Chapter 1

Visual Memory Systems

Andrew Hollingworth
University of Iowa

Steven J. Luck
University of California, Davis

This volume is devoted to the study of visual memory. In this introductory chapter, we summarize the various visual memory systems that are discussed in this volume, focusing on the nature of the representations used by these systems, their temporal dynamics, their neutral substrates, and their functional role in visually guided behavior. First, however, we consider what it is that makes a memory a *visual* memory and why it is worth devoting a volume to the topic of visual memory.

1.1 DEFINING VISUAL MEMORY

1.1.1 What Makes a Memory a Visual Memory?

As a first pass, one could posit that a visual memory is any memory for which the stored information was acquired initially by the visual system. However, such a definition would be too broad. For example, in reading, although initial acquisition occurs through the visual system, memory for the semantic or syntactic content of the materials cannot qualify as a visual memory, because such representations are likely to be abstracted away from the specific visual properties of the printed text. To qualify as a visual memory, then, the memory must retain properties of the original perceptual states generated when the memory was encoded.

A second possibility is that visual memories are imagistic: a representation of a previously viewed stimulus that retains the topographic and metric properties of the original perceptual states. Such a view of visual memory has been popular in the literature on mental imagery, in which memory for previous perceptual states is proposed to give rise to quasi-visual experience in a manner that retains the metric properties of early vision. Although mental images certainly count as visual memories (and are the focus of Chapter 9), this definition is too narrow. First, visual memories that retain the precise metric properties of early vision are extraordinarily fleeting and could not form the substrate of longer-term visual memory. Second, higher-level perceptual systems discard precise metric information in favor of more abstract representations (see Chapter 6). Third, higher-level visual systems, such as visual short-term memory, clearly retain visual information but in a format that is not inherently imagistic (see Chapter 3).

Thus, visual memory encompasses memory representations that maintain information about the perceptual properties of viewed stimuli, but the format in which that information is encoded can vary from low-level imagistic representations generated in early visual areas to higher-level visual representations stored further along the ventral stream and abstracted away from the precise spatial structure of early vision.

In practice, researchers have isolated visual memory by careful choice of stimuli and by attempting to eliminate the coding of information in a nonvisual format. To this end, researchers have frequently used novel, complex stimuli that do not map onto existing object categories and would be difficult to code in a nonvisual format. For example, the classic study of Phillips (1974) used a set of checkerboard stimuli constructed by randomly filling half of the cells of an 8×8 matrix, which could not plausibly have been encoded verbally. In studies using common, familiar objects and easily nameable stimuli, such as color patches, a concurrent verbal memory load or articulatory suppression are typically used to minimize the possibility of encoding the visual stimuli verbally. Interestingly, a verbal memory load has little or no effect on visual memory performance, indicating that verbal and visual memory do not draw upon a common limited-capacity resource (Scarborough, 1972; Vogel, Woodman, & Luck, 2001).

1.1.2 The Independence of Visual Memories

We do not, however, mean to suggest that visual memories are entirely independent from more abstract, conceptual memories. In the context of long-term memory, researchers such as Barsalou (e.g., Barsalou, 1999) and Martin (e.g., Martin & Chao, 2001) have argued that abstract, conceptual representations are linked to sensorimotor representations, such that activating the concept of *lemon* may activate visual representations of the color of lemons, olfactory representations of the scent of lemons, and motor representations of the act of puckering in response to the sour taste of lemons. However, these links between abstract concepts and sensorimotor representations are fully consistent with the proposal that the modality-specific sensorimotor representations have properties specific to the individual modalities. Thus, visual memories can be studied as a related class of representations without denying that they are linked to a larger conceptual network.

1.1.3 Why Focus on Visual Memory?

It is possible to slice the memory pie in many different ways, so why does this volume pull out a slice defined by the visual modality? There are two main reasons for this. First, highly refined methods have been developed to assess visual functioning in general, leading to an enormous wealth of information about the visual system. The field of vision science therefore provides extremely firm footing for researchers who attempt to study visual memory, footing that is absent for most other varieties of memory. The methods and concepts of vision science also provide a large degree of unity to the different issues addressed in this volume.

A second reason to focus on visual memory is that excellent animal models are available to provide a link between human visual memory and the underlying neural substrates. In particular, visual memory has been studied fairly extensively in macaque monkeys using paradigms that are related to the memory paradigms used in humans, providing much more detailed information than can be obtained from human neuroimaging data. Research on monkey neurophysiology thus plays an important role in many chapters in this volume.

1.2 VISUAL MEMORY SYSTEMS

Researchers have subdivided visual memory into three main subsystems: visual sensory memory, visual short-term memory (VSTM),[1] and long-term memory (LTM). We will provide an overview of each subsystem and then consider how they are related.

1.2.1 Visual Sensory Memory

Visual sensory memory can easily be experienced by going into a dark room with a flash camera and taking a picture. Although the camera's flashbulb will provide only a few milliseconds of illumination, the perception of the illuminated room will fade over a period of about half a second. Scientists had noted this phenomenon for many decades, but it was initially just a curiosity based on introspection. In perhaps the single most influential study in the history of visual memory, Sperling (1960) applied partial report methods to the study of sensory memory, beginning a long series of studies that have documented the existence and properties of this variety of memory (see Chapter 2 for a detailed exploration).

Visual sensory memory—later named *iconic memory* by Neisser (1967)—was originally thought to be a simple, precategorical, spatially mapped, picture-like image, and this view still dominates many textbook descriptions. However, decades of research have subdivided visual sensory memory into two main categories of persisting visual information, called *visible persistence* and *informational persistence*. Visible persistence gives rise to the phenomenological experience of a fading visual image, and it reflects the persisting activity of photoreceptors and neurons in the early stages of the visual system (probably through area V1). This aspect of visual sensory memory corresponds to the

1 The terms *visual short-term memory* and *visual working memory* typically are used to refer to the same memory system.

typical view of iconic memory, in which a detailed, picture-like, but relatively unprocessed representation of the world fades over time.

Ironically, however, visible persistence does not play a central role in Sperling's partial-report method, which instead depends primarily on informational persistence. Informational persistence is, as the name implies, the persistence of information following the offset of the stimulus but, although this information can be used to perform various tasks, it does not give rise to any phenomenological experience of a fading visual image. It can be divided into two subcomponents. One component is a spatially organized and precategorical representation that likely reflects gradually decaying activity in intermediate stages of visual cortex (e.g., area V4). A second component is a more abstract, categorized, and amodal representation.

Visual sensory memory probably plays little or no role in complex higher-level aspects of human cognition. Because this type of memory is fleeting, it is unlikely to support the accumulation of visual information over timescales relevant to postperceptual processing. In addition, because visible persistence is masked by new perceptual processing, it cannot support the comparison or integration of perceptual information across longer perceptual episodes. However, a short-lived memory is important for early and intermediate stages of visual perception. In general, sensory memory can be described as a temporal smoothing of the input signal that can allow information to be extracted over more extended periods of time. For example, photoreceptors gain much greater sensitivity by integrating photons over a period of time rather than producing brief but tiny responses to each photon. Higher-level memory systems, such as VSTM, draw upon sensory memory when consolidating perceptual information into more stable forms of memory.

1.2.2 Visual Short-Term Memory

Visual short-term memory maintains visual information from a small number objects in a relatively abstract, object-based format. The capacity of VSTM is limited to three or four objects for simple stimuli and one or two objects for more complex stimuli. These object representations are significantly abstracted away from the precise metric structure of early vision. Whereas iconic memory is strongly disrupted by changes in the retinal position of stimuli, VSTM is highly robust over changes in absolute position (Phillips, 1974) and fairly robust over changes in relative position (Jiang, Olson, & Chun, 2000). Object representations in VSTM can be bound to spatial locations, but VSTM itself does not appear to be inherently spatial in the sense of an array-format image. This likely reflects the fact that VSTM draws upon visual representations in regions of the ventral stream with relatively large receptive fields that provide only a coarse coding of location.

More specifically, VSTM representations appear to be implemented by means of sustained neural firing in the inferotemporal cortex (in monkeys) and the lateral occipital complex (in humans), along with sustained firing in prefrontal cortex (see Chapter 8). These areas will produce a sensory response

to a stimulus and then maintain this response for several seconds after stimulus offset if required by the task. Earlier sensory areas, in contrast, appear to be unable to maintain high levels of activity following the termination of the stimulus.

The key functional feature of VSTM is its robustness across delay and across subsequent perceptual processing. Visual short-term memory representations can be maintained for many seconds and are largely impervious to visual masking. This allows VSTM to maintain information that spans perceptual episodes, bridging disruptions such as saccadic eye movements, blinks, and brief occlusion. This short-term bridging supports the comparison of perceptual information from objects separated in time and space, allowing people to compare visual objects to a remembered search target in the course of visual search, detect that an object has changed across a perceptual disruption, or note visual differences between two spatially separated objects. In addition, because consolidation of information into VSTM is strongly dependent on attention, the content of VSTM will reflect the moment-to-moment allocation of attention to goal-relevant objects during the performance of real-world tasks.

1.2.3 Visual Long-Term Memory

Visual short-term memory representations are robust over short periods of time, but the capacity limitations of VSTM preclude any significant accumulation of visual information over longer periods. In contrast, VLTM has a remarkably large storage capacity and highly robust retention. Single-trial learning of thousands of visual stimuli is possible, and such memory retains information about the specific visual form of objects and scenes. Visual long-term memory plays a central role in memory for the visual features of objects in the service of object and scene categorization (see Chapters 5 and 6). Visual long-term memory also is sensitive to statistical structure in visual input, which allows the visual system to utilize predictive contextual information in perception (see Chapter 7). For example, the learning of scene structure and object location facilitates visual search for an object, biasing search toward likely locations of the target. Finally, long-term environmental learning allows for the integration of visual and spatial information into large-scale, episodic representations of environments and events (see Chapter 5).

Whereas VSTM representations are maintained by means of sustained neural firing, VLTM representations are maintained by means of changes in the pattern and strength of the connections between neurons (see Chapter 8). These changes are thought to be implemented by means of structural changes in synaptic connections, which are responsible for the durability of VLTM representations. It has been notoriously difficult to find the actual storage locations of long-term memories in the brain, but most researchers believe that the memories are stored within the same systems that underlie perception (making it difficult to find cases of the loss of existing memories in the absence of perceptual deficits). Indeed, Chapter 6 argues that memory and perception are intrinsically interrelated.

1.2.4 Relations Among Visual Memory Subsystems

To what extent can these visual memory systems be considered independent? Each of the main visual memory systems—visual sensory memory, VSTM, and VLTM—has unique properties, especially in terms of capacity, duration, and abstractness. However, this does not mean that they are entirely independent. Sensory memory—especially insofar as it reflects an intrinsic aspect of perceptual representations—clearly provides information from which VSTM and VLTM representations are formed. Indeed, Chapter 3 argues that VSTM representations are best considered as high-level perceptual representations that have become stabilized by means of a limited-capacity attentional process. It is not yet known, however, whether these VSTM representations then become the starting point for VLTM representations. The influential working memory model of Baddeley and Hitch (1974) proposes that information need not pass through short-term memory to reach long-term memory, but this has not received much study in the visual domain. A few studies have asked whether VLTM representations influence VSTM storage, with most finding little or no interaction in this direction (see Chapter 3). However, we know of no direct tests of the hypothesis that VSTM representations provide the starting point for building VLTM representations. The alternative hypothesis is that VLTM representations are created directly from perceptual representations, whether or not they have been stabilized into VSTM representations. This will be an important question for future research.

REFERENCES

Baddeley, A. D., and Hitch, G. J. (1974). Working Memory. In *The Psychology of Learning and Motivation, Vol. VIII*, ed. G. H. Bower, 47–90. New York: Academic Press.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences* 22:577–660.

Jiang, Y., Olson, I. R., and Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory & Cognition* 2:683–702.

Martin, A., and Chao, L. L. (2001). Semantic memory and the brain: Structure and process. *Current Opinion in Neurobiology* 11:194–201.

Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century Crofts.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics* 16:283–290.

Scarborough, D. L. (1972). Memory for brief visual displays of symbols. *Cognitive Psychology* 3:408–429.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs* 74:(No. 498).

Vogel, E. K., Woodman, G. F., and Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance* 27:92–114.