Perceptual attributes in memory research

Robert Sekuler¹ and Allison B. Sekuler²

¹Volen Center for Complex Systems and Department of Psychology
Brandeis University, Waltham Massachusetts
²Rotman Research Institute, Baycrest; Department of Psychology
University of Toronto; Department of Psychology, Neuroscience &
Behaviour, McMaster University
Toronto and Hamilton, Ontario

To appear in
The Oxford Handbook of Human Memory
M J Kahana and A D Wagner (editors)

Final revision submitted December 2021

Acknowledgments. ABS supported by the Sandra A. Rotman Chair in Cognitive Neuroscience, Baycrest; the Canadian Institutes of Health Research; and the Natural Sciences and Engineering Research Council. RS supported by grants from the NSF, NIH, and Brandeis University Provost’s Fund for Research. We thank Chad Dubé and Mike Kahana for their consistently thoughtful comments, and Aidan J. S. Bennett for assistance in creating figures.
Abstract. The close linkage between perception and memory allows them to influence one another in numerous ways. Improved understanding of this linkage will enhance the reach and quality of research in both fields. This chapter emphasizes short-term memory, the form of memory for which perception's influence is clearest. The chapter begins with a brief overview of key perceptual attributes that enable memory to impact perceptual judgment, offering examples of perceptual phenomena that clearly bear memory’s stamp. We then turn to four ways in which memory research benefits when researchers design memoranda with perceptual factors firmly in mind. First, we consider stimulus similarity because it impacts memory in multiple ways, including its usefulness for framing and testing theoretical insights. Second, we examine how connections between perception and memory make it difficult to evaluate memory without some confound from one or more perceptual influences. These include individual or group differences in perception that can masquerade as differences in memory. Third, we consider how verbal representations can distort or support memory for what was seen, and summarizes research in which visual and auditory memory have been probed with random or quasi-random stimuli. Finally, we briefly describe how careful stimulus choice can aid modeling and understanding of memorial processes.
Textbooks and curricula alike encourage a separation between the study of perception and memory. But, perception and memory entwine, like two strands of DNA. After all, memory plays a significant role in every perceptual act, and, in turn, perception constitutes the gateway to memory. What is remembered almost always has roots in perception. In fact, without much hyperbole, memory is aptly characterized as "perception plus the passage of time" (MJ Kahana, personal communication). Just as transmission of genetic information reflects influences from both strands of DNA, most cognitive processes reflect influences from both perception and memory (e.g., Agam & Sekuler, 2007) and their complex and bidirectional mutual influences (Teng & Kravitz, 2019). For example, if perception were compromised by a stimulus that was too weak, or by aging or disease-related change in the eye or ear, the additional resources that must be diverted to sensory processing can limit resources available for memory processes (McIntosh et al., 1999; Bennett, Sekuler, McIntosh, & Della-Maggiore, 2001; Pichora-Fuller et al., 2016) or memory processes might be needed for top-down resolution of perception’s inherent ambiguity (Brady, Konkle, & Alvarez, 2009; Marr, 2010). Finally, it has become ever clearer that the functional organization of the brain does not respect traditional, convenient simple categories, like memory and perception (Buzsáki, 2019). So, a major impetus for this chapter has been our belief that those simple categories need to be reconsidered.

**What’s memory got to do with perception?**

Before turning to the chapter’s main focus, perceptual factors in memory, we will briefly consider the opposite direction of influence, memory’s role in perception. Although the influence extends to all sensory modalities, it is most easily appreciated in the context of vision. Our ability to see and recognize objects and events entails far more than merely detecting the presence of light or simply registering how much light is present. Unfortunately, our experience with cameras may cause us to underestimate the task’s difficulty. As Casati and Cavanagh (2019, p.55) remind us, “Cameras and pictures just collect patterns of light,
dark, and color. They do not know what those patterns represent."

Figure 1 suggests that inferential, probabilistic operations are required to transform a distribution of light on the retina into knowledge of which objects and object arrangements were responsible for that distribution. According to the figure, visual processing begins with the extraction of basic, local features, such as orientation and color. Complex inferential processes then organize these features into higher-order structures, such as surfaces and depth planes. Because the basic features that were extracted can be organized in many different ways, perception must solve an ill-posed inverse problem, which depends upon the availability and use of *a priori* constraints (Pizlo, 2001; Kersten, Mamassian, & Yuille, 2004; Zeki & Chen, 2019). Examples of such constraints (priors) include the Gestalt principles of organization (e.g., similarity, good continuation or smoothness, proximity) (Wertheimer, 1938). Whereas Köhler (1947, p.132) insisted that the Gestalt principles depend on innate physiological properties of the brain, Brunswik and Kamiya (1953) proposed an attractive alternative, namely that the Gestalt organizing principles mirror the statistics of the natural world, and that they may be the product of "generalized probability learning." Subsequently, Brunswik and Kamiya’s hypothesis has been bolstered by detailed, quantitative studies of the statistics of the visual world (Elder & Goldberg, 2002; Geisler, 2008), and by demonstrations that important properties of the visual system can be shaped by the statistics of the sense-stimulating environment (e.g., Ramachandran, 1988; Kersten et al., 2004).

There are many other examples of perception’s sensitivity to the memory of what had previously been experienced. A few perceptual phenomena will suffice to make that point. First, consider perceptual hysteresis. Best known from the physics of ferromagnetism, "hysteresis" refers to the dependence of the state of a system on its history. In perceptual hysteresis, a perceptual state is influenced by the perceiver’s history, that is, by some stored memory for what had just been experienced. For example, if you view a video in which one face slowly morphs into another, you will continue to perceive the initial face well
Figure 1. Simplified schematic illustrating Casati and Cavanagh (2019)’s description of how memory and knowledge can influence perception, in this case, object recognition. Visual processing begins by extracting measurements of simple local features, such as edges, colors, motion, and depth. This is followed by inferential processes in which priors, such as memory for and knowledge about objects’ materials, lighting, and surface structure, guide the reconstruction of a three-dimensional scene most likely gave rise to the distribution of light on the retina. (Figure courtesy of Aidan J. S. Bennett)

beyond the morph’s mid-point -- the initial percept, with which you have a history, persists. Perceptual hysteresis has been demonstrated in many domains including perception of motion (Williams, Phillips, & Sekuler, 1986; Nawrot & Sekuler, 1990; Hock & Ploeger, 2006), binocular depth perception (stereopsis) (Fender & Julesz, 1967; Piantanida, 1986), and binocular rivalry, which results when different stimuli are presented to each eye (Buckthought, Kim, & Wilson, 2008). Notably, perceptual hysteresis is not limited to vision. Comparable effects have been reported for pitch perception (Chambers & Pressnitzer, 2014), for temporal coordination of auditory and visual stimulation (Martin et al., 2014), and for judgments of tactile stimuli (Thiel et al., 2014).

Second, memory influences the perception of illusory, subjective contours (Kanisza figures; Wallach & Slaughter, 1988), and can cause a
two-dimensional stimulus to appear to rotate in three dimensions (the kinetic depth effect; Wallach, O’Connell, & Neisser, 1953). Subjects also misperceive the true direction of motion of a moving stimulus if, while viewing that stimulus, they are holding a different motion direction in visual working memory (M.-S. Kang, Hong, Blake, & Woodman, 2011).

Finally, memory’s impact on perception is represented by scores of examples of lingering effects of previous experience, as seen in visual and auditory after-effects, perceptual learning, and priming (Tulving & Schacter, 1990). In some cases, the effects are relatively short-lived, as in some priming experiments where exposure to one stimulus influences a response to a subsequent stimulus (Biederman & Cooper, 1991), sometimes without conscious awareness (e.g., A. B. Sekuler & Palmer, 1992; A. B. Sekuler, Palmer, & Flynn, 1994). In other cases, the effects can be remarkably long-lasting. For example, several minutes’ viewing during which two different combinations of color and orientation alternate, can alter color perception for days or even months (the McCollough effect; McCollough, 1965; Jones & Holding, 1975). Also, even relatively limited time spent training to discriminate textures or faces can produce effects lasting well over a year (Hussain, Sekuler, & Bennett, 2011). Having highlighted just a few examples of memory’s influence on perception, we turn now to our main topic: ways in which perceptual factors are important for memory and memory research.

**Choosing the right stimulus**

When it comes to psychophysical or perceptual investigations, choosing the right stimulus might be the single most consequential choice a researcher makes (Stevens, 1951, p.31). That choice is equally important for studying other aspects of cognition, including various forms of memory. Knowing that perception research has benefitted considerably from careful attention to stimulus choice and control, R. Sekuler and Kahana (2007) argued that a stimulus-oriented approach to memory research could offer significant theoretical advantages. This chapter expands on that point, emphasizing developments since that publication, and
extending the discussion to auditory memory, and to paradigms beyond measurements of simple short-term memory.

Some key factors to consider when choosing stimuli to be employed in a memory study include:

- The overall perceptibility of stimuli
- Effects of stimulus discriminability and similarity, and
- Ways in which stimuli may interact with verbal representations.

After appropriate stimuli have been identified, there are multiple ways in which their perceptual attributes can be manipulated for additional insights into the mechanisms that support memory. Here we explore several such approaches, including adding noise (variability) to stimuli, parametrically varying stimulus discriminability, and going beyond traditional memory assays to estimate what a memory looks like.

**Controlling stimulus perceptibility**

One can remember a stimulus only if it had first been perceived. To guard against imperceptible memoranda a researcher could measure subjects’ detection thresholds for the memoranda, and then present them at some multiple of threshold. That solution would be a good first step, but is flawed. To gauge memory under optimal conditions, stimuli must be not just seen or heard, but easily seen or heard. This is particularly important for memory research whose focus is developmental, as the following shows.

**Life-span challenges to controlling stimuli**

Sensory systems, at either end of the life-span, are likely to diverge from those of the young adult humans whose performance is most commonly the target of memory researchers. Early on in life, sensory systems have not yet reached adult levels; and later in life, sensory systems tend to undergo a host of age- or disease-related changes. In vision, these changes affect visual acuity and contrast sensitivity (Owsley, Sekuler,
& Siemsen, 1983), and many other perceptual attributes (R. Sekuler & Sekuler, 2000). In audition, the best known analogous changes are ones that disproportionately impact high-frequency sensitivity (presbycusis), which, in everyday activities diminishes ability to distinguish consonants from one another (Phatak, Yoon, Gooler, & Allen, 2009). Most evidence for a connection between age-related visual changes and age-related changes in cognition is correlational, and therefore open to alternative explanations. However, Monge and Madden (2016) summarize evidence from interventional studies showing that manipulation of vision’s quality affects the cognitive performance of both younger and older adults.

Toner et al. (2012) have made a strong case for the use of what they call "vision-fair" tests in neuropsychological assessment. Fully implementing vision-fair stimulus materials would take account of stimulus factors including age-related changes in retinal illuminance, sensitivity to spatial details and contrast, and slowed reading speed when spacing of text is suboptimal (Liu, Patel, & Kwon, 2017). Additionally, it is important to consider other factors that limit perception as one ages, including declines in older adults’ ability to multitask (A. B. Sekuler, Bennett, & Mamelak, 2000), and the decreased rate at which older adults’ brains encode perceptual information (Kline & Birren, 1975; Rousselet et al., 2009, 2010). Such age-related limitations would impact processing of a wide-range of stimuli and tasks, including the processing of words used as memoranda, or for delivery of instructions to subjects (Akutsu, Legge, Ross, & Schuebel, 1991; Yu, Cheung, Chung, & Legge, 2006).

So, unless age-related sensory changes are taken into account, or even better, compensated (Amick, Cronin-Golomb, & Gilmore, 2003; Cronin-Golomb, Gilmore, Neargarder, Morrison, & Laudate, 2007; Gilmore, Cronin-Golomb, Neargarder, & Morrison, 2005; Groth, Gilmore, & Thomas, 2003), those changes can easily be misinterpreted as deficits in cognition. For example, some age-related decrease in face recognition ability --seemingly a decrease in memory-- actually arises from decreased visual function (Konar, Bennett, & Sekuler, 2013; Boutet, Taler, & Collin, 2015).

Analogous challenges arise when auditory memory is assessed in
subjects of various ages. Because some hearing loss is so common in aging, older adults’ poor memory for auditory memoranda is easily misattributed to central cognitive changes when it actually results from changes peripherally, in the ear (Lin & Albert, 2014). Moreover such causal misattributions can arise with even relatively modest age-related hearing loss (Rabbitt, 1990).

It is clear that an individual who has a difficult time seeing or hearing something will be challenged when having to remember that thing. However, simply ensuring that both instructions and stimuli are visible or audible may not be enough to level the playing field for some older adults. Sometimes older subjects manage to compensate for moderate sensory loss by exerting extra effort in processing a stimulus (Pichora-Fuller et al., 2016; Wingfield, 2016; Ayasse, Lash, & Wingfield, 2016). That extra effort can cut into the cognitive resources available for other cognitive domains, including memory (Piquado, Benichov, Brownell, & Wingfield, 2012; Lin & Albert, 2014; Phillips, 2016; Peelle & Wingfield, 2016). Even when performance on a simple perceptual discrimination task, with minimal memory requirement, is comparable across age groups, the neural networks that support performance can differ significantly. Older brains may undergo compensatory reorganization so that a simple perceptual task engages regions such as the hippocampus, which normally are thought to be associated with memory processes rather than perceptual ones (McIntosh et al., 1999; Della-Maggiore et al., 2000; Bennett et al., 2001). Importantly, if older adults must devote greater cognitive resources to perception, they may miss opportunities to acquire and process information. Over time, this could promote memory and cognitive decline (Ge et al., 2021; Kim et al., 2020).

**Stimulus discriminability and similarity**

In addition to ensuring that individual stimuli are easily perceptible for a given participant, it is critical that the researcher also consider the related concepts of stimulus discriminability and stimulus similarity. It is not enough to ensure that stimuli can be detected, they also must
be discriminable from other memoranda under consideration. The next two sections address the issues of discriminability and similarity in more detail, highlighting the potential difficulties in interpretation that arise, as well as the opportunities for exploiting stimulus similarity to derive novel insights about memory.

**Importance of Discriminability**

Weak or noisy sensory processing can undermine the quality of any stimulus representation, thereby promoting confusion among imperfectly perceived stimuli, and producing errors in recall or recognition that may masquerade as imperfections of memory (e.g., Majaj, Pelli, Kurshan, & Palomares, 2002; Boutet, Dawod, Chiasson, Brown, & Collin, 2019). As a result, stimulus discriminability should be assessed and verified before a memory experiment begins. Individual differences in accuracy of sensory processing can compound the problem, with differences in perception being mistaken for individual differences in memory quality. This can be particularly challenging when a researcher wants to compare memory quality across age groups. Without taking differences in discriminability into account, any subject’s performance would conflate differences among individuals’ memory ability with differences in their perceptual ability. To avoid this confound, before memory testing begins, perceptual ability can be assessed individually for each subject, for example, by measuring discrimination performance with some or all the stimuli that will be used as memoranda. In addition to ensuring that results from memory experiments with visual stimuli are unambiguously interpretable, these pre-experiment, psychophysical measurements allow memoranda to be scaled in perceptual units, such as just noticeable difference (JND) units of discriminability.

Consider one application of this approach, a recognition test in which a list of study items, $S$, is followed by a single probe item, $P$. The subject makes a binary ("Yes" or "No") judgment of whether $P$ replicates one of the study items. In such experiments, performance depends on variables such as the number of items contained in $S$, the retention interval, between the presentation of items in $S$ and the presentation of
PERCEPTUAL ATTRIBUTES

$P$, the serial position within $S$ that was occupied by an item that matched $P$, the characteristics of $P$, and so on. Importantly, the ability of a non-matching $P$ to attract false positive responses will vary systematically with the number of JNDs in memory that separate it from a member of $S$. A recognition study by Zhou, Kahana, and Sekuler (2004) is an example of this general approach. On each trial in their experiment, subjects saw two brief study textures ($S_1, S_2$), which were presented one after the other. After a short retention delay, a test probe, $P$, appeared, and the subject judged whether $P$’s vertical spatial frequency (essentially, the density of the vertical bars) matched the vertical spatial frequency in either study item. (Although the horizontal bars’ spacing varied randomly from trial to trial, their spacing was consistent within a trial, but irrelevant to subjects’ task.) Zhou et al. used their results to examine how the study stimuli were represented in memory, that is, whether memory comprised distinct exemplars (e.g., Nosofsky, 1984; Hintzman, 1988), a prototype reflecting a weighted mean of items in $S$ (e.g., Reed, 1972), or some combination of these two.

Zhou et al. (2004) fit their results with a simple three-parameter model. The model assumed that the remembered spatial frequency ($S'$) of some study item ($S$) is a random deviate from a normal distribution centered on the spatial frequency of $S$. Because Weber’s Law holds for spatial frequency (Campbell, Nachmias, & Jukes, 1970), $\epsilon'$, the standard deviation of the distribution of $S'$ can be expressed as a proportion of $S'$ (the coefficient of variation). On any trial, $S_1' \sim N(S_1, \epsilon_1)$, and $S_2' \sim N(S_2, \epsilon_2)$, where $\epsilon_1$ and $\epsilon_2$ are the coefficients of variation, and $N$ is a zero-mean Gaussian random distribution. For simplicity’s sake, the covariance of $S_1'$ and $S_2'$ was set to zero. Subjects’ binary judgments reflect the relationships among $P'$, $S_1'$ and $S_2'$. When $|P'-S_1'| \leq \kappa$, the model generates a "Yes" response; otherwise the model produces a "No" response.

With $P$’s spatial frequency varied systematically in equal discriminability steps (JNDs), Zhou et al.’s model generated what that they called a “mnemometric function” for each study item. In this function, the probability of a “Yes” response (that is, a judgment that $P$ matches
Figure 2. Sample texture stimuli from two trials. In the upper panel, the spacing of vertical bars in the probe stimulus \( P \) differs from the spacing in the study items \( S_1, S_2 \). In the lower panel, the spacing of vertical bars in \( P \) replicates the spacing in one study item, \( S_2 \). As noted in the text, horizontal spatial frequency varied randomly from trial to trial, but within a trial was both constant and task-irrelevant. (Figure from Zhou et al., 2004)

one of the study items) is plotted against the JND difference between \( P \) and that study item.

Mnemometric functions are shown in Figure 3 for the experiment’s eight conditions. Essentially, the mnemometric function comprises a snapshot of the distribution of memory strength associated with the study item, and the characteristics of mnemometric functions track changes in experimental conditions. For example, mnemometric functions show a substantial recency effect, with the mnemometric function for the more recent study item \( S_2 \) displaced less than the mnemometric function corresponding to the earlier study item \( S_1 \).

In a follow-on experiment, Huang and Sekuler (2010) used a pre-stimulus cue to direct subjects’ attention to one of two adjacent Gabor study items. After a brief delay, subjects adjusted a variable test Gabor to match the remembered spatial frequency of the attended stimulus. The resulting matches showed that the remembered appearance of
**Figure 3.** Mean Pr("yes") responses versus probe spatial frequency in each of eight conditions. Each trial’s data were normalized by setting the arithmetic mean of $S_1$’s and $S_2$’s frequencies on that trial to zero. $P$’s position is plotted against multiples of threshold units. In the four graphs at the left, $S_2$’s spatial frequency was higher than $S_1$’s; in the four graphs at the right, the reverse was true. In each panel, the normalized spatial frequencies of $S_1$ and $S_2$ are indicated by dark vertical lines. In panel A-D show results for various differences between values of the two study items, $|S_1 - S_2| = 1, 2, 4, \text{ and } 8$ threshold units, respectively. Results, represented by filled circles for lures and open circles for targets, were averaged across ten subjects, and sorted into bins one threshold unit wide. Values predicted by the three-parameter model are shown by the dotted line (parameters: $\epsilon_1 = 0.2, \epsilon_2 = 0.115, \text{ and } \kappa = 1.3$). Error bars show 95% confidence intervals. (Figure from Zhou et al., 2004)
the attended Gabor was influenced by: (1) the relationship between its spatial frequency and the spatial frequency of the accompanying, task-irrelevant non-target stimulus; and (2) the average spatial frequency of Gabors seen on previous trials. These two influences, one reflecting the accompanying non-target stimulus and another reflecting a previous visual experience, represent quite different time scales, but were approximately additive in their effects, each operating as an attractor for remembered appearance (Anton-Erxleben & Carrasco, 2013).

With paradigms and memoranda that differed from those in Huang and Sekuler, Orhan and Jacobs (2013) used mnemometric functions for their initial characterization of subjects’ performance in a change-detection paradigm (Luck & Vogel, 1997). On each trial, two brief stimulus presentations were separated by just one second. Some stimulus attribute might undergo a subtle change from the first presentation to the second; the subject’s task was to detect whether there had been a change. In one experiment, each stimulus comprised a pair of squares whose horizontal and vertical locations could vary; in another experiment, each stimulus comprised a pair of Gabor patches whose orientations could vary. Modeling results from both experiments revealed a correlation between the memory representations of a stimulus’ two distinct components, and showed that the correlations decreased with either the spatial distance between components (in one experiment), or the orientation difference between them (in the other experiment).

Not all memoranda are as easily scaled in terms of perceptual discriminability as the stimuli were in Zhou et al. or Orhan and Jacobs. In fact, complex multidimensional visual memoranda, such as images of faces, natural images of scenes, challenge that approach. Of course, perceptual differences among such stimuli involve not just obvious, high-level features, but also low level dimensions (e.g., color, contrast, brightness, and statistics for orientation and spatial frequency content). Because memory for complex visual memoranda is sensitive to differences among the low-level stimulus properties of such stimuli, those differences can introduce confounds into experiments. For example, when using face stimuli in a memory task, researchers want their mea-
measurements to reflect memory for face identity or emotion, not simply a face’s contrast, size, or orientation (J. Gold, Bennett, & Sekuler, 1999; A. B. Sekuler, Gaspar, Gold, & Bennett, 2004a). Researchers wanting to study memory for complex visual memoranda could exploit available tools that measure and equate those low level properties (Willenbockel, Sadr, et al., 2010; Bainbridge & Oliva, 2015).

Exploiting Similarity

Although a lack of discriminability among target stimuli can lead to difficulties in disentangling the relative impacts of perception and memory, researchers can use perceived similarity of stimuli as an assay in probing memory. As many investigators have noted (e.g., Blough, 2001; Edelman & Shahbazi, 2012; Jiang, Lee, Asaad, & Remington, 2016), similarity is central in many areas of psychology, including perceptual memory. Moreover, similarity shapes a good deal of our behavior both inside and outside the laboratory. It underlies generalization, various kinds of errors of memory, and perceptual organization. When coupled with an appropriate theoretical perspective, “[S]imilarity provides an elegant diagnostic tool for examining the structure of our mental entities and the processes that operate on those entities... [providing] a powerful, if indirect, lens onto representation-process assemblies” (Goldstone, 1994, p. 3).

For many theoretical purposes, measurement and parametric manipulation of similarity confer substantial advantages. For example, parametric manipulation of stimulus similarity can aid the formulation and examination of links between memory and its physiological substrates (van Vugt et al., 2009). But any attempt to define “similarity” must deal with the fact that similarity is a psychological construct, like hue, saturation or brightness. Similarity should not be defined solely in terms of physical properties; how those physical properties are perceived matters. And numerous demonstrations show that the psychophysical transformation can be complex. Points in case: the complex relationship between wavelength and hue, between physical and perceived visual contrast, or between patterns of facial muscle
contractions and emotional expression.

Similarity’s influence extends over multiple stages of a memory task. For example, similarity can be influential during initial, encoding stages of memory, but it also can be influential during a subsequent, recognition stage (R. Sekuler & Kahana, 2007). Although the distinction between early and later stages was initially suggested by detailed modeling of behavioral results, it has been supported by analysis of subjects’ EEG signals. With intracranial recordings from humans, the two stages of similarity generated distinctive neural signatures (van Vugt et al., 2009).

Depending upon a study’s theoretical and empirical goals, it could be advantageous to explore memory with stimuli that are embedded in a known similarity space. The geometry of such a space represents each item’s similarity to every other item (R. Sekuler & Kahana, 2007). To specify the degree of “similarity” among stimulus objects requires psychophysical measurements, such as ones used for multidimensional scaling (MDS; Kahana & Bennett, 1994; Hout, Papesh, & Goldinger, 2013). As suggested by Shepard’s universal law of generalization for psychological science (1987), there is virtually no limit to the dimensions along which stimulus similarity can be measured, controlled and exploited for research (Hebart, Zheng, Pereira, & Baker, 2020). Unsurprisingly, memory researchers have made good use of stimulus similarity defined along a variety of dimensions, including spatial frequency (Kahana & Sekuler, 2002; Nosofsky & Kantner, 2006; Viswanathan, Perl, Visscher, Kahana, & Sekuler, 2010; Dubé, 2019), brightness and sequences of brightnesses (J. M. Gold, Aizenman, Bond, & Sekuler, 2014), temporal patterns of stimulation in multiple modalities (H. Kang, Lancelin, & Pressnitzer, 2018), direction of visual motion (Pavan et al., 2019), rate of acoustic spectrotemporal variation (Visscher, Kaplan, Kahana, & Sekuler, 2007), facial expression of emotion (Galster, Kahana, Wilson, & Sekuler, 2009), auditory pitch and pitch sequences (Cousineau, Demany, & Pressnitzer, 2009), and subtle acoustic patterns in random noise (Agus, Thorpe, & Pressnitzer, 2010; Agus & Pressnitzer, 2013), to cite just a few. MDS can be time-consuming when standard methods like pairwise-comparison or
triadic judgments (Kingdom and Prins, pp. 49-50) are used with many stimuli; less-well known, but quite efficient alternatives are available (Levin, 2000; Knoblauch & Maloney, 2012; Hout, Goldinger, & Ferguson, 2013).

![Figure 4](image.png)

*Figure 4.* Sample face-name pairs from Expt. 1 of Pantelis, van Vugt, Sekuler, Wilson, and Kahana (2008)

To illustrate the usefulness of the stimulus’ similarity structure, consider Pantelis et al.’s study of memory for face-name associations (2008). The researchers wanted to know how similarity among faces influenced ability to learn and consistently connect an individual face with a name. For one of their experiments, stimuli were 16 achromatic male faces (for examples, see Figure 4), which were constructed following a method developed by Wilson, Loffler, and Wilkinson (2002). Similarity ratings, gathered using the method of triads, were processed via MDS, which yielded the similarity space shown in Figure 5. Each face was then paired with a common, one-syllable American male name. The experiment comprised ten blocks of alternating study and test phases. In study phases, each face was accompanied by a spoken presentation of its associated name. In test phases, faces were presented one at a
time, and subjects had to say the name that had been paired with each one. After ten study-test alternations, subjects produced the correct name at an average rate of 83%. Importantly, the actual proportion correct varied considerably from face to face, depending upon how many neighbors a face had in the MDS space. Faces with few neighbors were named correctly at a far higher rate than faces with many neighbors. Moreover, when a test name was incorrectly recalled, that name tended to be the name of a face that lay near the test face in MDS space. The authors interpreted their results as arising from associative interference among items in a subject’s study set, with the amount of interference seeming to follow the metric properties of the MDS similarity space. Thus, understanding the perceptual similarity among stimuli enabled a clear interpretation of the associative memory results, and could lead to predictions for memory performance with other stimulus sets.

Verbal-perceptual interactions

Because our focus here is on adult human memory, one particularly distinctive adult human ability must be taken into account. Adult humans are incorrigible in generating and relying on verbal descriptions of whatever they see or hear. Presented with a visual (or auditory) memorandum, an adult human subject is virtually certain to produce, store, and perhaps exploit not only some sensory representation of what was just experienced, but some symbolic or verbal description as well (Paivio, 1986; Sternberg, 2003). As a result, memory can be supported either by verbal or by sensory codes individually, or together -- at the same time, or in shifting, dynamic interactions over time.

Stimuli that seem to qualify as purely "visual" or "auditory," in the sense that they are delivered to the eyes or ears, trigger not only sensory responses, but verbal or categorical responses as well. For example, imagine that a subject in a memory experiment is shown a photograph of a cat. Yes, that image on the subject’s retina will be processed through multiple stages of the visual system, and is visual in that sense, but that processing also will produce a verbal label, "cat,"
or perhaps something more detailed, “small, calico American shorthair, with one black hind paw.” Stimuli are likely to give rise to dual codes: sensory and verbal. That is why the sight of a familiar scene, like the sandy shore of a pond, evokes a host of associations not immediately present in the visual stimulus. Those associations can aid memory of the scene, but can also generate false memories of things not actually present (Miller & Gazzaniga, 1998; Děchtěrenko, Lukavský, & Štipl, 2021), much like lists of words can do in the Deese-Roediger-McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995).

**Figure 5.** Sixteen face-stimuli plotted in their 4-dimensional MDS-derived face space, from Pantelis et al. (2008). The 4th dimension, which like another of the dimensions correlated with a combination of lip thickness and head shape, is indicated by the brightness of a face. Note that subjects were given no instructions about what stimulus attributes should be used as a basis for their judgments.
Verbal overshadowing

The classic study by Carmichael et al. (1932), showed how verbal labels can influence recall of previously seen visual stimuli. Subjects viewed a series of twelve ambiguous geometric figures, each accompanied by a meaningful verbal label. Groups of subjects received different labels, as shown by the sample stimuli and labels in Figure 6. After seeing all twelve figures and labels, subjects were asked to reproduce as many as they could, in any order they chose. Some, but not all reproductions showed evidence of what the researchers called "figure assimilation," that is, the reproduced figures seemed to be shifted away from what subjects had actually seen, toward greater conformity to the label that had accompanied the figure. However interesting Carmichael et al.’s results may be, their study fails to capture what happens when, after a stimulus has disappeared, subjects verbalize what they had seen in an effort to preserve their memory. That situation is more closely represented in the so-called verbal overshadowing (VO) paradigm. In one influential VO experiment (Schooler & Engstler-Schooler, 1990), subjects saw a video of a simulated bank robbery. Immediately afterward, some subjects were asked to describe the appearance of the robber’s face,
while other, control subjects performed an unrelated task. Later, in a recognition test, subjects tried to pick out the robber’s face from a set of similar, lure faces. Subjects in the verbalization condition did less well than subjects in the control condition (Schooler & Engstler-Schooler, 1990, Expt 1). A similar result was produced with color patches as stimuli. Although VO has been most widely studied with visual stimuli, such as faces and color (Souza & Skóra, 2017), there is some evidence that it can affect memory for auditory stimuli (such as voices) or even tastants (Hatano, Ueno, Kitagami, & Kawaguchi, 2015). In their report of a collaborative replication study, Alogna et al. (2014) noted that VO has significant implications for police practices, as well as for understanding the mechanisms underlying memory processes. Because verbal description can distort memory, Alogna and colleagues suggested that the testimony of eyewitnesses who previously were asked to describe a suspect should be weighted less than the testimony of witnesses who had not previously given a description.

The potential importance of verbalization was made explicit in Baddeley and Hitch’s influential model of working memory (1974), which included parallel perceptual stores: a phonological loop, which stores verbal content; and a visuo-spatial sketchpad, which holds visuo-spatial data. The recognition that working memory can reflect contributions from either of these paired buffers encouraged the development of dual task techniques in which modality-specific interference is used to clamp or suppress one of the buffers. This basic experimental strategy has even been used to selectively interfere with memory of sentences that had been learned with the aid of motor imagery (Saltz & Donnenwerth-Nolan, 1981). Of dual tasks interference techniques, articulatory suppression has been the most commonly used. The idea is to use an articulatory task to control verbalization of a stimulus, either during encoding or in a post-stimulus delay. Although articulatory suppression has been shown to be effective, recent studies suggest that it might be less effective than originally thought (Toppino & Pisegna, 2005; Doherty et al., 2018; Norris, Butterfield, Hall, & Page, 2018; Norris, Hall, Butterfield, & Page, 2019).
The twin buffers in Baddeley and Hitch’s model have rough parallels in dual-coding theory (DCT) of memory and cognition (Paivio, 1986). DCT postulates that information originating from visual stimulation can be represented visually and/or verbally, each reflecting distinct, but related, information. The redundancy afforded by dual codes boosts the chances that a memory can be recalled if one of the codes is lost or degraded. That would be particularly advantageous if the two codes each entailed considerable detail and if the two codes were well matched in coverage. At the same time, though, such redundancy can challenge researchers who want to connect performance on some memory task to the sensory properties of the instigating memoranda (e.g., Okazawa, Sha, Purcell, & Kiani, 2018). The challenge introduced by dual codes is compounded by the fact that very small changes in a stimulus or a task can radically alter the balance between the usefulness of the two codes. An excellent demonstration of this claim comes from study that purported to show that Russian and English language speakers differ in their perception of colors in the visible spectrum’s short-wavelength (“blue”) region.

Figure 7. Panel A. Examples of color samples like ones used by (Winawer et al., 2007) to test purported perceptual effects of linguistic categories. Note that color squares were separated by ≈5° visual angle. Panel B. Same stimuli abutting; language effect disappeared when color samples abutted one another. (Figure by Robert Sekuler)

Unlike English, the Russian language enforces an obligatory distinction between lighter blues (goluboy) and darker blues (siniy), but has no general term for “blue.” Winawer et al. (2007) asked whether
this linguistic difference led to differences in how color was perceived. For an answer, subjects saw arrays of three small squares whose colors were from the blue portion of the spectrum. For each array, subjects selected which of the two comparison squares matched the color of a standard square. One of the two comparison stimuli was a perfect color match to the standard; the other comparison stimulus was chosen either to lie on the same side of the *siniy/goluboy* border as the standard, or on the other side. Only Russian speakers’ color matches were slowed when the color samples differed but fell within the same Russian linguistic category; that was not true for English speakers. As Winawer put it in a personal communication, “If you have two codes, one visual and one verbal, and you are making a comparison, either across time or across space, then the dual code can speed your performance when it’s helpful (cross-category comparison).” Of course, when the two codes are not in agreement, comparisons over time or across space would be slowed, which is what the researchers found. A follow-up study eliminated the small spatial separation between stimulus squares (see 7B), which eliminated the difference in color matching between Russian and English speakers (Witthoff, 2007). Apparently, even a small separation between stimuli allowed linguistic categories to influence the speed of decision making, rather than influencing perception directly (Klemfuss, Prinzmetal, & Ivry, 2012).

The availability of multiple codes should be taken into account in the design or interpretation of memory experiments. For example, suppose that in attempting to compare visual and auditory memory, researchers opted to use familiar sounds as auditory stimuli and pictures of common objects or scenes as visual stimuli (Cohen, Horowitz, & Wolfe, 2009). Certainly, such comparisons will likely reflect more than sensory systems’ memory capacity; they will also reflect pre-experiment familiarity, the ease with which the stimuli can be given consistent labels, and the experimental tasks used to assess memory. To circumvent such complications, memory researchers might avail themselves of stimuli and tasks that are meant to curtail complications from verbal mediation. The next sections provides examples of such stimuli and tasks.
Probing memory with stimulus noise

Like members of many other animal species, humans possess an amazing ability to register, remember, and exploit the similarities among sequences of sensory experiences, and to spot when a pattern changes. One especially useful approach to studying memory for patterns examines the ability to distinguish random stimulus sequences from random sequences onto which some form of structure has been imposed. Because random sequences comprise a large, perceptually-similar pool of stimuli that lack semantic content, their use in research puts the spotlight squarely on pattern-recognition in sensory processing and memory for features that are challenging to identify and process categorically (Kaernbach, 2004).

Figure 8. Two faces that are corrupted by increasing levels of visual noise. Decreasing Signal:Noise ratios, from the top to the bottom of the figure, renders the faces less distinct. (Courtesy of Jason M. Gold)
In perception research, the term “noise” refers to some unpredictable variation in one or more dimensions of some stimulus (Allard, Faubert, & Pelli, 2015). For visual stimuli, noise might entail a random distribution of luminance over space (see Figure 8); for auditory stimuli, noise might entail random variation in sound pressure with time; for somatosensory stimuli, noise might entail random variation in the timing or intensity of successive tactile stimuli. For many decades, noise has been an important tool for sensory researchers. Because different kinds of noise probe different aspects of the computation, noise has been invaluable in identifying the information and computations used by various sensory tasks (J. M. Gold, Bennett, & Sekuler, 1999; Näsänen, 1999; J. M. Gold, Murray, Sekuler, Bennett, & Sekuler, 2005; Murray, 2011). We believe that noise has considerable potential for studying memory formation as well. There are many everyday conditions in which experience makes memorable stimuli that when first presented seemed featureless and unstructured. In fact, “the process by which noise becomes meaningful is foundational to how we interact with the world” (Seitz, 2010). Take, for example, the image in Figure 9. If you have never seen the image before, and we asked you describe the scene, you might answer that was a collection of meaningless black blobs. However, individuals who are familiar with the image, or have been given some prompts, would describe a Dalmatian dog that is facing to the left, head down, sniffing the ground, perhaps in search of her favorite tree. Importantly, once you see the dog emerge from the noise, the image becomes meaningful, and fundamentally changes the way you interpret the scene. Moreover, that change in perception will endure until you next see the stimulus.

One research group has put auditory noise to particularly good use, treating it as vehicle for understanding short-term and long-term memory and learning. In an effort to stimulate readers to consider how noise might be harnessed to serve their own particular research interests, we will summarize the basic methodology and some key findings of that program, followed by a brief consideration of roughly analogous work with visual noise.
Figure 9. Demonstration that a seemingly meaningless stimulus can develop into a recognizable object. See text for explanation. (Figure from van Tonder and Ejima, 2000)

Noise probes of auditory memory and learning In a series of behavioral and brain imaging experiments, Pressnitzer, Agus, and colleagues examined the development of memory for brief samples of acoustic random noise (e.g., Agus et al., 2010; Agus, Carrión-Castillo, Pressnitzer, & Ramus, 2014; Andrillon, Kouider, Agus, & Pressnitzer, 2015). In one of their characteristic experiments, auditory noise stimuli, each one second in duration, were generated from normally distributed random numbers. These stimuli, which the researchers designated N ("Noise") stimuli, had no reliably discernible features, and one exemplar sounded pretty much like any other. On 50% of trials, this base stimulus type was modified by concatenating the first 500 msec of acoustic noise with an exact duplicate of itself; these were called Repeated Noise RN stimuli (see Figure 10). Note that two halves of RN stimuli were presented without interruption. N and RN stimuli were generated afresh for each trial, and were randomly intermixed for testing. After each complete one-second stimulus, a subject judged whether the final 0.5 s of noise samples replicated or differed from the initial 0.5 s of noise samples. This judgment required that subjects compare the second segment of noise against their memory for the first, akin to a same-different task (Kingdom & Prins, 2016, p. 43). At the start of testing, subjects were only slightly better than chance in judging whether the two halves of a
stimulus were the same, and they showed little improvement over many trials. A notable aspect of the task is that it can be performed even if subjects do not know or cannot articulate the way in which one half of the stimulus might differ from the other half (DeCarlo, 2013). This property of the task was important because the nature of the noise stimuli made it impossible for the researchers to anticipate what acoustic "feature(s)" subjects might be able to exploit. In fact, results of a later study by some of the same researchers suggest that different subjects may have exploited different features (Andrillon et al., 2015).

To examine learning and memory for random acoustic noise stimuli, Agus et al. used a variant of Hebb’s repetition paradigm (1961). Randomly interspersed among R and RN stimuli, for each subject, a specific RN stimulus was made to repeat identically once every few trials. This special subspecies of RN stimuli were designated RefRN stimuli, for "reference Repeat Noise." Note that the identity of a RefRN stimulus varied from one subject to another, so that results related to this condition could not be linked to particular low-level stimulus characteristics. Although at the outset of the experiment, responses to RefRN stimuli were indistinguishable from responses to other noise stimuli, with successive repetitions, subjects grew increasingly adept at telling that the two, 500 msec-halves of these RefRN stimuli had repeated. Once these memories were formed, they emerged rapidly: subjects approached perfect performance after just a few repetitions. Moreover, once memory had formed for a particular RefRN stimulus, it was retained for more than a week without additional training or rehearsal.

It is surprising that memory for a RefRN stimulus developed as rapidly as it did. After all, several stimuli were randomly interspersed between successive occurrences of a RefRN stimulus. The random stimuli interspersed between successive presentations of the RefRN stimulus should have generated massive interference, which would have limited the build up of a persistent representation of a RefRN stimulus (Dubé, Zhou, Kahana, & Sekuler, 2014; Jonides et al., 2008). The absence of control measurements makes it impossible to gauge whether and how much interference was actually generated. Additionally, according
Figure 10. Spectrographic representations of noise stimuli presented on seven successive trials. In each spectrogram, the horizontal axis represents the stimulus’ one second duration, and the vertical axis represents acoustic frequency (on a log scale). The stack of colors at any one point along the time axis represent relative intensities at various frequencies. Note that for all stimuli whose labels include RN (the second, third, fourth and six stimuli), the first half of the stimulus is identically repeated without break in the second half. That is not true of stimuli designated N; for them, the two half stimuli differ. Finally, note that the stimuli designated RefRN, not only are the two half of a stimuli identical, but the same entire noise stimulus is replicated on different trials. See text for details. (Adapted from Agus, et al., 2010)
to some memory models (Kahana & Sekuler, 2002; R. Sekuler & Kahana, 2007), any interference that did build up would have had greatest impact on retrieval processes.

To identify the neural substrates for this unusual form of memory, Kumar et al. (2014) increased the durations of their noise stimuli. Multivoxel Pattern Analysis of signals from functional magnetic resonance imaging revealed memory for particular noise stimuli in two brain regions, the hippocampus and in the planum temporale, an area involved in music and language. Although it has long been known that the hippocampus was involved in various forms of memory, this may be the first demonstration that it is involved in memory for low level acoustic features as well. While acoustic noise stimuli are very different from stimuli more traditionally used to study memory and learning, including words and pictures, it is noteworthy that regardless of stimulus type, memory seems to entail some neural representation in hippocampus (Bennett et al., 2001; Della-Maggiore et al., 2000; McIntosh et al., 1999).

**Example studies with visual noise** A number of studies have been reported on memory for visual noise. Here we concentrate on those studies that closely parallel work just described for auditory memory. Stimulated by that research, J. M. Gold et al. (2014) examined vision’s ability to support both short term memory and robust, longer term memory, as was seen for auditory noise. Adapting the auditory noise paradigm described above, Gold and colleagues assessed subjects’ ability to learn arbitrary short sequences generated by temporally modulating stimulus luminance. Because of vision’s relatively limited temporal sensitivity, their stimulus sequences comprised a series of just eight luminances across a 1 s presentation, without gaps between luminances in a sequence. For half the sequences the last four luminances replicated the first four; these were analogues to the RN stimuli in Agus et al. Throughout the experiment, subjects had only to judge whether the first half of a one-second long stimulus was replicated in its second half. No other judgment was required. Each subject was assigned a unique special RefRN stimulus, which recurred intermittently, every few
trials. Luminances were sampled from a limited range, which controlled the possibility that subjects could perform the task by spotting some particularly distinct luminance. The three stimulus types are illustrated in Figure 11.

![Figure 11](image_url)

**Figure 11.** Examples of Experiment 1’s three kinds of stimuli (Noise (N), Repeated Noise (RN), and Reference Repeated Noise (RefRN)). Note that the second half of an RN stimulus recycles the luminances from the first half of the stimulus; note also that an entire RefRN sequence repeats identically from one trial to a later trial. Luminances in the stimuli as shown here do not depict the actual luminances presented in the experiment. (Adapted from Gold et al., 2014)

Over successive presentations of RefRN stimuli, correct judgments increased, though more gradually than had been found with auditory noise. Subjects retained this enhanced performance for RefRN stimuli without loss for at least 24 hours, unfortunately the only retention delay that was tested. There are several routes by which subjects could have performed the basic task of judging whether half of a stimulus replicated or not. Subjects might have made difficult item-by-item comparisons of the corresponding luminances in a sequence, comparing items 1 and 5, 2 and 6, 3 and 7, and 4 and 8; subjects also could have processed a sequence globally, as a complex pattern or temporal compound (Warren & Bashford, 1993). To evaluate these possibilities and to identify what cues subjects most relied on when making their judgments, Gold and colleagues applied reverse correlation analysis (Ahumada & Marken, 1975; Ahumada, 1996; Neri, Parker, & Blakemore, 1999; J. M. Gold, Murray, Bennett, & Sekuler, 2000; A. B. Sekuler, Gaspar, Gold, & Bennett, 2004b; Murray, 2011; Nagai et al., 2013). The analysis showed that in judging whether the last half of the stimulus replicated the first half,
subjects placed particular, though not exclusive, weight on the fourth and eighth (final) items in each eight-item stimulus sequence. In other words, instead of comparing item by item all the corresponding items in a stimulus’ first and second halves, subjects based their judgements largely on whether the luminances of just two key items matched, which tended to be the fourth and eighth items of a sequence. Finally, the researchers were careful to characterize subjects’ learning with the word "incidental" rather than the more specific word "implicit." In making that distinction, they acknowledged that they had not done the tests needed to establish that subjects had not explicitly recognized the special status of particular stimulus sequences. The visual noise memory paradigm just described has been extended to spatial arrays of random luminances (Maharjan, Gold, & Sekuler, 2017), and to sequences in which concurrent tone pitch and luminance was either correlated or uncorrelated (Keller & Sekuler, 2015; Aizenman, Gold, & Sekuler, 2018).

Case Studies: Stimuli as Tracers

Like a tracer in medicine (e.g., a dye or a radioactive isotope), a stimulus with well-defined, distinctive properties can be inserted into a memory task, and used to trace the course or define the characteristics of memory process(es) recruited by the task. Exploiting two known effects of stimulus similarity, Visscher, Kahana, and Sekuler (2009) traced the differential carryover of auditory information across successive trials. The study built on the finding that previously acquired information can interfere with memory for subsequently acquired information (Underwood, 1957; Jacoby, Debner, & Hay, 2001; Postle & Brush, 2004). This effect can have consequences in everyday life: inability to suppress or control information that was once task relevant, but is no longer, can significantly limit important cognitive functions (Jonides & Nee, 2006).

In Visscher et al.’s experiments, the stimuli were moving ripple sounds, which had been similarity scaled in terms of subjects’ JNDs. Moving ripple sounds are broadband sounds that vary sinusoidally both in time (with a period of $w$ cycles per second) and in frequency content.
(with a period of $\Omega$ per octave). Earlier, Visscher et al. (2007) showed that short-term memory for these stimuli exhibits strong parallels to short-term memory for visual stimuli, such as oriented sinusoidal gratings. Additionally, ripple sounds share spectrotemporal similarities with the sounds that characterize speech (e.g., Shamma, 2001). Thus, findings with these ripple sounds, which are difficult to verbalize, can be generalized to speech sounds without incurring the complications from verbal labels.

Using these stimuli, Visscher et al. examined carry-over for two distinct sources of information: the characteristics of individual study sounds (item information), and the relationships among study sounds, (study set homogeneity, the reciprocal of inter-item similarity), as previous research had shown that the homogeneity of items in a study set strongly influenced short-term recognition memory (Kahana & Sekuler, 2002; Nosofsky & Kantner, 2006; Pantelis et al., 2008; Viswanathan et al., 2010). On each trial, subjects heard two successive brief, broadband study sounds followed by a broadband probe ($P$) sound. They judged whether $P$ had been one of the study sounds. As the similarity of $P$ to stimuli presented on the preceding trial increased, so did the probability of false recognitions, from 0.4% to 4.4%. This evidence of proactive interference was highly dependent on similarity -- interference was limited to $Ps$ that were very similar to an item on the previous trial. In addition, on other trials, Visscher et al. varied the homogeneity of the study items. The intent was to manipulate the criterion subjects would use in making their recognition judgments. This manipulation’s effect was local; that is, it did not affect the subsequent trial. So, the observed proactive interference likely occurred with no appreciable contribution from a criterion change.

**What do memories look like?**

To estimate what a subject remembers, researchers draw on subjects’ responses as the bases for inferences about memory. Such inferences are necessarily limited by the granularity of the responses. Traditionally,
memory experiments have relied on binary responses, "yes" vs. "no", "old" vs. "new", etc. However, the information represented in visual memory certainly exceeds the one bit captured in a single binary response. As a result of binary responses' limited granularity, estimating the quality of a stimulus as remembered requires the aggregation of several dozen binary responses. An obvious alternative would be to increase the granularity of the available responses, that is, increasing the number of response alternatives from two up to a much larger number. However, it is not clear how well subjects can actually use a large number of categories, at least with good consistency (Yotsumoto, Kahana, McLaughlin, & Sekuler, 2008). In fact, subjects' inconsistency limits the mutual information (Watson, Rilling, & Bourbon, 1964) that can be transmitted with bounded, categorical responses.

In psychophysics, adjusting one stimulus to perceptually match another stimulus is a common way to assess stimulus appearance (Brindley, 1970; Kingdom & Prins, 2016). For example, instead of asking a subject to describe some color patch of fabric or computer display in words, subjects can select from a set of calibrated sample the one that matches perceptually. This operation can be carried out as a side-by-side comparison against calibrated color samples like those produced by Munsell (1905) or Pantone (Pantone Reference Library, 2012). With few notable exceptions, though, matching has not been used to characterize perceptual memories.

An early exception by Harvey (1986) illustrates the approach. Harvey’s subjects viewed a photo of a face and then, a few minutes later, recalled what they had seen. The aim was to learn how the remembered face differed from what had originally been seen. In one of several tests, subjects saw variants of the studied face and rated each one’s similarity to their memory of the studied face. Suspecting the importance of spatial frequency in face recognition, later verified by numerous studies (J. M. Gold et al., 1999; Kornowski & Petersik, 2003; Näätänen, 1999; Wenger & Townsend, 2000; Gaspar, Bennett, & Sekuler, 2008; Willenbockel, Fiset, et al., 2010), Harvey constructed several sample faces by eliminating various frequency bands from the original
face. A subject rated the similarity of each filtered sample face to the memory of the original, study face. The similarity ratings for the sample filtered faces showed that the remembered face resembled a face from which high spatial frequencies (fine details) had been dropped. So, it seems that over time, fine details of a face are selectively lost from memory, a result that is consistent with results with stimuli constructed from two-dimensional Gaussian noise patterns (J. M. Gold et al., 2005).

Harvey’s work constitutes a step toward capturing what the memory of a face or -- of any complex stimulus -- looks like, but it necessarily falls short. After all, no face is completely specified by spatial frequency content alone (Pachai, Sekuler, & Bennett, 2013), and the memory of a face undoubtedly entails more than just spatial frequency information.

Using the method of adjustment in a series of delayed recall experiments, Rivera-Aparicio, Yu, and Firestone (2021) discovered that subjects’ reproductions of previously seen scenes were not pale, faded imitations of their original selves, but instead exhibited an exaggeration of their vividness. Figure 12 illustrates a trial in one of their experimental protocols. A subject first saw a scene that varied in its degree of blur. Then, after a brief delay, the subject adjusted a variable image to reproduce the blur of the original. Subjects seemed to misremember scenes as more vivid (i.e., less blurred) than those scenes had actually appeared when initially presented, just moments earlier. Effects like these, and others described in this section of the chapter, are reminders that what a perceptual memory looks like reflects influences beyond those of early sensory processing.

In a way, the method used by Harvey and by Rivera-Aparicio et al. to define the characteristics of a perceptual memory is analogous to recall measures used to assessed memory for verbal stimuli. A influential study by Wilken and Ma (2004) extended this approach by measuring precision of memory for three different classes of stimuli: color, spatial frequency and orientation. Figure 13 shows in schematic form one trial used to study memory for color.

In the figure, the briefly presented study set comprised $n$ randomly chosen color patches ($n = 2, 4, 6, \text{ or } 8$). These were briefly displayed
**Figure 12.** Sequence of events on a trial in one protocol in Rivera-Aparicio et al. (2021). After a 2-sec. presentation of a blurry image and a 2-sec. delay, a subject adjusted the blur of a variable test image to reproduce what had been seen originally. A typical outcome is illustrated by the righthand image. (Figure courtesy of Aidan J. S. Bennett)

**Figure 13.** Schematic timeline of a typical trial in a task that Wilken and Ma (2004) used to gauge short term memory for color. See the text for explanation. (Figure courtesy of Aidan J. S. Bennett)
at equi-spaced locations around a calibrated color circle. Then, after a 1500 msec delay, a probe square cued the subject to reproduce from memory the color that had been shown at that probe’s location (in the example illustrated, the cued location had been a greenish patch). The subject indicated the remembered color by clicking on the appropriate location on a calibrated color annulus. Analogous experiments were done with memory for other stimulus attributes, orientation and spatial frequency. In those other experiments, the colored squares were replaced by Gabors that could vary in spatial frequency or orientation, and the color annulus was replaced by a probe Gabor whose spatial frequency or orientation could be adjusted by the subject. For all three stimulus types, study stimuli were designed to minimize subjects’ ability to code stimuli categorically or with verbal labels. To accomplish that important aim, study stimuli were chosen randomly from large, perceptually dense sets, that is, sets whose items varied gradually from one another in small steps. For example, study colors were chosen from a color palette of 252 different possible items, rather than from a more limited palette of highly distinctive, readily nameable colors. Spatial frequency study items were randomly selected from a set of 16 Gabor patches, equi-spaced in frequency over a narrow range of 2:1, impeding ability to apply distinct and consistent verbal labels to different stimuli. Similarly, orientations were chosen randomly from possible values spanning 0 to $2\pi$. In orientation measurements, subjects matched the cued orientation by using the right and left arrows on the keyboard to rotate the probe Gabor. In the spatial frequency experiment, subjects used the arrow keys on the keyboard to change the spatial frequency of the probe stimulus.

Figure 14 shows the basic results. Each distribution shows the number of responses on which each stimulus value (color, orientation, or spatial frequency) was chosen by the subject. Responses plotted against error = 0 correspond to memory reports that faithfully reproduce the cued study item; responses plotted against other x-axis values represent the amount by which the reported memory deviated from the cued study item. A symmetrical distribution centered at error = 0 means
that average memory-based recall did not deviate systematically from the cued study item, and that errors in recall, whatever their cause, are equally likely to occur in either direction away from the actual cued study time.

Wilken and Ma’s task design obscured potential, theoretical valuable interactions among the study items’ representations in memory. From trial to trial, the item that subjects would match was drawn randomly from the set of study items on each trial. As a result, subjects’ matches were mute about possible interactions among the multiple memories on a trial, like the interactions reported by Huang and Sekuler (2010), mentioned earlier in the chapter. Differences among the breadth of distributions in Figure 14 are consistent with the well-known limits on short-term memory capacity (Cowan, 2001). Subsequent research with the same basic protocol have drilled down to more detailed explanations, one of which is especially relevant here.

**Figure 14.** Main results from Wilken and Ma (2004) for the three different stimulus type and study set sizes = 2, 3, 4, or 5. "CLUT" stands for color lookup table, a matrix that holds the digital values used by the computer to produce color stimuli. (Adapted by Robert Sekuler after Wilken & Ma, 2004)

Tests of memory for color like those described by Wilken and Ma (2004) use stimuli that occupy known locations in a well defined perceptual color space. Knowing the perceptual properties of color stimuli expands the theoretical reach of the research, and has reinforced the idea that perception of color has a hierarchical structure, which entails both continuous and categorical dimensions. The categories correspond
to a small set of familiar, possibly universal color terms, and memory for color includes both continuous and categorical influences. For example, common naturally occurring or anthropogenic objects may have characteristic colors, which are remembered and influence their perception (e.g., Duncker, 1939; Hansen, Olkkonen, Walter, & Gegenfurtner, 2006). These robust influences, known as "memory colors", afford a useful tool for examining the interplay between perception and memory, and can illuminate the properties of each one. For example, when they analyzed color memory errors, Bae, Olkkonen, Allred, and Flombaum (2015) found that the size and direction of errors were non-uniform. In particular, memory for colors distant from the centers of the salient color categories were shifted toward the centers, introducing a category-dependent variation in error size and direction. Analogous effects are seen in memory-based estimates with other stimulus dimensions, such as orientation (Taylor & Bays, 2018) and spatial location (Huttenlocher, Hedges, & Vevea, 2000).

**Time is of the essence** Of course, when asking what our memories look like, it is important to recognize that our memories are not static, but can change over time. Memory exhibits two distinct forms of time dependence. One is transience (Schacter, 1999), the decreasing accessibility of information in memory over time. Transience is often represented graphically by the well-known forgetting curve, introduced in the earliest systematic studies of memory (Ebbinghaus, 1885). For some stimuli and memory tasks, the loss of information from memory tends to be relatively slow, which makes it easier to measure the time course of forgetting. Notable exceptions in which forgetting is rapid, include sensory memory, like the contents of iconic or echoic stores (Sperling, 1963; Lu, Williamson, & Kaufman, 1992), and verbal recall tasks, like the Brown-Peterson paradigm, which entail considerable interference (Brown, 1958; Peterson & Peterson, 1959; Lewandowsky, Oberauer, & Brown, 2009).

However, it can be more challenging to measure memory’s other form of time dependence, the time needed to encode the stimulus. This
is time associated with transferring the responses of sensory receptors into memory. This aspect of time dependence has been less well studied than forgetting, but because the quality of memory depends upon the quality of encoding, this process is arguably just as important as the process of forgetting. Although encoding can be remarkably fast (e.g., Biederman, Rabinowitz, Glass, & Stacy, 1974; Owens et al., 2018), its speed is not infinite: in fact, careful measurements do reveal a distinct dependence on time (Loftus, Duncan, & Gehrig, 1992).

An excellent demonstration of this dependence was provided by Bays, Gorgoraptis, Wee, Marshall, and Husain (2011)’s study of the temporal evolution of memory precision. On each trial in their study, subjects saw an array of $N = 1, 2, 4,$ or $6$ bars of different colors and random orientations for durations that ranged from $25$ to $1000$ msec. After the array’s presentation, further visual processing was terminated by the presentation of a random pattern mask (Breitmeyer & Öğmen, 2006). One second later, a color “probe” bar appeared at fixation, and subjects adjusted its orientation to match the remembered orientation of the bar in the array with which it shared a color. The response on each trial was converted into an error measure -- the difference in degrees between the orientation as reproduced from memory and the actual probe bar orientation. Then for each combination of stimulus duration and $N$, results were distilled into a measure of recall precision -- the standard deviation (in circular coordinates) of the distribution of errors for that condition. These distributions are similar to ones shown in Figure 14 for stimulus color, orientation, and spatial frequency. Figure 15 shows the results on temporal evolution of memory’s precision. Note that for each value of $N$ (1, 2, 4, or 6), precision clearly increased with exposure duration, rising rapidly at first, and then slowing as it approaches asymptote. A detailed analysis showed that at the shortest exposure durations, when precision was rising rapidly, that initial rise varied inversely with $N$. Bays et al. (2011) noted that this dependence is consistent with models in which multiple stimuli are processed in parallel, competing for transfer to short term memory (e.g., Bundesen, Habekost, & Kyllingsbaek, 2005). Importantly, this operation is distinct
from what would be seen with serial models in which stimuli are selected one at a time for transfer to short term memory. Note that being able to successfully trace the rapid evolution of memory precision over time depended upon a graded, momentary assay of memory -- in this case, reproduction -- rather than a simple binary response.

![Figure 15](image)

**Figure 15.** Temporal evolution of working memory precision, as described by Bays et al. (2011). Recall precision as a function of exposure duration and number of study items (N), with a subset of results shown here. Precision is defined as the reciprocal of the standard deviation of error in the reproduction task. Error bars indicate ±1 standard error. Dashed lines are from first-order differential equations whose parameters best fit the temporal evolution of recall precision at each array size. (Figure courtesy of Aidan J. S. Bennett)

Response distributions like ones used by Bays et al. (2011) and illustrated in Figure 14 are useful for constructing and testing theories of memory, provided that one understands the variables that shape the distributions. However, several obstacles stand in the way of linking empirical response distributions to the memory representations they are supposed to reflect. Unsurprisingly, considerable effort has been expended to understand some of those potential obstacles, including variables whose magnitude and impact affect how one interprets
response distributions like those shown in the figure. Here, we will mention a few of those variables, focusing on ones most closely related to perception.

One general source of error is the mix of phenomena commonly lumped together as "absent mindedness." These phenomena occur when "insufficient attention is devoted to a stimulus at the time of encoding or retrieval or because attended information is processed superficially" (Schacter, 1999). Progress has been made in identifying neural correlates of attentional failure, particularly during encoding (Weissman, Roberts, Visscher, & Woldorff, 2006), which can undermine performance on virtually any task. Analysis of psychophysical studies often includes a related, but somewhat distinct variable, the "lapse rate." Lapses are errors that are unrelated to failures of stimulus detection or discrimination; lapses can occur because subject is momentarily inattentive, or attends to the wrong (that is, a non-cued) stimulus, or makes a mistake when executing the response (Arditi, 2006; Kingdom & Prins, 2016). In tasks like the one represented in Figure 13, lapses would translate into guesses, producing responses totally unconnected to memory for the cued-study item. Lapses (or mind wandering), in tasks like the one illustrated in Figure 13 could account for extremely deviant responses. A theoretical challenge is how to estimate the proportion of those lapse trials that are present in response distributions, and to distinguish them from trials whose responses are connected to memory, but only imperfectly. One approach might attempt to manipulate the likelihood of mind wandering by varying a feature of the experimental design such as the inter-trial interval.

Another challenge to interpreting results like those shown in Figure 14 arises from possible mis-bindings of color and location information (Rajsic, Swan, Wilson, & Pratt, 2017), a phenomenon related to the perceptual effect known as "illusory conjunctions" (Treisman & Schmidt, 1982). In experiments like those of Wilken and Ma (2004), mis-binding would promote correct recall of a color, but from some location other than the cued-location. A mis-binding would then be misinterpreted as a recall that was very wrong (deviant). Finally, a simple account of
color memory assumes that a recall of the color at the cued-location is unaffected by memory for the other colors that accompanied the color in the cued-location. If correct, that assumption would simplify interpretation of data like those in Figure 13. Recently, though, several groups have raised doubts about the validity of this convenient assumption (Brady & Alvarez, 2015; Hardman, Vergauwe, & Ricker, 2017). Without question, there is more work to be done, perhaps aided by existing software that evaluates different variables’ impact on formal models of the architecture and capacity of working-memory (Suchow, Brady, Fougnie, & Alvarez, 2013).

Summary

Throughout this chapter, we have argued that the wall that separates perception from memory has outlived its usefulness and needs to come down. Instead of building walls, researchers should build bridges, recognizing and taking advantage of the inter-relatedness of perception and memory. They have more in common than surface characteristics of the scientific approaches and histories might suggest. Perception and memory are intrinsically intertwined, like two strands of DNA, and each supports and shapes the other.

Although perception is often thought to be an automatic, bottom-up, stimulus-driven process, it is clear that perception is much more complex than that. Our sensory systems initiate perception, a process that often requires integration of stored (memory) information to disambiguate under-specified stimuli. Our perception also constrains memory in a number of ways, and cognitive researchers should be mindful of those constraints, and the importance of choosing the right stimulus to answer the question at hand. Among the key factors to consider are the overall perceptibility of a stimulus, the discriminability of and similarity among different stimuli, and the ways in which perceptual and verbal representations may interact. Notably, these factors are influenced by a subject’s past experiences and age, so an individual differences approach is ideal when feasible.
Stimulus manipulations that have been useful in elucidating perceptual processes also can be applied to advance our understanding of memory. For example, manipulation of stimulus noise and similarity spaces can be used to probe the representations of memoranda and to provide insights into the mechanisms underlying development, organization, and maintenance of those representations.
References


Liu, R., Patel, B. N., & Kwon, M. (2017). Age-related changes in crowding and reading speed. Scientific Reports, 7(1), 8271. doi: 10.1038/s41598-017-08652-0


