

Noisy and hierarchical visual memory across time scales

Timothy F. Brady^{†1}, Maria M. Robinson¹, Jamal R. Williams¹

¹Department of Psychology, University of California San Diego, La Jolla, CA, USA

[†]email: tfbrady@ucsd.edu

Abstract

Both in everyday life and in memory research, one tends to think of items as being ‘held’ in mind, in the same way that a real-world object can be held in one’s hand. Inspired by this metaphor, traditional work on visual working memory and visual long-term memory focuses on understanding how many objects are remembered or forgotten, or held or lost in particular circumstances. By contrast, newer computational and empirical work on visual memory focuses on the role of noise in memory representations — conceiving of memories as continually varying in ‘strength’ or ‘precision’ — as well as the role of the visual hierarchy and priors in structuring memory. In this Review, we merge these contemporary theories and evidence. We describe how fundamentally noisy memory representations are instantiated at different levels of the visual hierarchy and support both visual-working and long-term memory. We also discuss how conceiving of memory in this way can direct further research and illuminate the nature of cognitive function more broadly.

[H1] Introduction

Consider what your dinner plate looked like last night, including the color, size, and positions of food on the plate. When you try to bring these images to mind, you are trying to retrieve the details of a particular visual experience, using your visual memory. Visual memory refers to information that can be successfully accessed from past visual experiences, and the structure of that information.

Visual memory is of broad interest because of its impressive capabilities, which can be appreciated in the method of loci, widely used to turn arbitrary words or playing cards into rich visual images by the best memorizers in the world¹. Visual memory is also of broad interest because it provides key insights into the nature of memory representations. While traditional memory research has often focused on identifying separate memory systems and on establishing a taxonomy of memory systems and their function^{2,3}, and focused less on examining the nature of stored memory representations. By contrast, research on visual memory places a major emphasis on the structure and format of the representations of objects and scenes in memory, and how properties of these representations impose limits on what information can be remembered accurately^{4,5}. Because the perceptual representations underlying visual processing are (comparatively) well understood compared to many other domains^{6–8}, research at the intersection between visual perception and memory can provide unique insights into memory processes by building on what is known about perceptual representations and examining how they change when held in memory.

A key insight is that visual memory representations are noisy – that is, that people can never remember them perfectly accurately, and that the accuracy of their memory tends to vary in a continuous manner. This insight derives in part from visual memory research's emphasis on visual features that themselves vary continuously, such as color and orientation, and the metrics of performance that asking people to remember these features allows in visual memory tasks. However, the finding that visual memory representations are noisy conflicts with the tendency to use a physical metaphor to describe and understand memory. That is, people tend to think of an object they are trying to remember as either in mind or not in mind, and to talk about items as being 'held' in mind to describe working memory, just like a real object can be held in the hand⁹. This physical metaphor often serves as a core mental model for how memory has traditionally been conceived: as all-or-none, discrete, and operating over entire objects or chunks of information.

Given their continuous and noisy nature, however, the idea that memory representations are physical objects that can be held, or can be defined as simply existing or not existing, is not a useful metaphor for memory. Therefore, instead of using this physically inspired mental model for memory, here we propose a model based on an analogy to the nature of processing in the visual system (Fig. 1). In particular, the visual system is usually conceived of in terms of population codes, with many neurons representing a given feature and a given location— and each doing so with substantial noise¹⁰. Furthermore,

the visual system contains a hierarchy of representations, where representations are richer and more complex as one moves from primary visual cortex up to more anterior visual regions.

In this Review, we detail how we conceive of visual memory in a similar way, as a noisy and hierarchical system. We begin by reviewing work that suggests that representations for objects or scenes are strongly related across visual working memory and visual long-term memory, and in both cases are best conceived of as noisy and variable in strength. We then describe how such noisy memories are stored in terms of hierarchical memory representations. Finally, we discuss how these ideas lead to a reconceptualization of the limits of visual memory at different timescales, and the role of visual memory in other cognitive capacities.

[H1] Visual memory across time scales

We begin with the question of whether it makes sense to consider visual memory as a single construct, or, whether the more traditional view, where long-term memories are considered fundamentally distinct from working memories, is most appropriate. As in more traditional memory research, visual working memory and visual long-term memory are often studied separately, using different paradigms, by different researcher groups, and are often thought of as distinct systems. Although we acknowledge that a systems taxonomy can have heuristic value, this view is grounded in theories which posit that visual working memory and visual long-term memory differ in their mental representations and intrinsic structure. By contrast, contemporary evidence suggests that visual working memory and visual long-term memory share mental representations, even though varying task demands can lead to the differential recruitment of additional cognitive processes that operate over those representations.

Mainstream theories of memory often postulate that visual working and long-term memory representations differ from each other in key ways. One reason for this is distinctive neural correlates. For example, neuropsychological research suggests a distinction between visual long-term and working memory based on purported evidence of selective and distinct impairments to visual long-term and visual working memory^{11,12}. It has also been suggested that short-term memory reflects transitory electrical activity, whereas long-term memory reflects durable neurochemical changes in the brain¹³. Broadly in line with this view, some evidence indicates there are distinct neural signatures associated with visual working memory and long-term memory, with the sustained and active firing of neurons involved particularly in working memory maintenance^{14,15} but not long term memory. In humans, such working memory-specific activity can be tracked with electrophysiology, for example with the contralateral delay activity^{16–19}, a robust component thought to track the amount of information ‘actively’ stored in visual working memory²⁰ and that disappears when items are available in long-term memory²¹. Similarly, information in working memory can be seen in parietal and frontal region activity⁹ and decoded from ongoing activity during the maintenance period using neuroimaging^{22–24}.

However, these neural coding differences do not necessitate differences in the fundamental representations or intrinsic structure underlying working memory and long-term memory. For example, the neuropsychology work must be interpreted with caution based on the observation that participants from clinical populations suffer from a wide range of processing deficits that make claims of selective deficits to a specific cognitive system difficult²⁵. Similarly, the evidence of distinct neural correlates for working memory, involving active maintenance processes, do not necessarily indicate distinct representations are used in working memory compared to long-term memory. Instead, the distinction between typical working memory and long-term memory tasks can reflect the amount of activation of a particular representation²⁶⁻²⁹ or the engagement of attentional maintenance processes to maintain such activation, with the representations themselves being fundamentally the same. For example, the classic modal model of memory posited that short-term or working memory exists as a buffer that maintains information in a readily accessible state^{30,31}. It is therefore often posited that working memory is capacity limited such that only a small amount of information can be maintain in an active state at once, perhaps due to limitations in attentional maintenance processes³². Thus, rather than requiring two systems, the neural evidence is consistent with the view that representations in working memory and long-term memory are fundamentally the same, but long-term memory is mostly passive use of these representations and working memory is a form of attentional demanding, activated memory, involving control systems that manipulate and maintain the same underlying memory representations. According to this account, once activated in working memory, a subset of information related to pertinent representations can be continuously protected from noise accumulation by attentional processes, but the representations remain fundamentally similar to those used in long-term memory.

Behavioral evidence is consistent with this account. Despite a long history of evidence that suggests long-term memory is often less precise than working memory (such as theories that suggest only the gist survives in long-term memory, causing false memories³³)—new evidence shows that there is not any fundamental difference in the range of levels of precision or memory strength that are possible in visual-working and long-term memory. Distributions of memory errors obtained from continuous reproduction visual-working and long-term memory tasks³⁴ fit with a one parameter model that captures memory strength across tasks³⁵ (Fig. 2). This suggests that as memory gets weaker in the two systems, it degrades in an identical fashion. Of course, typical working memory and long-term memory tasks may differ in key ways, even with the same representations underlying them. For example, online attentional selection might often be more necessary in working memory tasks because many items are often presented at once in such tasks, and there is often less spatiotemporal context in long-term memory tasks compared to working memory tasks, with many items all presented in the same spatial location being relevant. Thus, although long-term memories on average might be weaker than working memories in everyday-life situations, evidence suggests that this difference strength does not necessarily reflect a fundamental difference in memory

systems or their representations but rather the way they tend to be used and investigated.

Thus, evidence is consistent with a view that visual-working memory is a form of activated long-term memory, with similar representations that differ in their degree of activation or in the processes being used to manipulate and maintain them. Through this lens, attention ‘selects’ and upweights goal-relevant information, and provides protection against the accumulation of noise across populations of neurons throughout the visual hierarchy³⁶, and memories protected in this way are working memories.

Prior work has criticized the unidimensional view of visual memory in part because the idea that memories are active in working memory (and less active in long-term memory) has not been consistently operationalized. For instance, activation has been used to describe both the strength of mental representations, as well as dynamics in neural processing. Furthermore, the conception that activation alone differs between working memory and long-term memory is generally insufficient to explain the differences across working memory and long-term memory tasks. For example, many theories treat activation levels as varying even among different long-term memories, and some work suggests that working memory can make use of inactive states^{37,38} in addition to active ones. This notion of ‘activation’ leads to theories arguing that up to a hundred or more recently seen items can be active in long-term memory³⁹, as compared to ‘inactive’ items that are seen equally often but are not relevant in the current context⁴⁰. Reconciling all the notions of activation across both neural and cognitive meanings of the term and across both working memory and long-term memory remains a difficult and important task for unidimensional theories of memory, and one that has been the source of a number of proposals about the architecture of working memory^{32,41–44}.

In summary, within this unidimensional view of visual memory, visual working and visual long-term memory are not distinct systems, but instead both involve computations over the same types of memory representations. Importantly, the nature of these computations is constrained by the demands of the task, such that some visual memory tasks elicit processing constraints which the others do not. In the remainder of this article, we develop this unidimensional framework by integrating findings from the visual working and long-term memory literature on the nature of visual memory representations. We then detail how differential demands can elicit processing bottlenecks in visual memory tasks. Finally, we discuss how the unidimensional view fits parsimoniously with more ecological theories that take into account the role of prior knowledge, stimulus complexity and meaning in visual memory across time scales.

[H1] Memories are noisy and vary in strength

A major dichotomy in the visual memory literature has been between views in which memories are conceptualized as all-or-none—with bound objects that are either remembered with complete accuracy or fully lost^{45,46}—and views in which memory

representations are noisy, hierarchically structured, and distributed—with object features that degrade continuously and relatively independently^{47–49}. This question of how to construe the fundamental properties of visual memory representations have dominated the visual working memory literature for the last couple of decades, and physical analogies based on holding items in mind have long pushed researchers toward more discrete views of memory.

In the following sections, we describe behavioral, neural, and computational evidence that visual memory representations are noisy, which is synonymous with the view that they vary in strength. We also discuss how noise accumulates over short and long time-scales and how memory representations can vary at different levels of abstraction: from individual features to entire objects, as well as how the graded nature of memory representations underscores the role of decision making in mainstream memory tasks.

[H2] Graded memory strength in working memory tasks

The question of whether memory is noisy vs. all-or-none has been most often examined in the domain of short-term, working memory tasks. All-or-none theories of memory in which items are either present (and perfectly represented) or absent entirely, have traditionally dominated the analysis of visual working memory data from change detection tasks (Fig. 3a). In a typical change detection task, people are shown an array of simple stimuli that they have to remember over a brief delay. After a retention interval, people are shown a single stimulus and asked to make a discrete judgement regarding whether it changed or did not change compared to the stimulus shown in that location in the original memory array. The discrete nature of “change” or “no change” responses provide only a coarse measure of memory errors and often incorrectly leads to inferences like items are either “present” or “absent” in memory⁵⁰.

By contrast, most contemporary work uses tasks which yield a more fine-grained measure of visual memory representations. Continuous reproduction tasks (Fig. 3b) provide a way of assessing gradations in memory representations and are common in visual working memory. In such tasks, instead of providing a binary response^{51–53}, people are asked to reproduce their memory of a probed feature or object using a continuous response wheel, which provides fine-grained information about both the magnitude and direction of memory errors. In these tasks, responses generate a distribution of memory errors in a single, continuous feature dimension (such as color or orientation). For instance, like in the change detection task, people are shown a memory array with simple stimuli and are instructed to remember these stimuli and their locations over a brief retention interval. After the memory delay, however, people are shown a spatial probe along with a continuous wheel that shows all possible values in the continuous feature space. Accordingly, people are instructed to select the feature value they think most closely matches the value of the probed item. Therefore, unlike change detection tasks, the continuous reproduction wheel provides a fine-grained measure of people’s memory. Similar fine-grained performance metrics can also be arrived at when using forced-choice recognition tests that vary the similarity of the seen and unseen items^{35,54}.

A first major insight from continuous reproduction tasks is that the distribution of errors in visual working and long-term memory tasks are incompatible with fully all-or-none theories of memory (Fig. 3c). These tasks reveal that noise (variance) of memory errors increases with memory load⁴⁹ and delay^{35,48}, and decreases with more encoding time³⁵, as well as repeated study, and do so in a way that is not solely due to all-or-none failures³⁴. These results challenge a fully all-or-none interpretation because they indicate that simple task manipulations can expose the graded changes that exist in memory strength and accuracy.

A second major insight from tasks that allow continuous measurement of memory accuracy is that some task conditions introduce subtle memory biases, such that similar representations sometimes repulse towards or attract away from one another^{55–60}. For instance, when people are presented with many red items that slightly vary in hue the representations of these items can be pulled or attracted towards their average hue. Alternatively, when people are presented with just a few similar red items, their representations can become more distinct and repel from one another (Fig. 3c). These results imply that items are not represented and recalled independently of each other, but are represented in a way that causes subtle interactions between them. Both of these insights are broadly inconsistent with the typical physical analogy of holding discrete items in mind and therefore with all-or-none theories of memory representation.

Thus, all contemporary theories of visual working memory have moved beyond the concept of complete discreteness in memory, broadly agreeing that information in memory is represented imperfectly and that this imprecision varies with a number of factors such as working memory load, encoding time, delay, and several others. Most neural and cognitive theories broadly endorse the view that there is internal noise in perceptual, attentional and memory processes, as well as sources of external noise from the environment. These sources of noise can be affected by how attention is distributed during encoding⁶¹, memory load⁶², retention intervals⁶³. All of these sources jointly determine the extent to which fluctuations in internal and external noise impact the fidelity of memory representations.

While nearly all models agree that memories must be considered noisy and imprecise, an increasingly large number of theories of visual working memory now subscribe to the view that noise accumulation among items alone is fully sufficient to explain the limits on performance and the pattern of errors obtained in such tasks; without any additional factor of all-or-none loss of items^{35,64,65}. However, some hybrid ‘mixture models’ include additional factors that postulate that continuously graded memories exist, but also that all-or-none loss of items can occur in some conditions^{66,67}. These models posit that some items that were seen and processed by the visual system while they were visually present are nonetheless completely unavailable and that observers have no remaining information about them, even less than 1 second later, such that when people are probed on those items, they are forced to give completely information-less guesses^{66,67}. Although such hybrid models were initially influential because they were unique in providing an account of the distribution of errors in reproduction tasks⁶⁸, later work showed that

continuous models can parsimoniously account for these distributions of errors just as well⁶⁹. For example, the finding that people tend to have little information about the 5th or 6th items when asked to report a set of 6 briefly-shown items follows naturally from models without all-or-none failures⁶⁹. Thus, modern models without any added assumptions about all-or-none failures generally account for data as well or better than such hybrid models^{35,64,65}. Moreover, hybrid models that attempt to predict how often pure, informationless ‘guesses’ arise when an item is not in memory, across different set sizes (e.g., how many putative “guesses” there are when shown 4 vs 6 items)⁶⁸, generally fail to accurately predict the full pattern of errors as memory load increases⁷⁰. Thus, modern hybrid models generally cannot sufficiently explain performance across set size, as they do not tend to provide any quantitative account of how often all-or-none losses are expected to occur across different conditions or set sizes, which was initially considered a major strength of hybrid models (for further discussion, see⁷¹). Overall, current findings about how visual working memory varies as a function of different variables, such as memory load and delay, can be parsimoniously explained with a unidimensional view according to which memory representations vary in noisiness, and do not require postulating additional assumptions about all-or-none coding that are necessary in hybrid models.

[H2] Noise across time

If working memory and long-term memory representations are fundamentally similar, we would expect models that focus on long-term memory rather than working memory to also be focused on understanding memory noise and imprecision. However, one notable difference between traditional theories of visual working and long-term memory is how they conceptualize noisy memory representations. Theories of visual long-term memory typically conceive of noise as affecting the ‘strength’ of memory representations rather than the noisiness or precision of those memories. This construct of memory strength follows from a long tradition of signal detection theory approaches for measuring memory^{72,73}. Within the signal detection framework, previously seen and unseen items give rise to two different distributions of familiarity signals. When asked to make a memory judgment (such as which of two items you saw), familiarity signals that can arise from comparing memories to stimuli along different dimensions are collapsed into a unidimensional continuum of signal strength, with seen items having higher average signal strength than unseen items⁷⁴. Accordingly, the standardized distance between the seen and unseen distributions (traditionally referred to as ‘noise’ and ‘signal+noise’ distributions are often captured by the signal detection measure d') is assumed to reflect the relative ‘strength’ of remembered items (i.e., how much more familiar, on average, previously seen items feel compared to previously unseen items).

By contrast, most visual working memory models, as described in the previous section, conceive of noise in memory representations in terms of precision (e.g., how tight the errors are when asked to reproduce an item) rather than strength^{66,67,75}. The construct of visual memory precision follows from characteristics of the continuous reproduction task.

In this task, the variance of the error distribution around the true item value is typically interpreted as a measure of how precisely a given feature or object was remembered⁵¹.

The constructs of memory strength and precision ultimately rest on the same fundamental assumption: that memory representations are noisy and therefore vary in their match to the information that was actually seen. Additionally, the two constructs can be formally linked by using signal detection theory with an entire population of signals³⁵ (Box 1). Importantly, how confident someone is in their memory tracks both precision and strength of memory representations, which provides another empirical bridge between the two constructs. In the long-term memory domain, it is generally found that confidence tracks memory performance extremely well, and that such confidence judgments are generally unaffected by demand characteristics⁷⁶. Similarly, in working memory, confidence also tracks performance accurately even in continuous feature spaces^{77,78}, and this relationship can be naturally accounted for by models that relate precision to confidence^{64,69} or those that use signal-detection based measures of strength to predict confidence³⁵.

Other measures that serve as an index of confidence — for example, asking participants to report the full range of feature values they think an object might have had rather than a single value^{78,79} — also suggest that participants have an internal sense of uncertainty that tracks with memory performance in working memory tasks, similar to the relationship between confidence and accuracy that is fundamental to measures of performance in long-term memory tasks. Collectively, this suggests that there might be a direct mapping between people's confidence judgments and a latent, continuous memory signal that is used on both visual working and long-term memory tasks. Together, confidence judgments and other judgments of uncertainty^{50,80} can be seen as providing convergent support for the view that precision and strength ultimately describe the same fundamental properties of visual memory representations, which is that visual memory representations are noisy and vary continuously in both working memory and long-term memory tasks.

[H2] Noise across multiple features

The view that visual memory representations for single features, like color or orientation, are noisy raises the question of how to think about items that consist of a conjunction of multiple features such as shape, orientation, and color. Although this 'binding problem'⁸¹ has traditionally been framed in discrete terms (e.g., are both features remembered or is just one forgotten?), in the context of noisy memory representations the question is whether features accrue noise independently or jointly or both. This question continues to be actively debated, but the latest evidence suggests that features seem to accrue noise largely independently⁸². For instance, memory errors in a continuous reproduction task are largely uncorrelated across feature dimensions^{83,84}. Likewise, independent-feature resource models, which postulate that a separate set of resources are used to support memory for different features, such as color and orientation, outperform shared resource

models, where color and orientation compete for the same pool of resources, in a version of the change detection paradigm⁸⁵.

Evidence for independent noise accrual of features has also been reported for real-world objects. People commit more binding errors (swapping features such as color across objects) when state (object configuration, like a full or empty mug of coffee) and exemplar (category, like mug or wallet) are jointly manipulated, suggesting the two features can be forgotten or misbound separately from each other⁴⁷. This is consistent with the finding from continuous reproduction tasks that different stimulus dimensions accrue noise independently, and provides convergent support for the view that object features are stored at least somewhat independently and not as all-or-none, holistic units that accrue noise holistically and are forgotten holistically. Similarly, in long-term memory tasks, different object features also seem to be represented at least to some degree independently⁸⁶.

Despite this evidence for independent noise accrual across stimulus dimensions, visual working memory studies do tend to find a clear memory benefit when visual features are part of the same object⁸⁷. For example, it is easier to remember two features when these are bound to the same object (such as, color and orientation), compared to separate objects⁸⁵. This benefit of objecthood might reflect the benefits of attending to a single as opposed to multiple spatial locations⁸⁸. However, classic work on visual attention also reveals evidence for benefits of object-based attention mechanisms, even when spatial distance is controlled⁸⁹ (for a critical analysis of object-based attention effects see⁹⁰). Collectively, the joint finding for independent noise accrual and spatial and object-based benefits for memory for features is consistent with the view that different features generally accumulate noise independently, but spatial and object-based attentional mechanisms can introduce additional, correlated sources of noise accrual during encoding or modulate external and internal noise accrual across feature dimensions.

[H2] Decision making under noise

A critical implication of the view that memories are always noisy and imprecise is that people must make probabilistic inferences based on noisy evidence to decide how to respond on memory Tasks^{91,92}. Thus, unlike in traditional all-or-none views, according to which people make memory judgments in a straightforward way by reading out whether an item is or is not in memory, within a continuous memory framework people have to use additional criteria for reporting whether they do or do not remember an item well enough to endorse it, and, accordingly, researchers must always consider the decision-making processes that underpin memory tasks. This point has often been made previously in the context of recognition memory tasks^{77,93–95}, such as old/new and change detection tasks in long-term and working memory research, respectively. These tasks are frequently mis-used and mis-interpreted because the decision process is not carefully considered⁷⁷. Decision-making processes are integral in such tasks because each stimulus elicits a noisy familiarity signal, people must set a decision criterion for responding whether an item is old or new, or whether a change occurred or did not

occur⁸⁰ and variations in people's criterion setting can lead to large differences in measures like overall accuracy without reflecting any underlying change in the memory representation itself.

The study of decision-making processes in visual memory tasks is especially central in more applied settings, like eyewitness memory tasks in which people must select the face of the guilty suspect amongst a lineup of faces or report that the guilty suspect is not part of the lineup⁹⁶. Much work suggests that in such tasks people do not simply process each face independently, but might use higher-level decision-making strategies, such as discounting features that are common to all faces⁹⁷. The development of computational models to capture such strategies is extremely important to help dictate how to construct lineups that increase the likelihood that the guilty suspect is correctly identified and decrease the likelihood that an innocent suspect is incorrectly identified.

Decision models have been rigorously applied in many visual memory tasks beyond eyewitness memory^{98–101}. For instance, in one study researchers applied sequential sampling models to examine the decision processes that underpin performance in continuous reproduction tasks¹⁰². This modeling work provided insight into how memory precision relates to a noisy process of evidence accumulation in perception and memory that underpins memory based decisions.

In sum, the view that visual memory representations are fundamentally noisy highlights that both short and long-term memory representations are continuous and share the same fundamental properties (Fig. 2). These assumptions apply to memory representations at different levels, from features to objects, and fits parsimoniously with neuro-computational population coding theories of perception and memory. Finally, the continuous representation view highlights that people must have ways of using noisy evidence to make memory-based judgments and actions, which highlights the fundamental role of decision-making processes in laboratory and real-world memory tasks.

[H1] Capacity limits

Given evidence that memory representations are continuous, a critical research goal has been to characterize limits in visual memory through a continuous framework. The working memory and long-term memory literatures have taken different approaches to this question. Whereas in long-term memory much work has accepted that no single fixed capacity limit can be found because the effect of interference between items and the role of retrieval cues will be complex and stimulus dependent, the working memory literature has been much more focused on attempts to find a fixed resource limit that explains performance in simplified settings.

One major difference between visual working and long-term memory is often framed as differences in their putative capacities. Unlike visual long-term memory, which is thought to be virtually unlimited, visual working memory is thought to be extremely limited, with

performance dropping quickly as more information must be maintained^{53,103}. While there is no doubt that there are limits on how much information can be actively maintained in working memory, especially due to limits on attention or maintenance processes that work to reduce noise in task relevant representations, most tasks designed to tap into working memory limits also artificially constrain memory capacity by intentionally employing situations in which the most useful memory cues are limited in their availability (due to interference or cue overload). In such tasks, one would expect a greater accumulation of memory-based noise than in tasks with reduced attentional demands, where there is less interference and more diagnostic retrieval cues. Since attentional mechanisms improve the quality of actively maintained information, these paradigms are informative about the capacity of attention and working memory maintenance processes, which is what they are designed to probe. However, they may not be very reflective of real-world uses of working memory, which will rarely occur under conditions designed to reflect solely the limits of attentional maintenance processes. Instead, many working memory tasks reveal primarily how accurately people can maintain information over short intervals only in quite unfavorable and unnatural circumstances. For example, typical visual working memory tasks involve the simultaneous flash of many meaningless, single feature items for less than a second⁵³, whereas by contrast, typical visual long-term memory tasks involve the sequential presentation of meaningful items for several seconds each⁵⁴. As we explain in detail in the following sections, the contrast between these tasks makes it unsurprising that performance in the former would be limited relative to the latter, and many purported differences between memory systems can instead be interpreted through a unidimensional view, where different visual memory tasks place differential demands on attention and provide different amounts of task relevant information that can be used to retrieve memories with higher fidelity.

[H2] Capacity in the lab versus the world

Differences in presentation format and stimuli across many working memory and long-term memory tasks introduce differential effects of spatial and temporal context that do not always reflect the true differences in everyday uses of shorter- and longer-term visual memory. However, there are some differences between typical working memory and long-term memory paradigms that may in fact reflect genuine differences in the real-world use of memory at short and long duration. For instance, spatiotemporal cues about an object—such as the spatial location in which it was shown and the context of other items around it—might be readily available in the environment in shorter-term but not longer-term memory tasks¹⁰⁴ because such cues tend to be available and stable only at short durations in the real-world (e.g., while pouring tea, my mug will likely stay in the same place, but over the course of a day it will likely change position quite often). These effects of context might differentially aid retrieval through external cues^{104,105} and introduce biases into memory representations^{57,59}.

However, other differences between lab-based working memory and long-term memory tasks are not necessarily typical of shorter- and longer-term visual memory in the world. As an example, differences in encoding demands between typical working memory and

long-term memory tasks are introduced if stimuli are presented simultaneously in several spatial positions or sequentially in the same position. For instance, a typical visual working memory task involves the simultaneous presentation of many objects, and people must distribute visual attention broadly amongst multiple items at once with no priority difference between the items and little time for elaborative encoding. By contrast, in sequential tasks, where objects are presented one at a time, like those typically used to assess long-term memory may be more representative of real-world situations where selective priority and elaborative encoding can take place. These and other differences in presentation format can drastically change estimates of working memory capacity for meaningful stimuli; people seem to have a higher capacity for meaningful objects presented serially¹⁰⁶ and under some conditions people seem to have no upper bound in how many objects can be recognized after serial presentation¹⁰⁷. Indeed, in contrast to prominent claims^{51,53,108,109} that visual working memory tasks are not encoding limited, but only limited by maintenance capacity: additional encoding time significantly enhances performance in working memory tasks even for objects with a single feature^{35,110,111}. Critically, in many realistic situations people might largely use working memory by sequentially focusing on a small number of items^{112,113} implying that visual working memory tasks where many items, all of equal relevance, are simultaneously presented at once for a brief duration may be less characteristic of real-world demands. In summary, it is important to consider the role of such attentional limits and differential encoding demands in visual working and visual long-term memory tasks, before attributing performance differences in these tasks to true differences in capacity or to core difference between distinct memory systems.

[H2] Visual working memory resources

In the working memory domain, many attempts have been made to formalize a single, limited capacity. This limited capacity view was popularized by the famous report that people seem to maintain only about 7 'chunks' of information¹¹⁴. The idea that working memory capacity is set in terms of discrete chunks or items dominated the traditional view of visual working memory capacity through the lens of 'slot' theories. According to these theories, the architecture of working memory is composed of a discrete number of slots that store objects composed of single or bound features and these slots are directly responsible for the limits of working memory capacity^{45,46,68,115–119}. This framework provides an intuitive way of thinking about capacity limits as a discrete number of items that can be 'held' in memory. Despite this intuitive appeal, strong versions of this view also construe memory representations as all-or-none, which cannot account for the extensive evidence that memory representations vary in strength or noise.

However, while there is overwhelming evidence that memory representations are continuous, it has proved conceptually challenging to provide a rigorous characterization of working capacity in terms of resources^{120,121}. Attempts to measure and define a single resource limit often hold many of the factors that would be expected to influence performance in working memory tasks fixed (such as encoding time, delay, contextual cues, and/or presentation format). Once fixed, they manipulate and measure how many

simultaneously presented simple items can be remembered. These tasks thus investigate a highly limited subset of what would be required to truly understand what constrains performance in tasks in which attention is used to protect items in memory over short delays. It is thus unclear if any putatively fixed resource limits are meaningful in more real-world contexts or across even small variations in task. Within these confines, there is substantial computational modeling work that is devoted towards developing concrete definitions of resources. For instance, some researchers define resources as ‘neural real estate’^{75,122} and suggest that visual memory representations are distributed across populations of neurons in the visual cortex. In such models, the main source of resource limits arises from the fact that a fixed firing rate is used to represent both a few and many items in memory (the basic concept of divisive normalization¹²³).

At the level of cognitive representations^{102,124}, researchers use principles from signal detection and sampling theory¹²⁵ to define resources as a limited number of ‘evidence samples’¹²⁶, which determine performance across variations in memory load. Although sampling theory makes more principled predictions than theories that simply postulate ‘resources’, the construct of samples is also poorly defined. It will therefore be important to ground out the idea of samples more rigorously, such as a new proposal that retrieval of representations from visual working memory reflect draws of samples from noisy spiking activity in a population of neurons⁷⁰. It will also be key to expand the understanding of resources beyond an understanding of how well people represent different numbers of simple objects flashed up briefly.

There has also been a resurgence of work using information-theoretic modeling approaches that has yielded new insights into properties of resource limited memory systems. These approaches frame memory systems in terms of information theoretic concepts like compression¹²⁷ or rate distortion theory¹²⁸. For instance, rate distortion theory, which aims to explain how a capacity limited system should store information, can be coupled with prior knowledge and task-relevance within deep neural network modeling¹²⁹ to simulate many fundamental aspects of people’s memory errors, such as biases towards the ensemble of remembered items⁵⁶. Broadly, this modeling approach falls in line with ‘resource rational’ models^{130,131} according to which people select computations that optimize outcomes but minimize the resource cost of implementing these computations. Such models have promise for understanding resources in a more general way, and across a broader set of tasks including visual working and long-term memory tasks.

As noted above, many of these computational modeling frameworks of capacity were developed using laboratory tasks that require short-term memorization of simple stimuli presented simultaneously. Such a simultaneous presentation format does not necessarily reflect how information is encoded over the short-term in real-world scenarios, where each item might generally be encoded serially and where one might often re-sample items rather than push the limits of working memory capacity^{112,132–134}. Indeed, estimates of visual working memory capacity when people encode meaningful, but not meaningless items serially rather than simultaneously increase substantially, likely because this format

provides time to connect online memory representations to existing knowledge¹⁰⁶. Even for simple stimuli, claims that encoding time was not a limit on performance^{51,53} have not held up to more psychophysical methods, which reveal that performance increases relatively smoothly with increasing encoding time, even with simple stimuli^{35,110,111}. These findings underscore that there are many experimental levers which can, and should, be used to investigate and build ecological models of visual memory processes across time (Box 2).

[H2] Attention and encoding limits

Another major source of capacity limits in both visual working and long-term memory tasks are due to attentional processes during the initial encoding of items rather than maintenance processes. This insight initially stems from research on change blindness: A phenomenon wherein people fail to detect changes in an image after a brief transient, such as an eye movement or a flicker of the display¹³⁵. Importantly, these changes can be large and appear in salient parts of the display, meaning that change blindness does not simply reflect failures of the perceptual system¹³⁶. Change blindness effects have been used to motivate theories of consciousness^{137,138} and working memory^{139,140} because they suggest that people remember only a small portion of what they see despite having a rich phenomenological experience of their immediate environment. Critically, one major way to nearly eliminate change blindness effects is to direct people's attention to objects^{127,141}, which ensures that objects are encoded into memory.

Other work suggests that encoding in visual working memory predicts the bandwidth of visual long-term memory. For instance, the effects of interference can be offset via attentional processes^{142–146} that induce repulsive biases at encoding and individuate similar memory representations in visual long-term memory^{57,147}. Other researchers have found that individual differences in performance on visual working memory tasks predicted performance on visual long-term memory tasks, but only under conditions in which visual working memory load is taxed¹⁴⁸. These results suggest that effects on visual long-term memory were due to failures of encoding rather than maintenance capacity, and underscores how controlling for task demands aligns with a unidimensional model of visual memory across timescales. Other work also provides converging support for the view that manipulations of attention can help upweight and improve long-term memory for objects¹⁴⁹. However, attentional processes might only be used to deprioritize irrelevant memory representations indirectly, under conditions of biased competition¹⁵⁰. More precisely, some evidence suggests that beneficial effects of attention occur only when multiple (two) stimuli were presented simultaneously and the irrelevant item was paired with an item than needed to be attended. This work directly connects to the finding that simultaneous as opposed to serial presentation in visual memory tasks can introduce attentional capacity limits and uncover structural bottlenecks¹⁵¹.

Thus, attentional limits at encoding are common to both visual working and long-term memory despite being generally understudied in visual long-term memory. Attentional limits at encoding, as well as online activation, fit well with the view that memory

representations are fundamentally noisy. These limits can therefore be viewed as one kind of resource that places limits on memory performance when encoding demands are high either in short or long-term visual memory tasks. More broadly, poor performance in most visual working memory tasks might reflect differences in encoding demands of visual working-memory and visual long-term memory laboratory tasks in addition to or instead of differential memory maintenance abilities or differential memory representations over different timescales per se. Thus, rather than viewing visual working and long-term memory as distinct systems that differ in capacity, a cognitive architecture in which attentional maintenance supports working activation of a subset of items that otherwise share the same underlying representation as long-term memory is a useful framework.

[H1] Influence of knowledge on visual memory

Visual memory is often assumed to be inherently perceptual in nature. However, existing knowledge and hierarchical knowledge structures help scaffold memory of real-world visual stimuli in both working-memory and long-term memory. This influence is taken for granted in visual long-term memory tasks, in which participants tend to be shown real-world scenes, objects, or faces as stimuli^{144,152} and where conceptual knowledge has long been known to scaffold visual memories. For example, one of the most classic studies in constructive memory had people repeatedly draw an ambiguously face-like visual stimulus from memory and found that these drawings slowly morphed over repeated reproductions to be more like a genuine face²⁶. These results demonstrate that people scaffold their memory for visual features using their knowledge of faces. Similarly, classic work has found that visual reproductions of ambiguous images are pulled in the direction of a label that was associated with them²⁷.

By contrast, the role of knowledge in visual working memory tasks has largely been underappreciated. In the following sections, we begin by describing how considering the role of meaning in shorter-term visual memory tasks has transformed views on how visual memory representations are maintained in everyday memory tasks. We then disentangle the joint contribution of stimulus complexity and meaning in shaping visual memory representations. Next, we integrate these points by discussing cognitive and neural evidence for the hierarchical nature of visual memory and discuss how such hierarchical structures as well as priors, scaffold memories for visual information in the real-world.

[H2] Meaning in working memory

In traditional visual working memory tasks, people tend to be shown meaningless stimuli, such as circles, defined by simple features, such as color^{49,53,65}. These simple, meaningless stimuli are often assumed to best assess the core capacity of working memory because they have no semantic associations and therefore require ongoing active maintenance to remember them^{45,153}. Sensory recruitment models based on such tasks reinforce the idea that people store visual working memories within the sensory

regions that process such stimuli (mainly low-level visual areas³⁷), which can lead to inferences because they are stored solely in an inherently perceptual format.

Using simple, single feature stimuli likely does increase the need for attentional maintenance to protect stimuli from noise, as such stimuli provide very poor long-term memory traces¹⁵⁴. However, reducing the number of features that can be stored about an object, for example by using meaningless stimuli such as colored shapes, also reduces how distinctive each object is in memory. This lack of distinctiveness leads to inter-item interference and confusability at retrieval which inevitably reduces performance in working memory tasks, leading to an underestimation of both visual working memory performance and the role of knowledge in scaffolding such performance. While such simple stimuli help to ensure the use of attentional maintenance, their use does not necessarily uncover the true architecture or computations that these visual working memory tasks rely on^{155–158}.

Consider memory for Mooney faces, two tone images that are sometimes perceived as a face and other times appear as a meaningless black and white blob (Fig. 4). In long-term memory, such stimuli have been used to show that recognition memory for the exact same item can be improved when they are perceived as meaningful (as faces) vs. not (as shape blobs)^{106,159}. Later work built on this idea in the study of working memory, demonstrating that perceiving a Mooney face as a face rather than a meaningless blob improves working memory performance and leads to an increase in the contralateral delay activity, a neural index of active maintenance engagement¹⁶⁰. This result indicates that there is more active maintenance activity when a stimulus is perceived as meaningful (such as a face or real-world object) than when it is not. There are further examples in which visually identical stimuli are better remembered in working memory tasks when they can be processed in a way that connects them to higher-level features or previous knowledge^{51,160–162}. These benefits extend to visual working memory performance whenever prior knowledge and expertise^{163,164} can be used to scaffold memory, such as for real world objects¹⁶⁵, famous faces¹⁶⁶, and functional relationships between objects¹⁶⁷. Overall, this points to the fact that visual working memory, like visual long-term memory, is not based solely on perceptual features. Instead, more meaningful stimuli are better remembered than less meaningful stimuli that are perceptually identical.

[H2] Complexity and meaning

In contrast to the predictions of rudimentary information theoretic models, people are far better at representing complex, meaningful stimuli in comparison to their performance for simple meaningless stimuli, an idea that dates to very early work on 'chunking'¹¹⁴ (Fig. 4a). Although this early work subscribed to very discrete views of chunking and capacity limits that are inconsistent with modern noise-based theories, its core message remains critical and underappreciated as an explanatory factor in visual working memory studies.

One explanation for the memory benefits of perceiving a meaningful and complex, multidimensional stimulus¹⁶⁰ is that doing so enhances the ability to extract features from that image. One specific hypothesis is that people are not completely flexible in their

encoding, relying heavily on pre-learned features (for instance, using phonological features to store both binary and decimal digits, rather than adaptively switching to a code that takes advantage of the additional compressibility of binary digits¹²⁷). Thus, when remembering an unfamiliar, meaningless image one can only encode it in terms of its low-level shape, spatial frequency, and other low-level and mid-level features (Fig. 4b). However, when perceiving the same stimulus as meaningful, one unlocks higher level visual features, such as face-specific features like eye position, age, and nose angle. This additional complexity likely improves performance because they make each stimulus more distinctive from the other stimuli (Fig. 4c). This role of featural distinctiveness is commonly modeled in the long-term recognition memory literature¹⁶⁸ but rarely in working memory tasks. Furthermore, recognizing features at a higher-level uncovers relations between features and makes it possible to efficiently encode and chunk multiple lower-level features jointly instead of separately, which also improves memory^{127,169}.

Importantly, however, it is not the case that adding arbitrary features to an object improves memory performance when those features are not part of one's library of pre-learned features. For instance, although scrambled or inverted stimuli can be more complex and perceptually rich than simpler single feature stimuli, they do not offer the opportunity to make use of existing higher-level features in the same way as realistic, meaningful stimuli presented in their pre-learned configuration. Indeed, memorizing more versus less visually complex meaningless stimuli hurts rather than helps visual memory performance^{106,170,171}. Thus, the benefits for extracting additional features from a visual object are only observed with stimuli with existing ways to encode high-level visual and semantic information, not scrambled yet perceptually complex stimuli^{162,167}. This concept is sometimes formalized in working memory as a kind of 'model mismatch'¹²⁸. Stimuli that match an internal model of the world are easier to remember than those that do not.

As we elaborate below, the beneficial role of prior knowledge bridges many findings on how complexity and meaning can jointly improve visual working memory performance. These include benefits when multiple features are chunked into a single object as compared to separate ones⁸⁷, better performance at remembering recognized Mooney faces than unrecognized Mooney faces¹⁶⁰ and other aspects of model mismatch¹²⁸, as well as improvements in working memory performance after learning about which features tend to go together¹²⁷. Conceptualizing chunking in a more continuous way, consistent with noisy representations, remains an important goal for future work on visual memory¹⁷². Conceptions of chunking construed in terms of more sophisticated formal models of information compression or rate-distortion theory^{120,122}, which have the potential to take into account differential effects of stimulus complexity, will likely be critical to future progress in understanding these phenomena. Through this more contemporary lens, visual memory chunking can be seen as a form of 'lossy' compression of continuous multidimensional variables, that is, compression of memory representations that results in some information loss by forming more efficient, more abstract representations that are linked to prior knowledge structures. More generally, although item-based theories have often been used to explain chunking in memory¹⁷³⁻¹⁷⁵, continuous theories of memory strength are equally compatible with the view that

memory representations can be translated into more efficient formats, and that the common link between discrete-slot models and compressibility is historical rather than logical.

[H2] Elaborative encoding and hierarchical memories

The idea that recognizing and connecting an object to high-level visual and semantic features improves memory is closely related to the popular idea of 'levels of processing' which is more commonly studied for verbal materials^{176,177}. According to this view, processing of stimuli at a 'deeper' level leads to more elaborated memory traces¹⁷⁸. Such memory traces are then more distinctive^{179,180} and easier to retrieve because they are more connected to multiple distinct knowledge structures that can serve as retrieval cues^{176,181}. Elaborative encoding in visual memory can occur when encoding semantic features, but also mid-level (such as shape) and high-level (such as the distance between two eyes on a face) visual features. Thus, the hierarchical structure of visual objects and scenes, in terms of low-level to high-level visual features and semantic information can enable improved memory performance.

Currently, how the hierarchy of visual features scaffold memory performance and improve the ability to remember realistic stimuli has not been explained by most models of visual working memory performance. Instead, even models that focus on the prior knowledge tend to focus on memory capacity rather than memory representations (e.g., rate distortion or information theoretic models, as discussed above). Thus, these models tend to focus on how performance varies as a function of the number of simple visual features shown. Important recent work has attempted to model representations of realistic stimuli using generative deep learning neural nets to capture how stimuli are represented at multiple levels of the visual hierarchy¹⁸². This model attempts to mimic the hierarchical structure of the ventral visual pathway, capturing the encoding of low-level visual features, such as orientation, and mid-level features, such as shape. Critically, by using recurrent feedback mechanisms, the model can be used to reconstruct important memory phenomena, such as more resource-efficient representation of familiar items, and efficient recognition of novel objects that share high-level features with familiar items. These findings and related work^{183,184} highlight that interactions across the visual hierarchy are bidirectional, meaning that signals from early areas of the visual hierarchy influence representations at higher levels, and vice versa and studying these interactions is an important step toward expanding models of visual working memory to account for a feature hierarchy.

In contrast to the visual working memory literature, the role of elaborative encoding and the availability of high-level or semantic features has more frequently been studied in the domain of visual long-term memory. For example, the role of categories in memory, such as the finding that memories might be biased towards commonly encountered instances of a category, suggests that semantic features, in addition to perceptual features, play a substantial role in memory encoding¹⁸⁵⁻¹⁸⁷. Recent work has also shown that normal, unmodified faces, activate higher-dimensional neural representations (including the involvement of high-level face processing regions) and results in better memory

compared to modified faces¹⁸⁸. The literature on memorability also taps into these concepts. Memorability is based on the empirical finding that some stimuli are more often remembered successfully, compared to other—sometimes very similar—stimuli. Although this concept is sometimes labeled intrinsic memorability, classic memory theories^{176,179} would suggest that some stimuli are more memorable because they connect more directly to higher-level visual and semantic features, enabling more elaborative encoding and are thus more distinctive (have lowered summed similarity to the current memory set¹⁸⁹), rather than an intrinsic property of the stimulus itself. Teasing apart the relative contributions of elaborative encoding and distinctiveness within the set of items in memory to reliability in responses across individuals remains an active area of work in the memorability literature^{4,190}.

In addition to the fact that features of a visual stimulus can be represented at multiple levels in the visual hierarchy (e.g., more perceptual, or more conceptual), there is even a stronger sense in which visual memory can be considered hierarchical. Visual memory might be fundamentally structured, with distinct representations at different levels of abstraction. One reliable phenomenon that supports this view is that people's memories are biased towards, but are not replaced with, their priors^{56,191,192}. For example, people's memories for the size of specific instances of natural objects (such as fruits and vegetables) are biased towards the average size of the object categories (size of all apples) and superordinate categories (size of all fruits¹⁹³; Box 3). This effect of priors on memory even holds in tasks in which priors are not informative for the stimulus selection process. For instance, people use knowledge of color categories even in a task where items' colors are randomly chosen¹⁹⁴—such that even for extremely simple stimuli like colored dots, visual memories naturally drift toward prior knowledge of color categories¹⁹⁵. More precisely, evidence suggests that people see some hues of a color as being more representative of the color category, and their memory will tend to drift towards these representative feature values.

In addition to memory depending on knowledge acquired before encoding, memories also seem to be encoded hierarchically within an episode. For example, when faced with sequences of items, in visual working- or long-term memory tasks, people make use of similarity among multiple items to store abstractions and those abstractions influence subsequent memory for individual items^{196–198}. For instance, scenes help provide a structure to the objects within them as shown by evidence that if people are presented an array of objects that are embedded within a scene, people encode both the ensemble structure of the entire scene in addition to individual items^{101,199,200}. Similarly, when trying to remember real-world scenes, people encode the gist plus detailed information about some specific objects^{201,202}. Moreover, they use the gist to guide their choice of which specific objects to remember^{203,204} and are influenced by the gist when later trying to recall the details of the scene, tending to remember objects that are consistent with the scene gist even if they were not present^{205–207}. Even in randomly generated displays of simple colors, people are typically able to extract some spatial or feature-based structure from the display, which they use to encode the items with respect to each other rather than totally independently²⁰⁰. Finally, in some situations, people seem to encode relatively

separate texture or ensemble information in addition to individual items and use it to alter their responses²⁰⁸.

In long-term memory, a similar view prevails that episodic memories, are hierarchical. Episodic long-term memories integrate across several levels²⁰⁹ or storing both gist and detailed item information and trading them off³³. These episodic memory processes flexibly integrate disparate components (including details and related information) of an autobiographical event into a coherent representation of the past or future^{209,210}.

To summarize, visual memory in the real-world operates on meaningful and familiar items, such as real-world objects and scenes. Memory for real-world objects is jointly but differentially influenced by stimulus meaning and complexity, where meaning can bolster memory by distributing information across levels of the visual hierarchy, with memory for perceptual features influenced by and supported by memory for more abstract conceptual features, and added complexity can further scaffold memory by providing additional relational information that can be used to compress meaningful information, or serve as an additional retrieval cue. Likewise, such information is heavily influenced by prior knowledge, which can adaptively bias memory representations to optimize memory performance. Collectively, the study of memory for meaningful as opposed to meaningless, degraded visual information can help ground theories and models of visual memory in the real-world.

[H1] Summary and future directions

The visual system is usually conceived of in terms of a hierarchy of population codes, with representations that become more complex as we move from the primary visual cortex up to more anterior visual regions. In this Review, we have suggested that a similar set of ideas can provide a useful mental model for understanding visual memory representations, with representations for objects or scenes conceived of as population-based, noisy (variable in strength), and stored in terms of hierarchical and distributed memory representations.

This conception of visual memory raises major challenges for measuring memory performance. In particular, modeling the decision process people use to integrate across an entire hierarchy of noisy representations will be critical to successfully measuring and understanding the memory representations themselves. In this context, a fruitful direction for future research is to examine how model-based approximations of representations – such as those instantiated via convolutional neural nets^{211,212}, or probabilistic models of knowledge structures²¹³ – can yield insights into how continuous representations arise and are integrated across the visual hierarchy.

Greater synergy between the kinds of models and mechanisms proposed in visual working and visual long-term memory tasks will be critical to progress across both domains, but particularly in the domain of working memory. Existing working memory models are largely focused on manipulations of memory load for brief, simultaneous

presentations of low level visual features, and more cohesive models that use many of the same principles proposed in models of long-term memory are needed to examine and explain how encoding, delay, frequency of item presentation, context, testing conditions²¹⁴ and item influence visual working memory processes. Such modeling approaches can also help illuminate the intersection between visual working and long-term memory and the role of limits on attentional protection of items from noise accumulation in real world tasks. To this point, it is also important to continue work that precisely defines – at the cognitive and neural level – how differences in ‘activation’ or attentional engagement can distinguish visual working-memory and long-term memory.

Finally, a focus on precise quantification of memory representations in visual memory (as in continuous reproduction) and on research domains in which representations are well understood, such as in perception and other domains where neuroscience informs the precise representation structure of items, will enable the development of detailed computational models that can help elucidate performance. However, it is also critical to consider that impoverished stimuli like single visual features do not capture the full breadth of scenarios where memory is used, such as in the real-world, and so models—particularly of visual working memory performance—must carefully consider the role of elaborative encoding of more realistic stimuli.

Figure captions

Figure 1. Visual memory at different time scales. Features of a visual object (1) are extracted and processed, eliciting a noisy population of activity over feature values, such as edges, color, and shape (2). Here we depict a simple case in which the color of the object is relatively arbitrary. Perception of the relevant color is noisy, and multiple colors near the shown green color all feel somewhat familiar after encoding the item (3). Focusing on actively holding an item in mind using attention can slow, but not stop, the accumulation of noise in that item's representation (4). Sampling the same memory representation at different time points reveals that the accumulated noise corrupts and alters the qualities of the originally encoded information, resulting in a noisy and imprecise sense of the color of the original object (5-6).

Figure 2. Memory strength across time. Distributions of memory errors (gray) obtained from continuous reproduction visual-working and long-term memory tasks³⁴ fit with a model (blue line³⁵) that requires only a single parameter which captures memory strength. When the same stimulus features are used and are probed in the same way, representations across the two timescales can be strong and precise or weak and noisy and degrade in an identical fashion as memory strength weakens.

Figure 3. Categorical and continuous reproduction of memory. A. Early methods for measuring memory performance used binary tasks, such as old/new for visual working memory and change detection tasks for long-term memory. Data from binary tasks have most commonly been interpreted as providing support for discrete memory representations^{45,50}. B. Continuous reproduction tasks are memory versions of the psychophysics method of adjustment. These tasks provide a fine-grained measure of the magnitude and direction of memory errors, which reveal that representations are noisy and are rarely completely lost in discrete fashion. C. Major insights from tasks in B are that precision of memory representations declines monotonically with memory load, declines with increasing delay, and that memory representations can be biased (for example, repulsed away from one another) depending on the relationship between currently active items.

Figure 4 A. As individual items become more complex, a computer or other agent that encodes the stimulus relatively optimally and flexibly can store many fewer of the items (because they have a fixed capacity in bits). However, human performance remains relatively constant as complexity increases so long as the information is meaningful, or capable of being scaffolded by previously acquired information. People have flexibility in their encoding but rely heavily on pre-learned features (for instance, using phonological features to store both binary and decimal digits, rather than adaptively switching to a code that takes advantage of the additional compressibility of binary digits).] B. When

inverted Mooney faces (meaningless stimuli) are maintained in memory, human performance is relatively poor compared to when the same faces are upright (and recognized as faces). One hypothesis is that existing high-level features that apply to perceived faces but not shapes add meaningful dimensions to the stimuli, therefore making them more separable and robust to interference and noise (such as age of the face, shown as the Z axis).

BOX 1: Memory as a point or a population

Conceiving of memory representations as noisy helps bridge the gap between cognitive and neural models of visual working and long-term memory. Much neural evidence suggests that cognitive representations are instantiated in distributed patterns of activity across neural populations and that activity across and within these populations is corrupted by noise^{69,215,216}. By contrast, many cognitive models of memory — especially long-term memory — tend to treat memory for an object as being effectively a single unitized familiarity signal, perhaps aided by other context-based sources of memory when determining whether an item has been seen before^{74,217}. In the past few years, other work has suggested that in working memory and long-term memory, these two views can be naturally reconciled by taking the signal-detection-based approach common in understanding familiarity and turning it into a population of familiarity signals modulated by psychophysical similarity³⁵. On the neural level, psychophysical similarity can arise from the tuning functions in feature-selective populations of neurons⁶⁵, and memory retrieval involves decoding these neural patterns of activations of stored features. [Approaches like these therefore provide a common framework for thinking about noise across both cognitive decision models and neural models, bridging different levels of processing²¹⁸.

Conceptions of memory based on a population of signals for remembering a single color are quite different from views in which what is stored is just a single point representation ('I think it was red'). An actively debated question that is deeply connected to the population coding view of visual memory²¹⁹, is whether people truly represent perceptual and visual memories as point estimates or as a probability distribution over features values²²⁰. This question is difficult to address because versions of point estimates models can mimic models that postulate probabilistic cognitive representations²²¹. Thus, addressing this question might depend critically on understanding the linking function between neural activity and cognitive representations. For instance, while population coding neural models are strongly compatible with the idea that people represent visual memories probabilistically (in early visual cortex), it is possible that this activity is inaccessible to higher-level processing and thus cannot be used to make decisions in memory tasks⁹². Nevertheless, people's subjective uncertainty in their memories tracks qualitative and quantitative properties of their memory errors^{78,79}, indicating that at a minimum people do have access to and readout information regarding how noisy their memory representations are, which suggests the possibility that visual memories are truly populations, not points.

BOX 2: Continuity in memory demands richer models

The finding that performance increases relatively continuously across different visual working and long-term memory manipulations and tasks raises the question of whether a single capacity limit is a useful construct, particularly in the domain of working memory where it is often used to motivate models and debates^{46,69}.

Models of visual working memory are typically developed to explain how performance varies as a function of how many items are simultaneously, briefly presented^{64,68,222}. As such, these models explain just a single slice through the space of possible levels of memory performance (top graph in figure). For instance, modulating encoding time, set size, delay, and how distinctive test items are to those in memory influences performance, but these variables are not often manipulated in systematic ways because ‘resource’ models generally do not provide an integrative framework for understanding how these levers influence the availability of a resource or noise accrual.

Given these limitations, a fruitful direction for visual working memory research is one taken in the study of long-term recognition and recall memory. In this domain, researchers often build integrative architectural models that describe processes that limit performance as a function of the particular encoding and retrieval manipulations for a given set of stimuli^{105,168,223}, as opposed to simply using more vague constructs such as capacity or limited resources. For example, considerations of how the distinctiveness of stimuli relate to the memory cues used will likely be fruitful in working memory research because the similarity of alternative items presented at test plays an important role in shaping memory performance. Similarly, some stimuli spaces — such as real-world objects — have a higher upper bound on distinctiveness (right graph in figure) than those typically modeled (such as colors)¹⁰⁶. Taking into account the full set of manipulations that impact performance, as opposed to heavily prioritizing how many items are simultaneously present, will lead to a deeper understanding of how visual memory representations are affected by task demands and provide a common framework for thinking about visual memory at different time-scales.

BOX 3: Random guesses in memory

It is often proposed that memory performance includes many responses that are completely independent of the information held in working memory, with no relationship between the response and the encoded nor the remembered information. In such a view, quantifying such 'pure guesses' is critical to understanding working memory^{67,68} and long-term memory²²⁴⁻²²⁶. However, if we accept the view that memory is both hierarchical and continuous in strength, the utility of a construct such a pure guessing is unclear.

Consider a visual working memory task in which one sees several different apples presented on a display and must remember their sizes for a memory probe a few seconds later. The vast majority of evidence suggests that although noise will accumulate for each item independently, this noise cannot be characterized as instantaneous and infinite (which is what is necessary for an item to be a pure guess). Instead, every item will likely be represented, even if this is with a very large degree of noise^{35,50,64,82}. However, even if items are so noisy that they may functionally be completely gone from memory, continuous models do not conceive of this a distinct state of 'guessing'. Such models instead hold the core assumption that even under conditions where an item is encoded extremely poorly, people still have some information in memory about that item; that information is simply almost total noise. However, this noise still has content information, and it is this content information people use to make their response.

Even if, in contrast to this view⁶⁷, an item representation is fully lost from memory, it is still unclear whether considering a response to that item as a guess is useful (as in²²⁷). Even without any information about that specific item, one has substantial information about the presented items in a more general sense. For instance, one will know what size the items were in general on the display⁵⁶, will likely know all the items present were apples, will have strong expectations about the expected size of an apple in general (versus other fruit such as a watermelon^{193,228}), and will know the expected size of each item on a display with just 3 apples on it. All of this information was acquired during the episode and informs the response about the size of the particular item, even if the individual representation of the actual size of a single apple is indistinguishable from noise at test. Thus, rather than a focus on whether or not participants are 'guessing', researchers should focus on capturing the true contents of memory, taking into account the fact that memory is noisy, reconstructive in nature¹⁴¹ and depends on integration across multiple levels of representational abstraction^{56,199}.

1. Foer, J. *Moonwalking with Einstein: The art and science of remembering everything.* (Penguin, 2012).
2. Squire, L. R., Knowlton, B. & Musen, G. The structure and organization of memory. *Annu. Rev. Psychol.* **44**, 453–495 (1993).
3. Eichenbaum, H., Yonelinas, A. R. & Ranganath, C. The medial temporal lobe and recognition memory. *Annu. Rev. Neurosci.* **30**, 123 (2007).
4. Brady, T. F. & Bainbridge, W. A. *Visual Memory.* (Routledge, 2022).
5. Luck, S. J. & Hollingworth, A. *Visual Memory.* Oxford University Press. (2008).
6. Hubel, D. H. & Wiesel, T. N. Ferrier lecture-Functional architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond. B Biol. Sci.* **198**, 1–59 (1977).
7. Logothetis, N. K. & Sheinberg, D. L. Visual object recognition. *Annu. Rev. Neurosci.* (1996).
8. Yamins, D. L. K. *et al.* Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc. Natl. Acad. Sci.* **111**, 8619–8624 (2014).
9. Todd, J. J. & Marois, R. Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* **428**, 751–754 (2004).
10. Averbeck, B. B., Latham, P. E. & Pouget, A. Neural correlations, population coding and computation. *Nat. Rev. Neurosci.* **7**, 358–366 (2006).
11. Beschin, N., Cocchini, G., Della Sala, S. & Logie, R. H. What the Eyes Perceive, The Brain Ignores: A Case of Pure Unilateral Representational Neglect. *Cortex* **33**, 3–26 (1997).
12. Guariglia, C., Padovani, A., Pantano, P. & Pizzamiglio, L. Unilateral neglect restricted to visual imagery. *Nature* **364**, 235–237 (1993).
13. Hebb, D. O. *The Organization of Behavior: A Psychological Theory.* (Wiley New York, 1949).

14. Fuster, J. M. & Alexander, G. E. Neuron activity related to short-term memory. *Science* **173**, 652–654 (1971).
15. Goldman-Rakic, P. S. Working Memory and the Mind. *Sci. Am.* **267**, 110–117 (1992).
16. Berggren, N. & Eimer, M. Does contralateral delay activity reflect working memory storage or the current focus of spatial attention within visual working memory? *J. Cogn. Neurosci.* **28**, 2003–2020 (2016).
17. Ikkai, A., McCollough, A. W. & Vogel, E. K. Contralateral Delay Activity Provides a Neural Measure of the Number of Representations in Visual Working Memory. *J. Neurophysiol.* **103**, 1963–1968 (2010).
18. Kang, M.-S. & Woodman, G. F. The neurophysiological index of visual working memory maintenance is not due to load dependent eye movements. *Neuropsychologia* **56**, 63–72 (2014).
19. Klaver, P., Talsma, D., Wijers, A. A., Heinze, H.-J. & Mulder, G. An event-related brain potential correlate of visual short-term memory. *NeuroReport* **10**, 2001–2005 (1999).
20. Pomper, U., Ditye, T. & Ansorge, U. Contralateral delay activity during temporal order memory. *Neuropsychologia* **129**, 104–116 (2019).
21. Carlisle, N. B., Arita, J. T., Pardo, D. & Woodman, G. F. Attentional templates in visual working memory. *J. Neurosci.* **31**, 9315–9322 (2011).
22. Barbosa, J., Lozano-Soldevilla, D. & Compte, A. Pinging the brain with visual impulses reveals electrically active, not activity-silent, working memories. *PLOS Biol.* **19**, e3001436 (2021).
23. Harrison, E. M. & Gorman, M. R. Changing the Waveform of Circadian Rhythms: Considerations for Shift-Work. *Front. Neurol.* **3**, (2012).

24. Serences, J. T., Ester, E. F., Vogel, E. K. & Awh, E. Stimulus-Specific Delay Activity in Human Primary Visual Cortex. *Psychol. Sci.* **20**, 207–214 (2009).

25. Morey, C. C. The case against specialized visual-spatial short-term memory. *Psychol. Bull.* **144**, 849–883 (2018).

26. Bartlett, F. C. & Bartlett, F. C. *Remembering: A study in experimental and social psychology*. (Cambridge university press, 1995).

27. Carmichael, L., Hogan, H. P. & Walter, A. A. An experimental study of the effect of language on the reproduction of visually perceived form. *J. Exp. Psychol.* **15**, 73 (1932).

28. Logie, R. H. *Visuo-spatial Working Memory*. (Psychology Press, 2014). doi:10.4324/9781315804743.

29. Logie, R. H., Belletier, C. & Doherty, J. M. Integrating theories of working memory. in *Working memory: State of the science* 389–429 (Oxford University Press, 2021).

30. Atkinson, R. C. & Shiffrin, R. M. Human Memory: A Proposed System and its Control Processes. in *Psychology of Learning and Motivation* vol. 2 89–195 (Elsevier, 1968).

31. Baddeley, A. D. The influence of acoustic and semantic similarity on long-term memory for word sequences. *Q. J. Exp. Psychol.* **18**, 302–309 (1966).

32. Cowan, N. What are the differences between long-term, short-term, and working memory? *Prog. Brain Res.* **169**, 323–338 (2008).

33. Brainerd, C. J. & Reyna, V. F. Fuzzy-trace theory and false memory. *Curr. Dir. Psychol. Sci.* **11**, 164–169 (2002).

34. Miner, A. E., Schurgin, M. W. & Brady, T. F. Is working memory inherently more “precise” than long-term memory? Extremely high fidelity visual long-term memories for frequently encountered objects. *J. Exp. Psychol. Hum. Percept. Perform.* **46**, 813 (2020).

35. Schurgin, M. W., Wixted, J. T. & Brady, T. F. Psychophysical scaling reveals a unified theory of visual memory strength. *Nat. Hum. Behav.* **4**, 1156–1172 (2020).
36. Serences, J. T. & Yantis, S. Selective visual attention and perceptual coherence. *Trends Cogn. Sci.* **10**, 38–45 (2006).
37. Sprague, T. C., Ester, E. F. & Serences, J. T. Restoring latent visual working memory representations in human cortex. *Neuron* **91**, 694–707 (2016).
38. Stokes, M. G. ‘Activity-silent’working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* **19**, 394–405 (2015).
39. Cunningham, C. A. & Wolfe, J. M. The role of object categories in hybrid visual and memory search. *J. Exp. Psychol. Gen.* **143**, 1585 (2014).
40. Wolfe, J. M., Boettcher, S. E., Josephs, E. L., Cunningham, C. A. & Drew, T. You look familiar, but I don’t care: Lure rejection in hybrid visual and memory search is not based on familiarity. *J. Exp. Psychol. Hum. Percept. Perform.* **41**, 1576 (2015).
41. Baddeley, A. Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* **4**, 829–839 (2003).
42. Chun, M. M., Golomb, J. D. & Turk-Browne, N. B. A Taxonomy of External and Internal Attention. *Annu. Rev. Psychol.* **62**, 73–101 (2011).
43. Postle, B. R., Druzgal, T. J. & D’Esposito, M. Seeking the neural substrates of visual working memory storage. *Cortex J. Devoted Study Nerv. Syst. Behav.* **39**, 927–946 (2003).
44. Oberauer, K. Access to information in working memory: exploring the focus of attention. *J. Exp. Psychol. Learn. Mem. Cogn.* **28**, 411 (2002).
45. Cowan, N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav. Brain Sci.* **24**, 87–114 (2001).
46. Luck, S. J. & Vogel, E. K. Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn. Sci.* **17**, 391–400 (2013).

47. Markov, Y. A., Utochkin, I. S. & Brady, T. F. Real-world objects are not stored in holistic representations in visual working memory. *J. Vis.* **21**, 18–18 (2021).

48. Rademaker, R. L., Park, Y. E., Sack, A. T. & Tong, F. Evidence of gradual loss of precision for simple features and complex objects in visual working memory. *J. Exp. Psychol. Hum. Percept. Perform.* **44**, 925–940 (2018).

49. Wilken, P. & Ma, W. J. A detection theory account of change detection. *J. Vis.* **4**, 11 (2004).

50. Williams, J. R., Robinson, M. M., Schurgin, M. W., Wixted, J. T. & Brady, T. F. You cannot “count” how many items people remember in visual working memory: The importance of signal detection–based measures for understanding change detection performance. *J. Exp. Psychol. Hum. Percept. Perform.* **48**, 1390–1409 (2022).

51. Alvarez, G. A. & Cavanagh, P. The Capacity of Visual Short-Term Memory is Set Both by Visual Information Load and by Number of Objects. *Psychol. Sci.* **15**, 106–111 (2004).

52. Awh, E., Barton, B. & Vogel, E. K. Visual Working Memory Represents a Fixed Number of Items Regardless of Complexity. *18*, 7.

53. Luck, S. J. & Vogel, E. K. The capacity of visual working memory for features and conjunctions. *Nature* **390**, 279–281 (1997).

54. Brady, T. F., Konkle, T., Alvarez, G. A. & Oliva, A. Visual long-term memory has a massive storage capacity for object details. *Proc. Natl. Acad. Sci.* **105**, 14325–14329 (2008).

55. Bae, G.-Y. & Luck, S. J. Dissociable Decoding of Spatial Attention and Working Memory from EEG Oscillations and Sustained Potentials. *J. Neurosci.* **38**, 409–422 (2018).

56. Brady, T. F. & Alvarez, G. A. Hierarchical Encoding in Visual Working Memory: Ensemble Statistics Bias Memory for Individual Items. *Psychol. Sci.* **22**, 384–392 (2011).

57. Chunharas, C., Rademaker, R. L., Brady, T. F. & Serences, J. T. An adaptive perspective on visual working memory distortions. *J. Exp. Psychol. Gen.* (2022).

58. Golomb, J. D. Divided spatial attention and feature-mixing errors. *Atten. Percept. Psychophys.* **77**, 2562–2569 (2015).

59. Lively, Z., Robinson, M. M. & Benjamin, A. S. Memory Fidelity Reveals Qualitative Changes in Interactions Between Items in Visual Working Memory. *Psychol. Sci.* **32**, 1426–1441 (2021).

60. Scotti, P. S., Hong, Y., Golomb, J. D. & Leber, A. B. Statistical learning as a reference point for memory distortions: Swap and shift errors. *Atten. Percept. Psychophys.* **83**, 1652–1672 (2021).

61. Pertzov, Y., Bays, P. M., Joseph, S. & Husain, M. Rapid forgetting prevented by retrospective attention cues. *J. Exp. Psychol. Hum. Percept. Perform.* **39**, 1224 (2013).

62. Makovski, T. & Jiang, Y. V. Distributing versus focusing attention in visual short-term memory. *Psychon. Bull. Rev.* **14**, 1072–1078 (2007).

63. Marini, F., Scott, J., Aron, A. R. & Ester, E. F. Task-irrelevant distractors in the delay period interfere selectively with visual short-term memory for spatial locations. *Atten. Percept. Psychophys.* **79**, 1384–1392 (2017).

64. van den Berg, R., Shin, H., Chou, W.-C., George, R. & Ma, W. J. Variability in encoding precision accounts for visual short-term memory limitations. *Proc. Natl. Acad. Sci.* **109**, 8780–8785 (2012).

65. Bays, P. M. Noise in Neural Populations Accounts for Errors in Working Memory. *J. Neurosci.* **34**, 3632–3645 (2014).

66. van den Berg, R., Awh, E. & Ma, W. J. Factorial comparison of working memory models. *Psychol. Rev.* **121**, 124–149 (2014).

67. Adam, K. C. S., Vogel, E. K. & Awh, E. Clear evidence for item limits in visual working memory. *Cognit. Psychol.* **97**, 79–97 (2017).

68. Zhang, W. & Luck, S. J. Discrete fixed-resolution representations in visual working memory. *Nature* **453**, 233–235 (2008).

69. Schneegans, S., Taylor, R. & Bays, P. M. Stochastic sampling provides a unifying account of visual working memory limits. *Proc. Natl. Acad. Sci.* **117**, 20959–20968 (2020).

70. Pratte, M. S. Set size effects on working memory precision are not due to an averaging of slots. *Atten. Percept. Psychophys.* **82**, 2937–2949 (2020).

71. Bays, P., Schneegans, S., Ma, W. J. & Brady, T. F. *Representation and computation in working memory*. <https://osf.io/kubr9> (2022) doi:10.31234/osf.io/kubr9.

72. Swets, J. A. Form of empirical ROCs in discrimination and diagnostic tasks: implications for theory and measurement of performance. *Psychol. Bull.* **99**, 181 (1986).

73. Wixted, J. T. Dual-process theory and signal-detection theory of recognition memory. *Psychol. Rev.* **114**, 152–176 (2007).

74. Wixted, J. T. & Mickes, L. A continuous dual-process model of remember/know judgments. *Psychol. Rev.* **117**, 1025 (2010).

75. Taylor, R. & Bays, P. M. Theory of neural coding predicts an upper bound on estimates of memory variability. *Psychol. Rev.* **127**, 700 (2020).

76. Delay, C. G. & Wixted, J. T. Discrete-state versus continuous models of the confidence-accuracy relationship in recognition memory. *Psychon. Bull. Rev.* **28**, 556–564 (2021).

77. Brady, T. F., Robinson, M. M., Williams, J. R. & Wixted, J. T. Measuring memory is harder than you think: How to avoid problematic measurement practices in memory research. *Psychon. Bull. Rev.* **30**, 421–449 (2023).

78. Honig, M., Ma, W. J. & Fougner, D. Humans incorporate trial-to-trial working memory uncertainty into rewarded decisions. *Proc. Natl. Acad. Sci.* **117**, 8391–8397 (2020).

79. Jabar, S. B. *et al.* Using a betting game to reveal the rich nature of visual working memories. *bioRxiv* (2020).

80. Yoo, A. H., Acerbi, L. & Ma, W. J. Uncertainty is maintained and used in working memory. *J. Vis.* **21**, 13–13 (2021).

81. Treisman, A. Features and Objects in Visual Processing. *Sci. Am.* 13 (1986).

82. Schneegans, S. & Bays, P. M. Neural architecture for feature binding in visual working memory. *J. Neurosci.* **37**, 3913–3925 (2017).

83. Bays, P. M., Wu, E. Y. & Husain, M. Storage and binding of object features in visual working memory. *Neuropsychologia* **49**, 1622–1631 (2011).

84. Fougner, D. & Alvarez, G. A. Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *J. Vis.* **11**, 3–3 (2011).

85. Shin, H. & Ma, W. J. Visual short-term memory for oriented, colored objects. *J. Vis.* **17**, 12–12 (2017).

86. Utochkin, I. S. & Brady, T. F. Independent storage of different features of real-world objects in long-term memory. *J. Exp. Psychol. Gen.* **149**, 530 (2020).

87. Fougner, D., Cormiea, S. M. & Alvarez, G. A. Object-based benefits without object-based representations. *J. Exp. Psychol. Gen.* **142**, 621–626 (2013).

88. Wang, B., Cao, X., Theeuwes, J., Olivers, C. N. & Wang, Z. Location-based effects underlie feature conjunction benefits in visual working memory. *J. Vis.* **16**, 12–12 (2016).

89. Egly, R., Driver, J. & Rafal, R. D. Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *J. Exp. Psychol. Gen.* **123**, 161 (1994).

90. Francis, G. & Thunell, E. Excess success in articles on object-based attention. *Atten. Percept. Psychophys.* **84**, 700–714 (2022).

91. Chater, N., Tenenbaum, J. B. & Yuille, A. Probabilistic models of cognition: Conceptual foundations. *Trends Cogn. Sci.* **10**, 287–291 (2006).

92. Ma, W. J. Organizing probabilistic models of perception. *Trends Cogn. Sci.* **16**, 511–518 (2012).

93. Kellen, D., Winiger, S., Dunn, J. C. & Singmann, H. Testing the foundations of signal detection theory in recognition memory. *Psychol. Rev.* **128**, 1022 (2021).

94. Rotello, C. M., Heit, E. & Dubé, C. When more data steer us wrong: replications with the wrong dependent measure perpetuate erroneous conclusions. *Psychon. Bull. Rev.* **22**, 944–954 (2015).

95. Starns, J. J. *et al.* Assessing theoretical conclusions with blinded inference to investigate a potential inference crisis. *Adv. Methods Pract. Psychol. Sci.* **2**, 335–349 (2019).

96. Wixted, J. T. & Mickes, L. Theoretical vs. empirical discriminability: the application of ROC methods to eyewitness identification. *Cogn. Res. Princ. Implic.* **3**, 9 (2018).

97. Wixted, J. T. & Wells, G. L. The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychol. Sci. Public Interest* **18**, 10–65 (2017).

98. Adler, W. T. & Ma, W. J. Comparing Bayesian and non-Bayesian accounts of human confidence reports. *PLoS Comput. Biol.* **14**, e1006572 (2018).

99. Ma, W. J. Bayesian decision models: A primer. *Neuron* **104**, 164–175 (2019).

100. Pouget, A., Beck, J. M., Ma, W. J. & Latham, P. E. Probabilistic brains: knowns and unknowns. *Nat. Neurosci.* **16**, 1170–1178 (2013).

101. Brady, T. F. & Tenenbaum, J. B. A probabilistic model of visual working memory: Incorporating higher order regularities into working memory capacity estimates. *Psychol. Rev.* **120**, 85–109 (2013).

102. Smith, P. L., Saber, S., Corbett, E. A. & Lilburn, S. D. Modeling continuous outcome color decisions with the circular diffusion model: Metric and categorical properties. *Psychol. Rev.* **127**, 562 (2020).

103. Ma, W. J., Husain, M. & Bays, P. M. Changing concepts of working memory. *Nat. Neurosci.* **17**, 347–356 (2014).

104. Schurgin, M. W. & Flombaum, J. I. Visual working memory is more tolerant than visual long-term memory. *J. Exp. Psychol. Hum. Percept. Perform.* **44**, 1216 (2018).

105. Howard, M. W. & Kahana, M. J. A distributed representation of temporal context. *J. Math. Psychol.* **46**, 269–299 (2002).

106. Brady, T. F. & Störmer, V. S. The role of meaning in visual working memory: Real-world objects, but not simple features, benefit from deeper processing. *J. Exp. Psychol. Learn. Mem. Cogn.* (2021).

107. Endress, A. D. & Potter, M. C. Large capacity temporary visual memory. *J. Exp. Psychol. Gen.* **143**, 548 (2014).

108. Tsubomi, H., Fukuda, K., Watanabe, K. & Vogel, E. K. Neural limits to representing objects still within view. *J. Neurosci.* **33**, 8257–8263 (2013).

109. Vogel, E. K., Woodman, G. F. & Luck, S. J. The time course of consolidation in visual working memory. *J. Exp. Psychol. Hum. Percept. Perform.* **32**, 1436 (2006).

110. Ricker, T. J. & Hardman, K. O. The nature of short-term consolidation in visual working memory. *J. Exp. Psychol. Gen.* **146**, 1551 (2017).

111. Li, X., Xiong, Z., Theeuwes, J. & Wang, B. Visual memory benefits from prolonged encoding time regardless of stimulus type. *J. Exp. Psychol. Learn. Mem. Cogn.* **46**, 1998 (2020).
112. Draschkow, D., Kallmayer, M. & Nobre, A. C. When natural behavior engages working memory. *Curr. Biol.* **31**, 869–874 (2021).
113. Williams, J. & Störmer, V. S. Working Memory: How Much Is It Used in Natural Behavior? *Curr. Biol.* **31**, R205–R206 (2021).
114. Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.* **63**, 81 (1956).
115. Balaban, H. & Luria, R. The number of objects determines visual working memory capacity allocation for complex items. *NeuroImage* **119**, 54–62 (2015).
116. Barton, B., Ester, E. F. & Awh, E. Discrete resource allocation in visual working memory. *J. Exp. Psychol. Hum. Percept. Perform.* **35**, 1359–1367 (2009).
117. Cappiello, M., Xie, W. & Zhang, W. The mental muscle: Effects of concurrent effortful physical action on visual working memory. *J. Vis.* **18**, 705–705 (2018).
118. Rouder, J. N. *et al.* An assessment of fixed-capacity models of visual working memory. *Proc. Natl. Acad. Sci.* **105**, 5975–5979 (2008).
119. Pratte, M. S. & Rouder, J. N. Hierarchical single- and dual-process models of recognition memory. *J. Math. Psychol.* **55**, 36–46 (2011).
120. Navon, D. Resources—A Theoretical Soup Stone? 19.
121. Franconeri, S. L., Alvarez, G. A. & Cavanagh, P. Flexible cognitive resources: competitive content maps for attention and memory. *Trends Cogn. Sci.* **17**, 134–141 (2013).
122. Bays, P. M. Spikes not slots: noise in neural populations limits working memory. *Trends Cogn. Sci.* **19**, 431–438 (2015).

123. Carandini, M. & Heeger, D. J. Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* **13**, 51–62 (2012).

124. Smith, P. L., Lilburn, S. D., Corbett, E. A., Sewell, D. K. & Kyllingsbæk, S. The attention-weighted sample-size model of visual short-term memory: Attention capture predicts resource allocation and memory load. *Cognit. Psychol.* **89**, 71–105 (2016).

125. Green, D. M. & Swets, J. A. *Signal detection theory and psychophysics*. vol. 1 (Wiley New York, 1966).

126. Palmer, J. Attentional limits on the perception and memory of visual information. *J. Exp. Psychol. Hum. Percept. Perform.* **16**, 332 (1990).

127. Brady, T. F., Konkle, T. & Alvarez, G. A. Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *J. Exp. Psychol. Gen.* **138**, 487–502 (2009).

128. Orhan, A. E., Sims, C. R., Jacobs, R. A. & Knill, D. C. The Adaptive Nature of Visual Working Memory. *Curr. Dir. Psychol. Sci.* **23**, 164–170 (2014).

129. Bates, C. J. & Jacobs, R. A. Efficient data compression in perception and perceptual memory. *Psychol. Rev.* **127**, 891–917 (2020).

130. Tomić, I. & Bays, P. M. Internal but not external noise frees working memory resources. *PLOS Comput. Biol.* **14**, e1006488 (2018).

131. Griffiths, T. L., Lieder, F. & Goodman, N. D. Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Top. Cogn. Sci.* **7**, 217–229 (2015).

132. Kristjánsson, Á. & Draschkow, D. Keeping it real: Looking beyond capacity limits in visual cognition. *Atten. Percept. Psychophys.* **83**, 1375–1390 (2021).

133. Hayhoe, M. & Ballard, D. Eye movements in natural behavior. *Trends Cogn. Sci.* **9**, 188–194 (2005).

134. Ballard, D. H., Hayhoe, M. M. & Pelz, J. B. Memory Representations in Natural Tasks. *J. Cogn. Neurosci.* **7**, 66–80 (1995).

135. Simons, D. J. & Rensink, R. A. Change blindness: Past, present, and future. *Trends Cogn. Sci.* **9**, 16–20 (2005).

136. Simons, D. J. & Ambinder, M. S. Change blindness: Theory and consequences. *Curr. Dir. Psychol. Sci.* **14**, 44–48 (2005).

137. Dretske, F. What change blindness teaches about consciousness. *Philos. Perspect.* **21**, 215–230 (2007).

138. O’regan, J. K. & Noë, A. A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* **24**, 939–973 (2001).

139. O’regan, J. K. Solving the “real” mysteries of visual perception: the world as an outside memory. *Can. J. Psychol. Can. Psychol.* **46**, 461 (1992).

140. Rensink, R. A. Change blindness. in *Neurobiology of attention* 76–81 (Elsevier, 2005).

141. Hollingworth, A. Scene and position specificity in visual memory for objects. *J. Exp. Psychol. Learn. Mem. Cogn.* **32**, 58–69 (2006).

142. Jost, K. *et al.* Controlling conflict from interfering long-term memory representations. *J. Cogn. Neurosci.* **24**, 1173–1190 (2012).

143. Konkle, T., Brady, T. F., Alvarez, G. A. & Oliva, A. Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychol. Sci.* **21**, 1551–1556 (2010).

144. Konkle, T., Brady, T. F., Alvarez, G. A. & Oliva, A. Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *J. Exp. Psychol. Gen.* **139**, 558 (2010).

145. Robertson, E. M. New insights in human memory interference and consolidation. *Curr. Biol.* **22**, R66–R71 (2012).

146. Wais, P. E. & Gazzaley, A. Distractibility during retrieval of long-term memory: domain-general interference, neural networks and increased susceptibility in normal aging. *Front. Psychol.* **5**, 280 (2014).

147. Drascher, M. L. & Kuhl, B. A. Long-term memory interference is resolved via repulsion and precision along diagnostic memory dimensions. *Psychon. Bull. Rev.* **1**–15 (2022).

148. Fukuda, K. & Vogel, E. K. Visual short-term memory capacity predicts the “bandwidth” of visual long-term memory encoding. *Mem. Cognit.* **47**, 1481–1497 (2019).

149. Sundby, C. S., Woodman, G. F. & Fukuda, K. Electrophysiological and behavioral evidence for attentional up-regulation, but not down-regulation, when encoding pictures into long-term memory. *Mem. Cognit.* **47**, 351–364 (2019).

150. Desimone, R. & Duncan, J. Neural Mechanisms of Selective Visual Attention. **30**.

151. Beck, D. M. & Kastner, S. Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Res.* **49**, 1154–1165 (2009).

152. Rugo, K. F., Tamler, K. N., Woodman, G. F. & Maxcey, A. M. Recognition-induced forgetting of faces in visual long-term memory. *Atten. Percept. Psychophys.* **79**, 1878–1885 (2017).

153. Lin, P.-H. & Luck, S. J. Proactive interference does not meaningfully distort visual working memory capacity estimates in the canonical change detection task. *Front. Psychol.* **3**, 42 (2012).

154. Olson, I. R. & Jiang, Y. Visual short-term memory is not improved by training. *Mem. Cognit.* **32**, 1326–1332 (2004).

155. Lu, X., Dai, A., Guo, Y., Shen, M. & Gao, Z. Is the social chunking of agent actions in working memory resource-demanding? *Cognition* **229**, 105249 (2022).

156. Snow, J. C. & Culham, J. C. The treachery of images: how realism influences brain and behavior. *Trends Cogn. Sci.* **25**, 506–519 (2021).

157. Vestner, T., Over, H., Gray, K. L. & Cook, R. Objects that direct visuospatial attention produce the search advantage for facing dyads. *J. Exp. Psychol. Gen.* (2021).

158. Brady, T. F. *et al.* Scaling up visual attention and visual working memory to the real world. in *Psychology of Learning and Motivation* vol. 70 29–69 (Elsevier, 2019).

159. Wiseman, S. & Neisser, U. Perceptual organization as a determinant of visual recognition memory. *Am. J. Psychol.* 675–681 (1974).

160. Asp, I. E., Störmer, V. S. & Brady, T. F. Greater visual working memory capacity for visually matched stimuli when they are perceived as meaningful. *J. Cogn. Neurosci.* **33**, 902–918 (2021).

161. Ngiam, W. X. Q., Brissenden, J. A. & Awh, E. “Memory compression” effects in visual working memory are contingent on explicit long-term memory. *J. Exp. Psychol. Gen.* **148**, 1373–1385 (2019).

162. Starr, A., Srinivasan, M. & Bunge, S. A. Semantic knowledge influences visual working memory in adults and children. *PLoS One* **15**, e0241110 (2020).

163. Moore, C. D., Cohen, M. X. & Ranganath, C. Neural Mechanisms of Expert Skills in Visual Working Memory. *J. Neurosci.* **26**, 11187–11196 (2006).

164. Xie, W. & Zhang, W. Familiarity increases the number of remembered Pokémon in visual short-term memory. *Mem. Cognit.* **45**, 677–689 (2017).

165. Brady, T. F., Störmer, V. S. & Alvarez, G. A. Working memory is not fixed-capacity: More active storage capacity for real-world objects than for simple stimuli. *Proc. Natl. Acad. Sci.* **113**, 7459–7464 (2016).

166. Jackson, M. C., Link to external site, this link will open in a new window & Raymond, J. E. Familiarity enhances visual working memory for faces. *J. Exp. Psychol. Hum. Percept. Perform.* **34**, 556–568 (2008).

167. O'Donnell, R. E., Clement, A. & Brockmole, J. R. Semantic and functional relationships among objects increase the capacity of visual working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **44**, 1151–1158 (2018).

168. Nosofsky, R. M. Tests of an exemplar model for relating perceptual classification and recognition memory. *J. Exp. Psychol. Hum. Percept. Perform.* **17**, 3 (1991).

169. Wyble, B., Swan, G. & Callahan-Flintoft, C. Measuring visual memory in its native. *Trends Cogn. Sci.* **20**, 790–791 (2016).

170. Sahar, T., Sidi, Y. & Makovski, T. A metacognitive perspective of visual working memory with rich complex objects. *Front. Psychol.* **11**, 179 (2020).

171. Stojanoski, B., Emrich, S. M. & Cusack, R. Representation of semantic information in ventral areas during encoding is associated with improved visual short-term memory. *bioRxiv* 2019–12 (2020).

172. Allen, M. G., Destefano, I. & Brady, T. F. Chunks are not “Content-Free”: Hierarchical Representations Preserve Perceptual Detail within Chunks. in *Proceedings of the Annual Meeting of the Cognitive Science Society* vol. 43 (2021).

173. Mathy, F. & Feldman, J. What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition* **122**, 346–362 (2012).

174. Wood, J. N. Visual working memory for observed actions. *J. Exp. Psychol. Gen.* **136**, 639–652 (2007).

175. Shen, M., Gao, Z., Ding, X., Zhou, B. & Huang, X. Holding biological motion information in working memory. *J. Exp. Psychol. Hum. Percept. Perform.* **40**, 1332 (2014).

176. Craik, F. I. Levels of processing: Past, present... and future? *Memory* **10**, 305–318 (2002).

177. Craik, F. I. & Lockhart, R. S. Levels of processing: A framework for memory research. *J. Verbal Learn. Verbal Behav.* **11**, 671–684 (1972).

178. Bradshaw, G. L. & Anderson, J. R. Elaborative encoding as an explanation of levels of processing. *J. Verbal Learn. Verbal Behav.* **21**, 165–174 (1982).

179. Nairne, J. S. Remembering Over THE Short-Term. *Annu Rev Psychol* **53**, 53–81 (2002).

180. Nelson, D. L., Cermak, L. & Craik, F. Remembering pictures and words: Appearance, significance and name. *Levels Process. Hum. Mem.* 45–76 (1979).

181. Tulving, E. & Thomson, D. M. Encoding specificity and retrieval processes in episodic memory. *Psychol. Rev.* **80**, 352 (1973).

182. Hedayati, S., O'Donnell, R. E. & Wyble, B. A model of working memory for latent representations. *Nat. Hum. Behav.* **6**, 709–719 (2022).

183. van Kerkoerle, T., Self, M. W. & Roelfsema, P. R. Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. *Nat. Commun.* **8**, 13804 (2017).

184. Self, M. W., van Kerkoerle, T., Goebel, R. & Roelfsema, P. R. Benchmarking laminar fMRI: Neuronal spiking and synaptic activity during top-down and bottom-up processing in the different layers of cortex. *NeuroImage* **197**, 806–817 (2019).

185. Maxcey, A. M. & Woodman, G. F. Forgetting induced by recognition of visual images. *Vis. Cogn.* **22**, 789–808 (2014).

186. Konkle, T. & Alvarez, G. Deepnets do not need category supervision to predict visual system responses to objects. *J. Vis.* **20**, 498–498 (2020).

187. Naspi, L., Hoffman, P., Devereux, B. & Morcom, A. M. Perceptual and semantic representations at encoding contribute to true and false recognition of objects. *J. Neurosci.* **41**, 8375–8389 (2021).

188. Sheng, J. *et al.* Higher-dimensional neural representations predict better episodic memory. *Sci. Adv.* **8**, eabm3829 (2022).

189. Kahana, M. J. *Foundations of human memory*. (OUP USA, 2012).

190. Rust, N. C. & Mehrpour, V. Understanding image memorability. *Trends Cogn. Sci.* **24**, 557–568 (2020).

191. Hemmer, P. & Steyvers, M. A Bayesian Account of Reconstructive Memory. *Top. Cogn. Sci.* **1**, 189–202 (2009).

192. Huttenlocher, J., Hedges, L. V. & Vevea, J. L. Why do categories affect stimulus judgment? *J. Exp. Psychol. Gen.* **129**, 220 (2000).

193. Hemmer, P. & Steyvers, M. Integrating episodic memories and prior knowledge at multiple levels of abstraction. *Psychon. Bull. Rev.* **16**, 80–87 (2009).

194. Bae, G.-Y., Olkkonen, M., Allred, S. R. & Flombaum, J. I. Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *J. Exp. Psychol. Gen.* **144**, 744–763 (2015).

195. Destefano, I., Brady, T. & Vul, E. A Framework for Predicting Memory Errors with a Bayesian Model of Concept Generalization. (2021).

196. Brady, T. F., Schacter, D. L. & Alvarez, G. A. The adaptive nature of false memories is revealed by gist- based distortion of true memories. 35.

197. Bruning, A. L. & Lewis-Peacock, J. A. Long-term memory guides resource allocation in working memory. *Sci. Rep.* **10**, 1–10 (2020).

198. Dubé, C. Central tendency representation and exemplar matching in visual short-term memory. *Mem. Cognit.* **47**, 589–602 (2019).

199. Orhan, A. E. & Jacobs, R. A. A probabilistic clustering theory of the organization of visual short-term memory. *Psychol. Rev.* **120**, 297 (2013).

200. Brady, T. F. & Alvarez, G. A. No evidence for a fixed object limit in working memory: Spatial ensemble representations inflate estimates of working memory capacity for complex objects. *J. Exp. Psychol. Learn. Mem. Cogn.* **41**, 921–929 (2015).

201. Hollingworth, A. & Henderson, J. M. Testing a conceptual locus. *Mem. Cognit.* **31**, 930–940 (2003).

202. Oliva, A. Gist of the scene. in *Neurobiology of attention* 251–256 (Elsevier, 2005).

203. Friedman, A. Framing pictures: the role of knowledge in automatized encoding and memory for gist. *J. Exp. Psychol. Gen.* **108**, 316 (1979).

204. Hollingworth, A. & Henderson, J. M. Semantic informativeness mediates the detection of changes in natural scenes. *Vis. Cogn.* **7**, 213–235 (2000).

205. Miller, M. B. & Gazzaniga, M. S. Creating false memories for visual scenes. *Neuropsychologia* **36**, 513–520 (1998).

206. Brewer, W. F. & Treyens, J. C. Role of schemata in memory for places. *Cognit. Psychol.* **13**, 207–230 (1981).

207. Lampinen, J. M., Copeland, S. M. & Neuschatz, J. S. Recollections of things schematic: room schemas revisited. *J. Exp. Psychol. Learn. Mem. Cogn.* **27**, 1211 (2001).

208. Schurgin, M. W. & Brady, T. F. When “capacity” changes with set size: Ensemble representations support the detection of across-category changes in visual working memory. *J. Vis.* **19**, 3 (2019).

209. Moscovitch, M., Cabeza, R., Winocur, G. & Nadel, L. Episodic memory and beyond: the hippocampus and neocortex in transformation. *Annu. Rev. Psychol.* **67**, 105 (2016).

210. Moscovitch, M., Nadel, L., Winocur, G., Gilboa, A. & Rosenbaum, R. S. The cognitive neuroscience of remote episodic, semantic and spatial memory. *Curr. Opin. Neurobiol.* **16**, 179–190 (2006).

211. Jaiswal, S., Fernando, B. & Tan, C. TDAM: Top-Down Attention Module for Contextually Guided Feature Selection in CNNs. in *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXV* 259–276 (Springer, 2022).

212. Bates, C. J., Alvarez, G. & Gershman, S. J. Scaling models of visual working memory to natural images. 2023.03.17.533050 Preprint at <https://doi.org/10.1101/2023.03.17.533050> (2023).

213. Lake, B. M., Salakhutdinov, R. & Tenenbaum, J. B. Human-level concept learning through probabilistic program induction. *Science* **350**, 1332–1338 (2015).

214. Cleary, A. M. Dependent measures in memory research: From free recall to recognition. in *Handbook of research methods in human memory* 19–35 (Routledge, 2018).

215. Pouget, A., Dayan, P. & Zemel, R. Information processing with population codes. *Nat. Rev. Neurosci.* **1**, 125–132 (2000).

216. Rust, N. C. & Cohen, M. R. Priority coding in the visual system. *Nat. Rev. Neurosci.* **23**, 376–388 (2022).

217. Yonelinas, A. P. Receiver-Operating Characteristics in Recognition Memory: Evidence for a Dual-Process Model. 14.

218. Marr, D. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. (MIT Press, 2010).

219. van Bergen, R. S. & Jehee, J. F. Probabilistic representation in human visual cortex reflects uncertainty in serial decisions. *J. Neurosci.* **39**, 8164–8176 (2019).

220. Yeon, J. & Rahnev, D. The suboptimality of perceptual decision making with multiple alternatives. *Nat. Commun.* **11**, 1–12 (2020).

221. Rahnev, D., Block, N., Denison, R. N. & Jehee, J. Is perception probabilistic? Clarifying the definitions. (2021).

222. Bays, P. M., Catalao, R. F. G. & Husain, M. The precision of visual working memory is set by allocation of a shared resource. *J. Vis.* **9**, 7–7 (2009).

223. Shiffrin, R. M. & Steyvers, M. A model for recognition memory: REM—retrieving effectively from memory. *Psychon. Bull. Rev.* **4**, 145–166 (1997).

224. Harlow, I. M. & Yonelinas, A. P. Distinguishing between the success and precision of recollection. *Memory* **24**, 114–127 (2016).

225. Nilakantan, A. S., Bridge, D. J., VanHaerents, S. & Voss, J. L. Distinguishing the precision of spatial recollection from its success: Evidence from healthy aging and unilateral mesial temporal lobe resection. *Neuropsychologia* **119**, 101–106 (2018).

226. Sutterer, D. W. & Awh, E. Retrieval practice enhances the accessibility but not the quality of memory. *Psychon. Bull. Rev.* **23**, 831–841 (2016).

227. Pratte, M. S. Swap errors in spatial working memory are guesses. (2018).

228. Konkle, T. & Oliva, A. A familiar-size Stroop effect: real-world size is an automatic property of object representation. *J. Exp. Psychol. Hum. Percept. Perform.* **38**, 561 (2012).

