free at:
http://learnmem.cshlp.org/content/25/7/309.full.html#ref-list-1

Creative Commons License This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first 12 months after the full-issue publication date (see http://learnmem.cshlp.org/site/misc/terms.xhtml). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

© 2018 Schurgin and Flombaum; Published by Cold Spring Harbor Laboratory Press