CHAPTER

2

Sensational Memorability: Working Memory for Things We See, Hear, Feel, or Somehow Sense

Nelson Cowan

Department of Psychological Sciences, University of Missouri, McAlester Hall, Columbia, MO, USA

INTRODUCTION

I was pleased to be asked to visit a remote, beautiful, verdant, and earthy location to give the symposium lecture at a conference on something researchers rarely discuss as such: the sensory aspect of working memory. Is there such a thing? Why is it rarely referred to in this way? First I will discuss terms and what they mean, with a short modern history of the usage of these terms and a general theoretical framework to organize the data. Then I will discuss three categories of memory: (1) brief-duration sensory afterimages; (2) more persistent, processed sensory recollections; and (3) abstract units or chunks constructed across sensory modalities. In my theoretical framework, there is a great deal of commonality across sensory modalities despite the differences.

For someone like me with an interest in human experience, this sensory memory information is of obvious relevance. Sensory information ties together experience and memory. Imagine, for example, a child moving a flashlight in a circular motion in the dark. The trail of light that results appears to exist in the present, but it is actually a sensory record of the path that the light has taken in the previous second or so. Illustrating the general appeal of sensory processing to curious minds, it may have been the first aspect of psychology to capture my imagination, as shown in a selection of pages from a booklet that I assembled for a science assignment in grade school (Figure 1).

IS THERE A SENSORY WORKING MEMORY?

To address this new term, sensory working memory, one needs to consider in turn the definition of sensory memory and the definition of working memory.

Sensory Memory Definition

Sensory memory might be defined as any memory that preserves the characteristics of a particular sensory modality: the way an item looks, sounds, feels, and so on. Given that memories can be inaccurate, one might tend not to consider inaccurate information about sensory information to be sensory memory. For example, if one had a vivid memory of how an apple looked but incorrectly remembered the apple as red instead of green, one might consider that a constructed mental image and not a truly sensory memory. Yet, the constructed mental image is often useful; for example, reading a printed word may give rise to an acoustic image of how the word would be spoken. We know this because there are acoustic confusions among verbal items visually presented for recall (e.g., Conrad, 1964), but the acoustic imagery may be necessary for efficient covert verbal rehearsal.

There is evidence that items to be remembered that are presented in different modalities yield different patterns of neural response during a retention interval (e.g., Lefebvre et al., 2013), presumably with retention relying on at least

Mechanisms of Sensory Working Memory
http://dx.doi.org/10.1016/B978-0-12-801371-7.00002-8

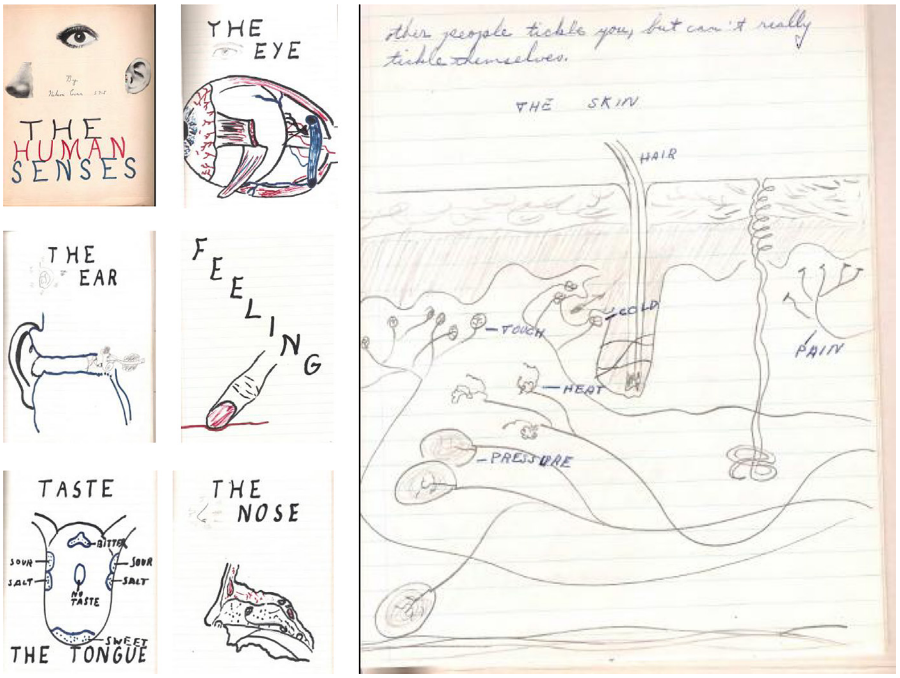

FIGURE 1 A selection of pages from a grade-school booklet by Nelson Cowan on the human senses. The page on human skin senses is enlarged to show the details.

some of the same regions in which the items were first processed, but there is also evidence that mental imagery yields patterns of neural activity similar to the imagined sensations if they had been actually perceived instead of imagined (Albers, Kok, Toni, Dijkerman, & de Lange, 2013; Farah, 1985). Mental imagery occurs not just in the case of visual stimuli, but also for other modalities (e.g., auditory stimuli such as tones). Crowder (1989) delivered a pure tone on each trial to be compared with a subsequent musical note in terms of pitch and found that the comparison was made more efficiently when participants were forewarned of the instrument in which the musical note would occur, which presumably allowed them to generate the appropriate timbre for the tone using auditory mental imagery.

It is not clear that it will always be possible to tell whether a memory is truly sensory in nature, given that mental imagery of sensations might sometimes get things right (e.g., correctly reconstructing through mental imagery that the apple was green). Given this theoretical difficulty, and given that I believe that memory of sensation over some seconds and mental imagery may play similar roles in processing, for the purpose of this chapter, I will consider both of them to be examples of sensory memory, which is thus broadly defined.

I should also add that I do not expect that there is an either/or situation as to whether a memory representation stemming from a stimulus is sensory or otherwise. Cowan (1988) introduced a rudimentary processing model in which there were several stages of memory. In the first, a brief afterimage of stimulation (in any modality) was said to last for several hundred milliseconds. Second, the afterimage impinging on the long-term memory system was said to give rise to activated features, which remain active for some seconds and include both sensory features such as color, angle, pitch, loudness, texture, smell, etc., depending on the nature of the stimulus, as well as more abstract or semantic features such as recognized phonemes, letters, shapes, or meanings (with many of these abstract or semantic features probably restricted to attended stimuli). The brain representation will depend on the constellation of activated features.

Sensory memory has traditionally been considered separate from working memory, in that sensory memory has a large capacity and a short duration, whereas working memory is more persistent but limited to a few items. For example, Sperling (1960) found that most characters in a briefly presented array were available for a short time when a partial report cue was presented (within a few hundred milliseconds), indicating the time limit of sensory memory, but he also found that the total number of characters that could be recalled on a trial was limited to about four regardless of the kind of report cue, indicating the item limit of working memory. According to Cowan (1988), Sperling (1960) was encountering the short phase of sensory memory, a literal sensory afterimage.

Capacity Limits for Sensory Memory?

It is likely that there are capacity limits for the second phase of sensory memory, defined as sets of sensory features that persist for some seconds. However, for several reasons, these capacity limits of sensory memory seem difficult to specify. It cannot be assumed that if a multifeatured item is presented, the observer's knowledge about the item is all or none. Luck and Vogel (1997) found that memory for one sensory feature implies memory for all perceived features, but this finding has been disconfirmed emphatically in subsequent work (Cowan, Blume, & Saults, 2013; Hardman & Cowan, in press; Oberauer & Eichenberger, 2013). These later articles show that if it is necessary to retain all features of each object in memory (e.g., color, orientation, length, and presence or absence of a gap), memory for any one feature is considerably impaired relative to trials in which only one feature must be retained.

Given this state of affairs, with the possibility of partial knowledge of an item, it is not clear how to measure the capacity limit in the longer phase of sensory information. Cowan (2001) suggested that capacity is limited to three to five items when each item is an integrated unit or chunk (Miller, 1956). When memory for one feature of an item is not accompanied by memory for other features of the same item, the item cannot be considered to be a single chunk; it may have to be remembered as multiple chunks (e.g., with orientation and color remembered separately), taking up more of the limited working memory capacity than would be the case if only one feature per object were needed. Also, there is the possibility of a sensory feature being remembered to some degree of specification, but not with enough precision to allow a correct response (e.g., memory for a spatial orientation of an item in an array to be remembered that is not specific enough to be distinguished from a probe item with which it is to be compared; see Bays & Husain, 2008; Zhang & Luck, 2008). So the characteristics of sensory memory, including its capacity, form a great area in need of further research.

Working Memory Definition

Working memory is a term that is used widely, but without much agreement between users. When I use the term, I try to do so in a general way, referring to any mechanisms of the mind that help to retain information temporarily. (For perhaps the first use of the term in this manner, see Newell & Simon, 1956, in their discussion of the use of computers for artificial intelligence.) Given that we cannot use the vast store of long-term memory information effectively all at once, but can only think about a small portion of it at any one time, there must be mechanisms to keep a small amount of information that is particularly task-relevant in a privileged state in which it can be accessed easily as needed for a short time. That information can come both from events recently experienced and from long-term memory of relevant past events and knowledge. Miller, Galanter, and Pribram (1960) used the term working memory in a way that does not seem too dissimilar from this, though they applied it specifically to memory for what one plans to do in the near future. Kane and Engle (2002) similarly have used the term working memory with specific reference to goal states to be maintained. In this sense of the term working memory, it is apparent that the activated sensory features of Cowan (1988) would count as part of working memory.

In contrast, in their seminal work, Baddeley and Hitch (1974) discussed working memory as a multicomponent system. By implication, the term included not only storage of information per se but also the manipulation of that information by central executive processes (an implication made more explicit by Baddeley, 1986). By definition, it appears that information coming directly from the senses cannot be manipulated or else it will no longer resemble what has come from the senses. Given that I have included sensory imagery as a kind of sensory memory, though, that kind of sensory memory could be manipulated. For example, Keller, Cowan, and Saults (1995) found that memory for the pitch of a tone was prolonged when the participant was free to rehearse the tone silently during the retention interval, rather than having to rehearse silently a distracting melody or a verbal sequence during that interval. The notion is that the participant was able to generate a mental image that matched the sensory memory and that the mental image could be maintained with attention, whereas the direct sensory memory of the tone otherwise would have decayed sooner.

Given my broad definition of "sensory," then, I agree that sensory working memory exists. Moreover, I would hasten to add that sensory working memory is a timely topic. It has been somewhat neglected by the standard body of work on working memory. That is probably because Baddeley and Hitch (1974) and others have been impressed with the finding that there are acoustic errors to printed verbal stimuli (e.g., Conrad, 1964), suggesting a nonmodality-specific form of memory coding. What has not received as much attention, however, is the well-researched finding that the recall of a verbal list is far superior at the end of the list (the recency effect) when that list is presented in a spoken as opposed to a printed form (auditory modality superiority effect: Cowan, Saults, & Brown, 2004; Murdock, 1968; Penney, 1989). Moreover, supporting the sensory memory basis of this effect, the auditory modality superiority is greatly reduced when the list is followed by a not-to-be-recalled final item, called a suffix, presented in the same voice as the list items (Crowder & Morton, 1969). The interference is less with a different-voice suffix and still less with a tone suffix. With a retention interval extended to 20 s, the voice-specific memory has faded, though the speech specificity of the memory remains (Balota & Duchek, 1986). Clearly, sensory memory plays an observable role.

I would see sensory memory, broadly defined here, as an integral part of working memory. Neurally, working memory may involve a limited number of attentional pointers from the parietal lobes that in themselves are abstract and not specific to a sensory modality, but that point to pools of neurons in sensory cortex and refer to them, integrating the sensory into the abstract (see, for example, Cowan, 2011; Harrison & Tong, 2009; Lewis-Peacock, Drysdale, Oberauer, & Postle, 2012; Serences, Ester, Vogel, & Awh, 2009).

Is Sensory Working Memory Veridical?

We have already discussed the fact that mental imagery of sensation is prone to being nonveridical, that is, different from the reality that it represents. Beyond that, it may be that all sensory memory is prone to error, just like other types of memory. Even the early stages of perception can be modified by attention (e.g., Woldorff et al., 1993), including modulation of the primary visual cortex (e.g., Pratte, Ling, Swisher, & Tong, 2013). Subjective properties as simple as the perceived duration of a stimulus can be affected by attention (Enns, Brehaut, & Shore, 1999). McCloskey and Watkins (1978) showed that when a line drawing was passed behind a narrow slit, what resulted was a perception of the entire object that was narrower than the actual object. Sensory or not, memory includes distortions that result from perceptual inferences; there is a little-explored topic of study called memory psychophysics, in which the distorting effects of memory on psychophysical functions are studied (Petrusic, Baranski, & Kennedy, 1998). Thus, if one distinguishes between the persistence of visible sensation and the persistence of information (Coltheart, 1980), one way in which they may differ is that the persisting sensation can incorporate some wrong information. For example, unlike sensory memory decay as a fading photograph, in Sperling's type of procedure, features from the array can survive but drift over time in memory relative to their initial positions (Mewhort, Campbell, Marchetti, & Campbell, 1981).

The terms *sense* and *sensation* are used in ways that are sometimes literal and sometimes metaphorical. The field of sensation is typically distinguished from perception on the grounds that sensations are not yet interpreted and categorized. One senses a red disk, perceives it to be a stoplight, and engages in further cognitive processing that leads to a decision to stop the car. This is the literal meaning of sense and sensation. More abstractly, instead of "in the literal meaning," I could have said, "in the literal sense," and that would reflect a usage of the word *sense* that is semantic rather than sensory. In *Sense and Sensibility*, a novel published "by a lady" in 1811 (in reality, Jane Austen), neither word is used in the sensory way; they mean, essentially, logic and emotion. The vocabulary usage demonstrates the special importance that people place on their sensory experiences, which they typically experience in a manner heavily infused with meaning and interpretation.

REVIEW OF SENSORY MEMORY RESEARCH

Having decided that sensory working memory broadly defined does exist, I will review the history of research on sensory memory because all of it seems important for the temporary maintenance of stimulation and is often neglected. Somewhere early in the history of cognitive psychology, sensory memory became divorced from other forms of information and excluded from what is typically called working memory. This can be traced to several influences. First, in a seminal investigation of temporary forms of memory, Sperling (1960) drew a clear distinction between a sensory memory that was time-limited but unlimited in capacity, and a short-term store that was limited to a small number of categorized items. When participants received an array of characters (3 rows × 4 columns in some critical experiments, for example) followed quickly by a tone cue indicating which row from the array was to be recalled, most of the row could be recalled. If the tone cue was delayed about a quarter of a second, performance declined to the point that only about four items from the entire array could be recalled (that is, just slightly more than one item per row in the 3 × 4 array mentioned). Without any cue, about four items could be recalled. This suggested that rich information about how the array looked was present, but that this information faded fast. Moreover, it could be used for recall only insofar as the information could be transferred into a short-term store limited to about four items. Subsequent experiments showed that for auditory stimuli, the sensory information lasted much longer yet still reached an asymptotic level of about four items when the sensory information faded (Darwin, Turvey, & Crowder, 1972; Rostron, 1974; Treisman & Rostron, 1972). Later I will present further analysis of this discrepancy between modalities, suggesting that there is actually a shorter and a longer store in all modalities.

In a second major influence on our understanding of sensory versus nonsensory forms of immediate memory, Conrad (1964) showed that our responses can be governed by something other than the sensory aspects of memory. He presented lists of letters to be recalled in the presented order and found that recall was impeded most by similarities, not in the way the letters looked (e.g., similarities in how *P* and *R* look), but in the way they would sound if pronounced (e.g., rhyming letter names for *B, C, D, E, G, P, T, V*, and *Z*). Based largely on this information, I think, Baddeley (1986) presented his model in a manner that tended to exclude sensory processes in favor of more processed phonological memory (regardless of the spoken or printed source of information) and visuo-spatial memory (which, despite its name, presumably could be formed even on the basis of the spatial qualities of nonvisual sensation).

This separation of sensory memory from the rest of working memory, however, leaves open the possibility that there are important commonalities between sensory and abstract types of working memory information. Cowan (1988) proposed that, regardless of sensory modality, there are two phases of sensory memory: a brief afterimage lasting several hundred milliseconds and a more processed phase lasting several seconds. Cowan further proposed that the latter phase consists of activated features from long-term memory, which can come from both past events and present sensory input. Activated features can include both sensory and categorical information.

Sensory memory of a very fleeting sort may be important in evolutionary terms because it allows the continued analysis of events that are very brief, such as flecks of moving color in the forest, brief growling sounds, or the first feeling of a spider crawling on your skin. A more persistent phase of sensory memory allows the careful comparisons of two stimuli, such as comparing your own trajectory when running with a rhinoceros coming in your direction, comparing the textures of two different stone surfaces to recognize your location in the dark, or learning a sound distinction in a new language. Sensory memory may be the glue that ties our experiences into a coherent stream of consciousness, allowing perception of the smooth motion of a deer as it runs between trees, the continuity of a talker's voice despite abrupt changes in the acoustics, or the feeling of your socks while you pull them up. This smoothness created out of discontinuity of the input is the basis of motion pictures.

Given these preliminary thoughts and definitions, I turn my attention to two questions that I will try to address. First, how much of working memory is sensory, that is, comprising the way things look, sound, feel, smell, etc.? The alternative is for much of working memory to be abstract, so that the modality of origin does not matter and sensory imagery does not dominate behavior. Second, how complicated is sensory memory? For instance, are there separate rules of operation for all modalities or are there common rules of operation across modalities?

What hinges on these questions is progress in scientific investigation of the mind. Finding regularities across modalities will allow us to establish more general principles of memory. Progress in understanding consciousness also depends on understanding sensory working memory. We somehow have to reconcile the vastness of the experienced world with the small limit of focal attention. The paradox can be seen, for example, in the phenomena of inattentional blindness (Bredemeiser & Simons, 2012; Simons & Rensink, 2005) or change blindness. Unattended aspects of the visual field, though clearly in plain view perceptually, can change without our noticing. If a flicker is introduced into the static showing of a photograph (which prevents transients from being used), one can make a dramatic change in the display without people perceiving it. For example, in a restaurant scene, a glass can alternate from being present to being absent, with most viewers needing quite some time to notice this alternation. The demonstrations and experiments of this sort highlight the difference between information that can be consciously perceived, as the entire scene in front of you can, in a holistic sense, and in contrast the selective entry of this information into working memory, which typically can occur for only a few separate objects in a scene at a time.

A Unifying Modeling Framework

The framework in which I will view sensory working memory is the one articulated by Cowan (1988), as shown in Figure 2. It is a framework rather than a well-worked out model, the intent being to summarize what we know and leave room for further details to be entered later, as we learn more.

Progression of Models

The model in Figure 2 can be put in context of the other models that preceded it. Stepping back, there are progressive changes in modeling conventions that are worth mentioning. (1) In the model of Sternberg (1966) and some other models of the era, phases of processing were depicted as if they occurred in a simple nonoverlapping sequence. Sternberg concluded that each item in a list is mentally scanned in turn. In many processing models of the era, first there was a large-capacity but brief-lived sensory store; then information was filtered by attention; then selected information entered a small-capacity short-term memory; and then that information was transferred to long-term memory for permanent storage. (2) Atkinson and Shiffrin (1968) kept that basic format but modified the modeling convention to focus on recurrent processes transferring information back from later stores to earlier ones, furthering a trend that was started in an offhand model sketched by Broadbent (1958). Presumably, information from long-term memory is used to enrich working memory, and information in short-term memory is recycled (refreshed) to

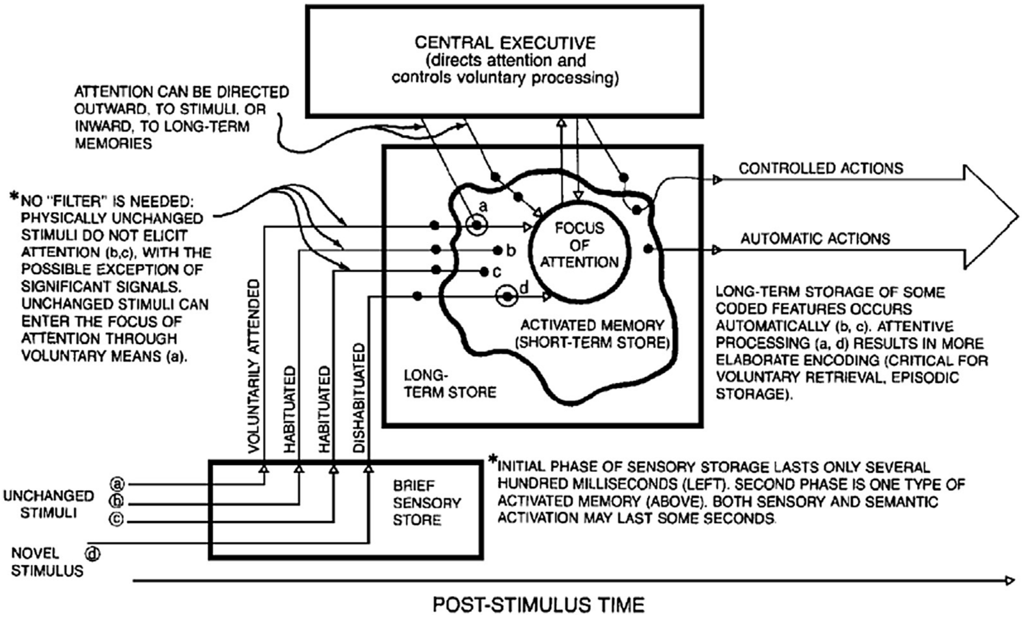

FIGURE 2 The modeling framework of Cowan (1988). *Reproduced from his Figure 1.*

counteract forgetting. (3) Baddeley and Hitch (1974), in developing the concept of working memory, changed the nature of transfer from one store to another. In their model, different aspects of information might be sent to the phonological store and the visuo-spatial store at the same time. The earlier notion of progression through the stores does not apply. Finally, (4) in the modeling framework of Cowan (1988), even the separation between components was abandoned, in ways that I will now explain.

The modeling framework of Cowan (1988), shown in Figure 2, leaves open the possibility that phonological and visuo-spatial memory operate separately but it leans toward the general principle that memory of any sort is susceptible to interference from subsequent items that are similar in their features. There are many forms of stimulation that clearly do not fit the phonological versus visuo-spatial distinction (e.g., spatial arrangements of tones, progressions of tactile or olfactory sensation). In the model, all forms are just represented as variants of activated sensory features from long-term memory. Moreover, until evidence can show otherwise, the assumption is that activated sensory features and semantic features of long-term memory may have similar properties, most notably decay over time and susceptibility to interference from subsequent input with similar features. Additionally, there is another important way in which components in this model are not separate. The limited-capacity storage system is conceived as the focus of attention, which is represented as a subset of the activated information from long-term memory. When information is no longer attended, the model states that it remains in an activated form for some seconds.

It is also the attentional filter of earlier models that is no longer separate in the model of Cowan (1988). In the modular way of thinking, it was unclear how information could sometimes get through the filter. Treisman (1964) suggested that there is an attenuating filter that only partly blocks unwanted stimulation. Instead, Cowan suggested that all information enters the memory system, whether it is attended or not. The information is used to form a neural model of the environment. That neural model includes sensory features of all sensory input streams but semantic features primarily of attended streams. When there is a change in the input that is perceived to differ from the neural model, either in perceived physical or perceived semantic characteristics, the result is a shift of attention accompanied by an orienting response (Sokolov, 1963). In the absence of an orienting response, the focus of attention remains on the task or sensory stream voluntarily selected. This model can account for why a change in the physical properties of an unattended aspect of the environment attracts attention. For example, the event can be a change from a low-pitched to a high-pitched voice in the background signifying that someone new just entered the room (see Cherry, 1953), or a sudden flash of light. There is no filter as such to exclude that information (at least, not when it results in the change in intensity of a feature; it may not apply when the abrupt change is in quality only, e.g., in color; see Folk, Remington, & Johnston, 1992). In short, the focus of attention is controlled partly by voluntary processes, but partly by orienting responses to changes in the environment.

Given evidence that there is not much semantic processing of truly unattended stimuli (Conway, Cowan, & Bunting, 2001; Cowan & Wood, 1997), Treisman's attenuation concept may not be needed. For example, Conway et al. showed

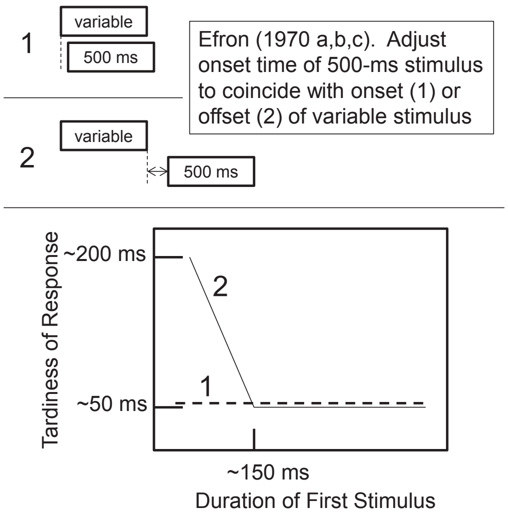

FIGURE 3 General method and schematic results of Efron (1970a, 1970b, 1970c). Number 1 reflects methods and results of onset estimation; number 2, of offset estimation.

that people noticing their names in an unattended channel in selective listening procedures (Moray, 1959) are primarily those with low working memory spans, whose attention may have wandered off of the assigned task to the channel with the name.

Within this framework, I will now provide more justification for the sensory and abstract aspects of working memory.

Phases of Working Memory Information, Both Sensory and Abstract

Cowan (1984) suggested that there are two forms of auditory sensory memory—a brief afterimage and a longer recollection of sensation—and Cowan (1988) extended that idea to all modalities (an idea presented also by Massaro, 1975). Why split something as elementary as sensory memory into two phases? The following research summary will explain why.

Brief Sensory Afterimages

There is abundant evidence of a kind of sensory afterimage. It is well known in the visual modality in which the chemical characteristics of the retina allow visual sensations to persist for some seconds, as in the case in which a candle is moved around in the dark. The kind of afterimage I mean, however, is not dependent on the retina, lasts only a fraction of a second, and seems to occur in every modality on which results have been reported.

One of the clearest indications of sensory afterimages was investigated in a remarkable series of experiments by Efron (1970a, 1970b, 1970c) on the minimal duration of a perception. His procedure and results are schematically illustrated in Figure 3. On every trial, two brief stimuli were presented. In some experiments, these were both visual or were both acoustic, whereas in other experiments there was one stimulus in each of these modalities. The stimulus being judged was of variable duration, whereas the other stimulus was of fixed duration. Sometimes, as shown in the top section of the figure, the task was to determine the temporal location of the fixed stimulus that would be perceived as synchronous with the onset of the variable stimulus. Other times, as shown in the second section of the figure, the task was to determine the temporal location of the fixed stimulus that would be perceived as synchronous with the offset of the variable stimulus. The difference between the two judgments was taken as the duration of the perception of the variable stimulus. The findings were similar regardless of the modalities of the stimuli: the point of offset of a very brief, variable stimulus was overestimated. The point of onset was only slightly overestimated and, putting together the two kinds of information (as shown schematically at the bottom of Figure 3), the minimum duration of a perception was about 200 ms. Therefore, the shorter the stimulus, the more its duration was overestimated. (A seasoned investigator may look at this schematic result and assume that the true result was much more variable, but it was actually surprisingly clean.)

The results of Efron (1970a, 1970b, 1970c) can be explained on the basis of a sensory afterimage of the variable stimulus. This sensory afterimage might occur as a side effect of the initial sensory processing of the stimulus. As shown at the top of Figure 4, in the critical condition, the indicator is supposed to be adjusted to the end of the variable-length stimulus but, unknown to the participant, it is actually adjusted to some point after the sensory memory of the variable stimulus has faded to a certain level at which it is no longer experienced as ongoing sensation.

FIGURE 4 Four consequences of following a brief stimulus with a second stimulus, all of which I interpret as the consequence of a brief sensory afterimage. The box figures represent stimuli, the curved and striped figures represent the sensory afterimage, and the right side of the figure shows the typical result for the second through fourth procedures. For the results of the first procedure, see Figure 3.

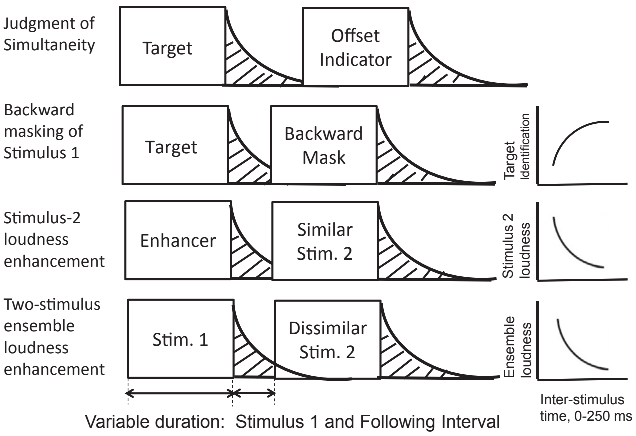

FIGURE 5 The results of Craig (1985, Figure 4) for forward and backward masking of vibrotactile sensation.

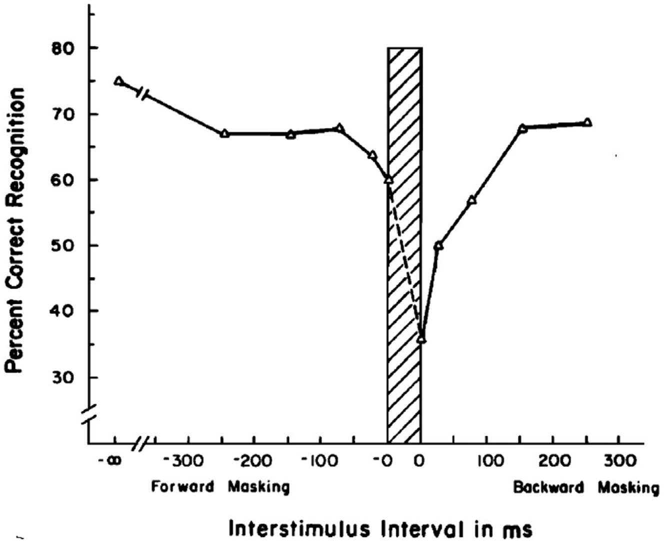

Figure 4 also shows other ways in which two-stimulus procedures are relevant to the notion that there is a brief sensory afterimage lasting several hundred milliseconds. In one type of procedure, recognition masking, two brief stimuli are presented in rapid succession and either the first stimulus is to be identified (backward recognition masking, shown in the figure) or the second stimulus is to be identified (forward recognition masking, not shown). The result is basically the same in vision (Turvey, 1973), audition (Massaro, 1975), or tactile sensation (Craig, 1985). To illustrate the finding, Figure 5 shows it for tactile sensation. Craig used pairs of very brief vibrotactile stimuli, each consisting of a pattern of pins that were thrust forward to stimulate a fingertip for 4.3 ms. The recognition of the second pattern was only minimally affected by presentation of the first pattern, though there was some effect when the patterns were within 50 ms of one another. In contrast, recognition of the first pattern was profoundly impaired by the second pattern (the backward masking effect) and this effect did not fully dissipate until the patterns were about 200 ms apart. The interpretation is that each pattern is followed by a sensory afterimage that contributes to the recognition process, and the second pattern interferes with recognition of the first if it interrupts the sensory afterimage of the first before that afterimage can be processed. Subsequent research has shown that the critical variable for the backward masking effect is the stimulus onset asynchrony, because for sufficiently long stimuli the recognition process can be completed while the stimulus is ongoing.

Cowan (1987) took the masking process a step further by demonstrating how the sensory afterimage might contribute to the process of loudness perception. On each trial in his Experiment 1, a pair of tones was presented, with a variable intertone interval. The loudness of each tone in the pair was to be judged on an eight-point scale. The tones came in three different brief durations (20, 30, and 40 ms), based on the fact that sensory integration results in louder judgments for longer tones. The expectation was that the sensory afterimage of each tone would contribute to that overall integration and that the second tone terminates the sensory afterimage of the first tone, limiting the neural

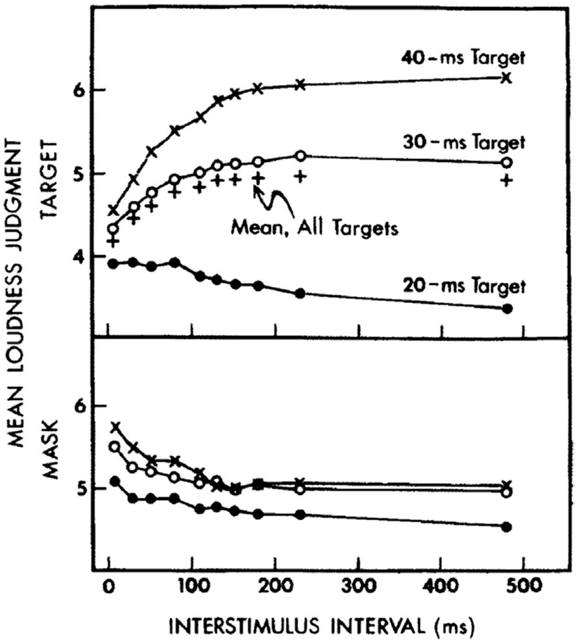

FIGURE 6 Masking of loudness, from Cowan (1987, Figure 2). The top panel represents the perception of the first tone in a pair; the bottom panel represents the perception of the second tone in a pair.

activity contributing to the perceived loudness of the first tone. If these assumptions are correct, the loudness of the first tone in a pair should increase as the interval between the two tones increases. Additionally, the scale should be used in such a way that the judgment of loudness for the three tone durations would diverge across intertone intervals, because each tone's loudness should be more clearly perceived at longer masking intervals. Those two principles together accounted well for the results. Figure 6 shows the results for the first tone in a pair (top panel) and for the second tone (bottom panel). For the first tone, the judgment of the 40-ms tone increased markedly as a function of the intertone interval, the judgment of the 30-ms tone increased more moderately, and the judgment of the 20-ms tone actually decreased slightly. The results were well-fit by a model that basically includes two factors as mentioned: (1) the integration of sensation across the stimulus and its sensory memory, up to the point at which another tone occurs to interrupt that sensory memory; and (2) a weighting function that produces more separation between the 20-, 30-, and 40-ms sounds as the interstimulus interval increases. For judgment of the second tone (or mask, as shown in the bottom panel of the figure), there was no backward masking, only forward masking, so responses were affected only at very short intervals, and not differentially for the three tone lengths. The results verify that the sensory afterimage adds to the sensation that is caused directly while the sound is ongoing.

Figure 4 also shows two other ways in which the sensory afterimage can be demonstrated using intensity perception (e.g., for tones, Arieh & Marks, 2003; for vibrotactile stimuli, Verrillo & Gescheider, 1975). When the two stimuli are very similar and close together, the sensory afterimage of the first stimulus can inappropriately contribute to the perceived intensity of the second stimulus, as shown in the third row of the figure. When the two stimuli are very different but close together, this effect on the judgment of the second stimulus does not occur. In that case, though, if participants are asked to judge the overall intensity of the two-stimulus complex, the judgment is higher when they are close together in time. Presumably, this phenomenon occurs because there are registers of stimulation (critical bands), and the overall perceived intensity is greater when more than one critical band is stimulated concurrently. I suggest that it is the concurrence of the sensory afterimage of the first stimulus with the reception of the second stimulus that causes this perceived intensity enhancement.

In sum, there is a brief phase of sensory memory lasting several hundred milliseconds from the onset of a stimulus. It appears to operate similarly across modalities, is tied to the process of stimulus recognition, and is experienced as the continuation of sensation in the modality of origin.

Processed Sensory Recollection

It seems to me that the field of sensory memory has always been the source of much confusion, largely because different procedures yield very different estimates of sensory memory. Whereas Sperling (1960) showed that there was a type of visual sensory memory lasting several hundred milliseconds, a much longer estimate of acoustic memory was obtained in a procedure similar to Sperling but using spatiotemporal arrays of spoken digits (Darwin et al., 1972).

Characters were spoken at left, center, and right locations concurrently in two successive temporal bursts for a total of six characters to be remembered per trial. A cue indicated which spatial location to recall. The finding was that sensory information took a lot longer to decline to a plateau than was found in vision: about 4 s. On the basis of this finding, it has often been argued that sensory information lasts a lot longer in the acoustic domain than in vision.

This interpretation of Darwin et al. (1972) as indicating a modality difference, however, would leave no explanation for the comparable results obtained for visual and acoustic stimuli by Efron (1970a, 1970b, 1970c), so there is a conundrum. The solution may be as follows. Results obtained by Massaro (1976) following up on Darwin et al. suggest that their sort of finding in the auditory domain is not comparable to what was obtained in the visual domain by Sperling. In vision, a postarray category cue (e.g., recall all letters but not numbers in the array) cannot be used to access a part of the sensory representation as efficiently as a postarray physical cue (e.g., recall the middle row). Yet, Massaro found that physical and category postarray cues worked equally well in audition. Darwin et al. carried out a comparable experiment with semantic cueing (their Experiment III) but it is difficult to compare with physical cueing because the whole-report results in these two experiments differ substantially for some reason. Overall, I conclude, with Massaro, that the ability to use the longer memory for acoustic stimuli may make this technique incapable of demonstrating the shorter memory of the type examined by Efron and others, and that there are two phases of sensory memory in each modality, with different time courses and characteristics (Cowan, 1984, 1988). It still remains to be discussed why the longer store does not similarly interfere with the findings in the visual modality.

Evidence of Two Phases of Sensory Memory

Although Sperling (1960) presented clean results on sensory memory, a further detail of the phenomenon leads to a different interpretation. Specifically, when one looks at the types of errors that occur with increasing cue delays, they turn out to be errors in the spatial location of items (Averbach & Coriell, 1961; Mewhort et al., 1981). This implies that information about the items continues to exist in memory, even in the visual modality, but not the same high-fidelity information that exists for a brief period of several hundred milliseconds and includes spatial location information.

Another well-known finding indicating two phases of visual storage, for a similar reason, was obtained by Phillips (1974). He presented on every trial a black-and-white grid with squares randomly assigned to black or white. After a variable retention interval, it was presented a second time, either intact or with one changed square color. The task was to indicate if the array was the same or changed. On some trials, however, the second array was also spatially shifted compared with the first array. The finding of interest here is the effect of the spatial shift. If the retention interval was about 300 ms or less, a spatial shift substantially impeded performance. If, however, the retention interval was longer (it could last up to about 600 ms), the spatial shift did not matter. This suggested that there is a literal representation that depends on an exact spatial match and lasts only about 300 ms, whereas there is a less literal representation that is more abstract but still contains information about many of the display elements. When the array contained as many as 64 squares, performance was still about 60% correct at the longest retention interval. The location-dependent and location-independent stores appear to be the same as the two phases of sensory memory that I have mentioned.

A different procedure has been used to suggest the existence of two stores in the visual modality. Kallman and Massaro (1979) presented three stimuli in the order *comparison-target-mask* or in an alternative order, *target-mask-comparison*. In either case, the target and comparison tones were to be compared. The time between the comparison and target tones was fixed but the time between the target and mask was varied. In the order *comparison-target-mask*, the masking interval could influence the recognition of the target by limiting access to its sensory afterimage. In the alternative order, *target-mask-comparison*, this was still the case but here there was an additional, second basis of interference. Specifically, regardless of the target-mask interval, the mask intervened between the target and comparison tones and could interfere with the longer phase of sensory memory of the target tone. Sure enough, the *comparison-target-mask* order showed an effect of the mask only at short masking intervals, whereas the *target-mask-comparison* order showed an effect of the mask regardless of the masking interval (because of the interference with both phases of sensory storage). The interference effects were seen as effects of similarity between the target and mask. This study provides clear evidence for the existence of two separate phases of sensory memory.

Although I cannot find any tactile study that similarly demonstrates the existence of two phases of sensory memory, there is evidence in the form of different procedures with different results. Recall that vibrotactile sensation yields evidence of recognition masking suggesting a store lasting several hundred milliseconds while the stimulus is identified. Other evidence, though, shows that the memory for a tactile stimulus is diminished across some seconds (e.g., across about 5 s for reproducing the location of the tactile event; Sullivan & Turvey, 1972).

Nature of the Longer Store

According to the model of Cowan (1988) shown in Figure 2, the second phase of sensory memory should be an activated portion of long-term memory. This point was demonstrated by Cowan, Winkler, Teder, and Näätänen (1993) using the mismatch negativity (MMN) response of event-related potentials. In the MMN, the participant is engaged in a visual task (such as reading a book) while ignoring acoustic stimuli. When a train of identical standard tones (e.g., 1 per second) is followed by a deviant tone, an MMN results. It is thought to result from the comparison of the standard representation to the deviant. However, this MMN is not produced under certain circumstances. If there is only one standard, there is no MMN to a change. If there are many standards but there is a long delay of over 10 s before the deviant is presented, it does not elicit an MMN. These results have been taken to suggest that the sensory memory representation of the standard must be built up across tones, and that this representation is lost by about 10 s. Cowan et al. reproduced this result but, in some conditions, after the long delay there was one reproduction of the standard tone, and then the deviant. Under these conditions, the MMN returned. The explanation was that the long phase of sensory memory is an activated set of elements from long-term memory, that the activation of the standard tone representation can be lost over the delay, and that a single reminder of the standard tone can reactivate this representation.

May and Tiitinen (2010) offered an alternative, non-memory basis of the MMN. In Cowan et al. (1993), they suggested that the neural populations that respond to the standard and deviant just following a long delay are in similar, unhabituated states, and therefore show no MMN. This interpretation appears to overlook the results of a roving standard control condition in which a standard train followed by a long delay had to be followed by not only one but several iterations of the standard before an MMN could occur. Perhaps the non-memory theory could be stretched to take that predelay basis of habituation into account. There are, however, subsequent data that would be even more challenging for a non-memory interpretation of the MMN. Winkler, Schröger, and Cowan (2001) found that the "long delay" that mattered was relative and not absolute. When standard tones presented in a fast train were followed by a 7-s delay, about half of the participants showed no MMN to a postdelay deviant. However, when standard tones separated by 7-s intervals were followed by a 7-s delay, those same participants did show an MMN to a postdelay deviant. Winkler et al. suggested that the processing system gauged the contextual relevance of the standards; a train of closely packed standards lost relevance after a long delay, but a train of separated standards did not because the context defined by the stimulus pace had not changed. It is difficult to see how the non-memory interpretation could account for those data.

The role of attention is probably different for the two phases of sensory memory (Winkler & Cowan, 2005). For the brief afterimage, Sperling (1960) showed that many items can be stored at once, though this information is short-lived except for attended items. The second phase of sensory memory lasts several seconds without attention, though it may be enhanced by attention. For example, Keller et al. (1995) found that same–different tone comparisons could be enhanced modestly by rehearsal across a retention interval of several seconds unless a distracting task was presented during that interval. This enhancement occurred even though the stimuli could not be easily categorized; tone differences still were less than a semitone.

In another investigation of the role of attention in the second phase of sensory memory, Cowan, Lichty, and Grove (1990) tested memory for syllables that were unattended during their presentation while the participant was busy reading. Memory for most syllables was not tested. After 1, 5, or 10 s of silent reading following the presentation of occasional target syllables, a light cue indicated that the participant should put down the reading and recall the most recent syllable. There was a marked decay rate of memory across this 10-s interval, with declining performance as a function of the retention interval. However, even a small amount of attention to the syllables abolished this decay function. Specifically, in one experiment, there was an added task to listen for instances of the syllable /dih/. This syllable was detected only 60% of the time, but this split attention was enough to abolish completely the decay of sensory information for the various syllables, which now held steady across 10 s. It seems likely that attention transfers information into a more abstract form, which we will discuss next.

It also might be, however, that the short phase of sensory persistence is influenced by attention as broadly defined. In Sperling's study, attention is distributed to the locations in space that contain all of the characters in the stimulus array, even though it may not be possible for the array items to be attended individually; there are too many of them. In other work, though, it has been shown that there is an effect of attention to versus away from part of a spatial field on the reported duration of a visual stimulus in that part of the field, a potential indication that attention extends the first phase of sensory memory (Enns et al., 1999). It is not clear, however, whether the measure of perceived duration that Enns et al. used provides insight into the same level of processing that Efron (1970a, 1970b, 1970c) provided by having participants compare the offsets of two stimuli. Enns et al. appear to have found an effect of attention from a 100-ms cue period on perception of a very short target, suggesting that the effect could result at an interpretive stage after the sensory afterimage has faded, rather than prolonging the sensory afterimage itself. This is fertile ground for further work.

In sum, the second phase of sensory memory is not literal; it is not an afterimage that continues the stimulus identification process, as the first phase is, but rather a vivid recollection of the sensation that can be used to compare stimuli. Examples are learning to distinguish between two slightly different words in a foreign language, and comparing two slightly different shades of paint (or two arrays, for that matter).

Constructed Abstract Units

It is abundantly clear that, in addition to sensory codes, there are abstract codes that do not directly depend on sensation after the stimuli have been encoded. For example, as mentioned previously, Conrad (1964) showed that memory for printed speech is still susceptible to acoustic confusions, suggesting that an acoustic-like code is internally generated. Even if this acoustic-like code is considered sensory, a more abstract phonemic category is also likely to be generated. Thus, words such as *metal* and *medal* can be acoustically identical and yet assigned to different categories. In other contexts, the categorical difference is not neutralized but heard (*metallic* vs. *medallion*).

One might suspect that all sensory codes are abstract, so that there would be no difference between the memory for acoustically presented speech and visually presented stimuli that are mentally converted to speech. That actually appears to be the expectation of Broadbent (1958), whose groundbreaking information-processing diagram showed a feedback arrow from working memory to what we now call sensory memory. An argument against this idea of equivalence, however, is that memory performance in the two cases is much different. In immediate free recall of lists of verbal items, memory for items near the end of the list is far better with acoustic presentation than with printed presentation (Murdock, 1968; Penney, 1989).

One commonality between the two phases of sensory memory is that they presumably each could include many elements at the same time (e.g., Phillips, 1974). In contrast to this, many believe that abstract information is limited to several distinct items at once. Cowan (2001) argued for this hypothesis. To identify working memory situations in which abstract units had to be relied upon, Cowan looked for procedures in which nothing else could be used: sensory memory could not be used (e.g., because of masking or a change in sensory qualities from stimulus to test), verbal rehearsal was prevented (e.g., by articulatory suppression), and items could not easily be combined to form larger chunks (e.g., because the presentation was too rapid). Under these circumstances, it was assumed that memory was based on abstract, nonsensory units. Normal adults could remember typically only three to five abstract units, with a somewhat smaller number in young children.

The term “abstract” is somewhat difficult to pin down, and the present section is an attempt to do so. The first part of the discussion covers a controversy regarding the capacity limit. One traditional distinction between sensory and abstract temporary memory has been that there is a capacity limit for abstract memory defined in terms of the number of chunks (after Miller, 1956) but no capacity limit for sensory memory; rather, a time limit (after Sperling, 1960). That distinction is challenged by a modern controversy in which some investigators assert that attention can be divided among any number of items in a stimulus field, with no mention of any abstract form of storage with a capacity limit.

After this controversy about capacity is discussed, the second part of the section asks how we might know that information is abstract. When verbal labeling of the information is possible it is easiest to know, but I argue that some information that cannot easily be labeled is still likely to be abstract as opposed to sensory. As mentioned previously, one reason why the distinction is difficult may be that we always have a sensory underpinning with any type of abstract processing. The two sections fit together inasmuch as it is possible that the belief in a capacity limited to several chunks is correct for abstract memory but that investigators who believe in a working memory based on attention spread thinly across the entire stimulus field are picking up on sensory aspects of memory.

Characteristics of Abstract Working Memory 1: Capacity Limits of Abstract Memory

Although the basic finding is clear, the interpretation is still an important issue. Much of the discussion is based on a procedure by Luck and Vogel (1997) in which a spatial array of objects is followed by a short retention interval and then a probe that is to be judged the same as in the array or different from it. Performance in adults is excellent with three items and declines markedly as the number of items increases. Simple formulas can be used to estimate the number of items held in working memory. Underlying them is the assumption that the participant answers correctly if the critical information is in working memory, and otherwise guesses with a certain bias. That bias depends on the test situation (see Cowan, 2001; Pashler, 1988; for an overall explanation of the formulas, see Rouder, Morey, Morey, & Cowan, 2011). The formulas show that on the average, between two and five items are held in working memory, and the average is typically close to three. Zhang and Luck (2008) modified the procedure to ask whether attention could actually be spread across an entire field. Participants received an array of items that could vary along a continuous dimension (color or orientation) and the answer was given by adjusting a wheel (color or orientation) to the correct value. The angle of disparity between the actual and recalled values was the imprecision of the answer. The findings along with a memory-plus-guessing model suggested that only a few items could be held in memory and that other responses were based on guessing.

Other recent evidence in favor of capacity limits comes from an investigation of reaction times and a sophisticated process model. Donkin, Nosofsky, Gold, and Shiffrin (2013) found that reaction times in change-detection tasks are as one would expect from a capacity-limited process: on some trials the participant knows the item being probed and responds quickly, whereas on other trials the participant must guess and responds more slowly.

In the other camp, Bays and Husain (2008) argued that working memory is a resource that actually can be spread across all items in a visual field. Anderson, Vogel, and Awh (2011) provided some strong evidence against that view, showing that imprecision increases when the number of stimuli increase from one to three, but then reaches a stable plateau, presumably because capacity has been filled. Zhang and Luck (2011) showed that participants could not be motivated to spread attention more thinly to more objects in the field. Nevertheless, recently, van den Berg, Shin, Chou, George, and Ma (2012) revived the hypothesis of working memory as a continuous resource, this time in a complex model in which the precision of the representation of the tested item can vary from trial to trial. van den Berg, Awh, and Ma (2014) took this line of argument further and tested a great number of models, finding that the winning model had variable capacity as well as variable precision. Nevertheless, there was an issue about some of the models being more flexible than others (i.e., more able to win the comparison even when fit to data generated by a process that does not match the model) so further work is needed. The jury is still out or, rather, is yet to be selected.

Although I favor the limited-capacity view, it is also the case that this view must be modified compared to where it was when defined by Luck and Vogel (1997). They carried out studies with objects that had up to four important features (color, orientation, length, and presence or absence of a gap) and concluded that if one had an object in working memory, it included all of the features. They found that people could remember three or four objects with all the features. In one sense, there could be nothing abstract about such a memory system because any item taking a slot in working memory would have to have attached to it all relevant sensory information. However, Hardman and Cowan (in press) investigated further with very similar materials and did not get the same results; the need to pay attention to all features in several objects consistently reduced the likelihood that any one feature of an object would be retained.

There are other findings concordant with Hardman and Cowan (in press), but note that none of them contradict the notion that there is a limit in the number of chunks that can cooccur in working memory. Cowan et al. (2013) found that attention to both color and shape resulted in fewer colors or shapes being available than if attention were directed only to color or only to shape. Oberauer and Eichenberger (2013) found that attention could be divided among even more than two features of an object, with tradeoffs between features. Both data sets, however, were fit to a model in which objects do play a role. In this model, the observer with capacity *k* apprehends *k* objects in one of the features (e.g., *k* colors) and then is able to encode other features, but only for these particular objects for which the first feature has already been encoded and, even then, often only for a subset of them. Therefore, there will be *k* objects with at least one feature encoded, though there also is a limit for the number of features that can be encoded concurrently. Hardman and Cowan (in press) also found that they could not rule out such a model in which there are two phases of working memory limits for complex objects: a phase in which features compete for attention, and a phase in which the incomplete objects in working memory are limited to about three objects.

Cowan (2001) summarized a great deal of other evidence suggesting that a capacity limit also applies in the case of the recognition or recall of verbal lists (further explored by Chen & Cowan, 2009; Cowan, Rouder, Blume, & Saults, 2012). There is additional evidence suggesting the possibility of similar capacity limits for tactile stimuli (e.g., Bliss, Crane, Mansfield, & Townsend, 1966) and even for odors (e.g., Laing & Francis, 1989).

Characteristics of Abstract Working Memory 2: The Distinction from Sensory Memory

For our understanding of human information processing it seems important to ask how much of what is considered abstract working memory is in fact sensory in nature. Abstract information would function in the same way regardless of the sensory modality of origin of the information. Abstract representations in working memory share neural areas even when the information comes from different modalities and codes (Cowan et al., 2011). Note though that there are sensory-specific areas of activation in these data too; it is just that a memory load area that cuts across sensory modalities can be found in the parietal lobes, the intraparietal sulcus (IPS).

One indication of abstractness comes in the nature of interference. Whereas memory for a tone's pitch sustains severe interference from additional presented tones, more than from other sounds (Deutsch, 1970), memory for a tone's pitch is indifferent to whether silent mental interference comes from an imagined digit sequence or an imagined tone sequence (Keller et al., 1995). Both interfere to the same degree relative to a no-interference condition in which imagery-based rehearsal is possible. This suggests that sensory input interferes with sensory memory but that abstract input from imagery may not interfere with sensory memory.

Under some situations, what Cowan (1984, 1988) called the second phase of sensory memory might be used to supplement abstract memory (e.g., see Sligte, Scholtel, & Lamme, 2008), but might not share some of the same capacity limits; the information could be spread over all array items. To investigate this issue, it is possible to impose a masking pattern between the array to be encoded into working memory and the test (cf. Saults & Cowan, 2007; Rouder et al., 2008). The masking pattern should be placed late enough to allow sufficient encoding of items into working memory in the first place; see Vogel, Woodman, and Luck (2006). Most of the studies in this field have not included a mask, so the data may in some ways suggest a continuous resource to some investigators because of the contribution of sensory memory, which makes the study inconclusive as to the properties of a more abstract working memory.

Last, one might wonder if the items that are abstract must be those that have verbal labels and therefore might be retained through verbal coding. This is a difficult issue because in the absence of an available verbal label, it is difficult to know whether the participant truly has a category for a stimulus or whether the stimulus might be perceived as an ensemble of parts, not integrated chunks or members of categories. Olsson and Poom (2005) argued that categories were needed, on the grounds that when their visual stimuli were difficult to categorize, participants could retain only about one such item.

Li, Cowan, and Saults (2013) addressed this question of categorical, yet nonverbal, working memory by developing a set of 12 tones that had musical timbres, which were generally too close together to be labeled as members of particular instruments; participants generally reported not labeling them. The frequencies of the tones did not fall on a musical scale, either. The presence of timbres presumably allowed categorization of the 12 sounds following repeated exposure to the sounds. In experiments using tones with and without the timbres, a sequence of tones to be remembered was followed by a mask to limit sensory memory. The presence of timbres increased the number of tones that could be remembered, with the more capable participants able to remember about three of them. Thus, there is some evidence suggesting that the use of abstract memory requires categorization even when the stimuli cannot be labeled, though further work is needed.

Brain studies could help to clarify what information is abstract and what information is sensory. On one hand, function magnetic resonance imaging studies with auditory verbal and visual spatial objects show that there are brain areas (most specifically the IPS) that respond to a working memory load regardless of its modality (Chein, Moore, & Conway, 2011; Cowan et al., 2011; Majerus et al., 2010). This leads to the notion that the stimuli of both types give rise to a very general neural representation of abstract information about the stimuli. On the other hand, there are certainly sense-specific areas of memory activation, even in these same studies. Lefebvre et al. (2013) emphasized very different representations for visual versus nonverbal auditory short-term memory (nonmusical tones) in electrophysiological studies. Looking across all of the research, it seems possible that the brain's parietal representation requires abstract categories, whereas the representation of sensory information that is not categorized requires more sustained attention, activating frontal areas. More work is needed crossing stimulus methods with imaging methods before we will have a full picture.

Another way to show nonverbal, yet categorical memory is to require categorical, verbal responses even for nonverbal items and to show that the representations remain distinctly nonverbal despite the verbal responses. Rowe and colleagues took this approach. Rowe, Philipchalk, and Cake (1974) presented identifiable sound effects or their labels and interfered with the memory with either poetry or music. The poetry interference hurt label memory more whereas the music interference hurt sound effect memory more, indicating two different sources of memory underlying the verbal response. (See also Rowe & Rowe, 1976.)

One could imagine a similar study of working memory for complex visual stimuli that become familiar over trials, yet cannot be named; I cannot think of such work. However, given that short-term memory for visual objects has similar properties in people and pigeons (Gibson, Wasserman, & Luck, 2011), it must not be completely dependent on verbal labeling. Further showing that verbal labeling cannot always account for capacity-limited performance, human studies of memory for simple objects in arrays show little or no effect of a rehearsal suppression task (e.g., Morey & Cowan, 2004).

Tradeoffs between Sensory Modalities

A key difference between sensory and abstract types of working memory is that abstract memory is supposedly based on a limited resource that is shared in common among modalities. Note that this is not necessarily true of sensory and abstract *long-term* memories. There are long-term memories that seem sensory in nature, such as one's memory for a person's voice that sometimes allows recognition over a telephone. One can remember the voice, however, presumably without impeding long-term memory for the way the person looks. In working memory, though, there is likely to be such a tradeoff across modalities. Remembering four exact colors should interfere with remembering four exact tones. There is some evidence in favor of that hypothesis (e.g., Cowan & Morey, 2007; Morey & Bieler, 2013; Morey, Cowan, Morey, & Rouder, 2011; Saults & Cowan, 2007). Still, it is a hypothesis that is in dispute; in his talk at the conference, Marois suggested that the tradeoff between modalities would disappear if all types of structural similarity between modalities were removed (e.g., a tone series versus a spatial layout of visual locations). The issue remains under debate.

Functional magnetic resonance imaging may help. Todd and Marois (2004) showed that there is a brain area that responds more strongly to a larger visual memory load, but only up to the capacity limit measured behaviorally: specifically, the IPS, bilaterally. Cowan et al. (2011) looked for areas across the brain that responded to memory loads consisting of arrays of colored squares, series of spoken letters, or combinations of both. Only one area was found that robustly responded regardless of modality: the left IPS (see Chein et al., 2011; Majerus et al., 2010). This area also responds heavily to attention demands even when working memory is not required (for a review, see Cowan, 2011). These data taken together may suggest that attention is used to form and hold abstract representations of items from different modalities, and that there is a limit in how many abstract representations can be held in this way. More work is needed to verify the prediction that this resource has a fixed limit of several abstract items regardless of the modality of origin, provided that sensory memory is disallowed.

Last, there are some complementary types of evidence suggesting abstract memory for information that may trade off when other modalities are combined (e.g., visual and tactile information: Gallace, Tan, & Spence, 2007).

CONCLUDING OBSERVATIONS

In sum, I have made the argument that there are striking similarities in the operation of a short-lived sensory memory across modalities. There is a brief-lived afterimage that lasts several hundred milliseconds from the beginning of the stimulus and is tied to the minimum duration of a perception (Efron, 1970a, 1970b, 1970c) and of the recognition process (Craig, 1985; Turvey, 1973; Massaro, 1975). There is also a longer recollection of sensation that consists of activated sensory features that seem to lose activation in a matter of seconds. The more common belief that sensory memory simply lasts far longer in audition than in vision (e.g., Darwin et al., 1972) appears problematic, inasmuch as it overlooks many procedures in which the very brief-lived afterimage exists in both of these modalities with comparable time constants.

Abstract representations are formed that have to do with the underlying type of information; for example, phonological representations are formed on the basis of either spoken or written items (Conrad, 1964). I have argued that there is a limit in this type of abstract working memory to three to five items (Anderson et al., 2011; Cowan, 2001; Luck & Vogel, 1997; Zhang & Luck, 2011).

More uncharted territory is in the memory of information that cannot be verbalized or easily categorized. We do not know, for example, if there is a fixed number of unknown voices or visual textures that can be held in mind at some level of precision, if memory for them depends on a fluid attentional resource, or if the limits are defined in some other way. The information of this sort can come either directly from a presented stimulus or perhaps, as I have suggested in this chapter, from sensory imagery.

There are also undeniable differences between the ways the modalities work. For example, information appears to be encoded with better spatial resolution in vision versus better temporal resolution in hearing (see Penney, 1989). The apparent duration of a sensory representation may depend on the type of information required, for example, on tests of temporal versus spatial distinctions (superior in the auditory vs. visual modalities, respectively). Sensory modalities also differ on whether the information basically pertains to what is going on inside the body (in the cases of proprioception, internal feelings, and pain), at the boundary of the organism and the environment (in the cases of touch, taste, and smell), and outside of the organism in the environment (in the cases of hearing and vision). A successful outcome of the symposium would be a reawakening of the field to the importance of sensory information for a better understanding of consciousness and behavior.

Acknowledgments

This work was completed with support from NIH Grant R01-HD21338. Thanks to Pierre Jolicœur, Christine Lefebvre, and an anonymous reviewer for important feedback on an earlier draft of this chapter.

This chapter based on the Symposium Lecture of 25th International Symposium on Attention and Performance, Station de biologie des Laurentides, Quebec, entitled “Mechanisms of Sensory Working Memory.” Proceedings editors: Christine LeFebvre, Pierre Jolicœur, and Julio Martinez-Trujillo.

References

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology*, *23*, 1427–1431.

Anderson, D. E., Vogel, E. K., & Awh, E. (2011). Precision in visual working memory reaches a stable plateau when individual item limits are exceeded. *The Journal of Neuroscience*, *31*, 1128–1138.

Arieh, Y., & Marks, L. E. (2003). Time course of loudness recalibration: Implications for loudness enhancement. *Journal of the Acoustical Society of America*, *114*, 1550–1556.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: a proposed system and its control processes. In K. W. Spence, & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2) (pp. 89–195). New York: Academic Press.

Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, *40*, 309–328.

Baddeley, A. D. (1986). Working memory. Oxford, England: Clarendon Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8) (pp. 47–89). New York: Academic Press.

Balota, D. A., & Duchek, J. M. (1986). Voice-specific information and the 20-second delayed suffix effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 509–516.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*, 851–854.

Bliss, J. C., Crane, H. D., Mansfield, P. K., & Townsend, J. T. (1966). Information available in brief tactile presentations. *Perception & Psychophysics*, *66*, 273–283.

Bredemeiser, K., & Simons, D. J. (2012). Working memory and inattentional blindness. *Psychonomic Bulletin & Review*, *19*, 239–244.

Broadbent, D. E. (1958). *Perception and communication*. New York: Pergamon Press.

Chein, J. M., Moore, A. B., & Conway, A. R. A. (2011). Domain-general mechanisms of complex working memory span. *NeuroImage*, *54*, 550–559.

Chen, Z., & Cowan, N. (2009). Core verbal working memory capacity: the limit in words retained without covert articulation. *Quarterly Journal of Experimental Psychology*, *62*, 1420–1429.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, *25*, 975–979.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, *27*, 183–228.

Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology*, *55*, 75–84.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: the importance of working memory capacity. *Psychonomic Bulletin & Review*, *8*, 331–335.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin*, *96*, 341–370.

Cowan, N. (1987). Auditory sensory storage in relation to the growth of sensation and acoustic information extraction. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 204–215.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, *104*, 163–191.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.

Cowan, N. (2011). The focus of attention as observed in visual working memory tasks: Making sense of competing claims. *Neuropsychologia*, *49*, 1401–1406.

Cowan, N., Blume, C. L., & Saults, J. S. (2013). Attention to attributes and objects in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 731–747.

Cowan, N., Lichty, W., & Grove, T. R. (1990). Properties of memory for unattended spoken syllables. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *16*, 258–269.

Cowan, N., Li, D., Moffitt, A., Becker, T. M., Martin, E. A., Saults, J. S., et al. (2011). A neural region of abstract working memory. *Journal of Cognitive Neuroscience*, *23*, 2852–2863.

Cowan, N., & Morey, C. C. (2007). How can dual-task working memory retention limits be investigated? *Psychological Science*, *18*, 686–688.

Cowan, N., Rouder, J. N., Blume, C. L., & Saults, J. S. (2012). Models of verbal working memory capacity: what does it take to make them work? *Psychological Review*, *119*, 480–499.

Cowan, N., Saults, J. S., & Brown, G. D. A. (2004). On the auditory modality superiority effect in serial recall: separating input and output factors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 639–644.

Cowan, N., Winkler, I., Teder, W., & Näätänen, R. (1993). Memory prerequisites of the mismatch negativity in the auditory event-related potential (ERP). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 909–921.

Cowan, N., & Wood, N. L. (1997). Constraints on awareness, attention, and memory: some recent investigations with ignored speech. *Consciousness and Cognition*, *6*, 182–203.

Craig, J. C. (1985). Tactile pattern perception and its perturbations. *Journal of the Acoustical Society of America*, *77*, 238–246.

Crowder, R. G. (1989). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception & Performance*, *15*, 472–478.

Crowder, R. G., & Morton, J. (1969). Precategorical acoustic storage. *Perception & Psychophysics*, *5*, 365–373.

Darwin, C. J., Turvey, M. T., & Crowder, R. G. (1972). An auditory analogue of the Sperling partial report procedure: evidence for brief auditory storage. *Cognitive Psychology*, *3*, 255–267.

Deutsch, D. (1970). Tones and numbers: specificity of interference in immediate memory. *Science*, *168*, 1604–1605.

Donkin, C., Nosofsky, R. M., Gold, J. M., & Shiffrin, R. M. (2013). Discrete slot models of visual working-memory response times. *Psychological Review*, *4*, 873–902.

Efron, R. (1970a). The relationship between the duration of a stimulus and the duration of a perception. *Neuropsychologia*, *8*, 37–55.

Efron, R. (1970b). The minimum duration of a perception. *Neuropsychologia*, *8*, 57–63.

Efron, R. (1970c). Effects of stimulus duration on perceptual onset and offset latencies. *Perception and Psychophysics*, *8*, 231–234.

Enns, J. T., Brehaut, J. C., & Shore, D. I. (1999). The duration of a brief event in the mind's eye. *Journal of General Psychology*, *126*, 355–372.

Farah, M. J. (1985). Psychophysical evidence for a shared representational medium for mental images and percepts. *Journal of Experimental Psychology: General*, *114*, 91–103.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1030–1044.

Gallace, A., Tan, H. Z., & Spence, C. (2007). Multisensory numerosity judgments for visual and tactile stimuli. *Perception & Psychophysics*, *69*, 487–501.

Gibson, B., Wasserman, E., & Luck, S. J. (2011). Qualitative similarities in the visual short-term memory of pigeons and people. *Psychonomic Bulletin & Review*, *18*, 979–984.

Hardman, K., Cowan, N. Remembering complex objects in visual working memory: do capacity limits restrict objects or features? *Journal of Experimental Psychology: Learning, Memory, and Cognition* http://dx.doi.org/10.1037/xlm0000031.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*, 632–635.

Kallman, H. J., & Massaro, D. W. (1979). Similarity effects in backward recognition masking. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 110–128.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective. *Psychonomic Bulletin & Review*, *9*, 637–671.

Keller, T. A., Cowan, N., & Saults, J. S. (1995). Can auditory memory for tone pitch be rehearsed? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*, 635–645.

Laing, D. G., & Francis, G. W. (1989). The capacity of humans to identify odors in mixtures. *Physiology & Behavior*, *46*, 809–814.

Lefebvre, C., Vachon, F., Grimault, S., Thibault, J., Guimond, S., Peretz, I., et al. (2013). Electrophysiological indices of maintenance in auditory and visual short-term memory. *Neuropsychologia*, *51*, 2939–2952.

Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, *24*, 61–79.

Li, D., Cowan, N., & Saults, J. S. (2013). Estimating working memory capacity for lists of nonverbal sounds. *Attention, Perception, & Psychophysics*, *75*, 145–160.

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

Majerus, S., D'Argembeau, A., Martinez Perez, T., Belayachi, S., Van der Linden, M., Collette, F., et al. (2010). The commonality of neural networks for verbal and visual short-term memory. *Journal of Cognitive Neuroscience*, *22*, 2570–2593.

Massaro, D. W. (1975). *Experimental psychology and information processing*. Chicago: Rand McNally.

Massaro, D. W. (1976). Perceptual processing in dichotic listening. *Journal of Experimental Psychology: Human Learning & Memory*, *2*, 331–339.

May, J. C., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, *47*, 66–122.

McCloskey, M., & Watkins, M. J. (1978). The seeing more than is there phenomenon: Implications for the locus of iconic storage. *Journal of Experimental Psychology: Human Perception & Performance*, *4*, 553–564.

Mewhort, D. J. K., Campbell, A. J., Marchetti, F. M., & Campbell, J. I. D. (1981). Identification, localization, and "iconic memory": an evaluation of the bar-probe task. *Memory & Cognition*, *9*, 50–67.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart and Winston, Inc.

Moray, N. (1959). Attention in dichotic listening: affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, *11*, 56–60.

Morey, C. C., & Bieler, M. (2013). Visual short-term memory always requires attention. *Psychonomic Bulletin & Review*, *20*, 163–170.

Morey, C. C., & Cowan, N. (2004). When visual and verbal memories compete: evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review*, *11*, 296–301.

Morey, C. C., Cowan, N., Morey, R. D., & Rouder, J. N. (2011). Flexible attention allocation to visual and auditory working memory tasks: manipulating reward induces a tradeoff. *Attention, Perception, & Psychophysics*, *73*, 458–472.

Murdock, B. B., Jr. (1968). Modality effects in short-term memory: storage or retrieval? *Journal of Experimental Psychology*, *77*, 79–86.

Newell, A., & Simon, H. A. (1956). *The logic theory machine: A complex information processing system*. Santa Monica, CA: Rand Corporation.

Oberauer, K., & Eichenberger, S. (2013). *Visual working memory declines when more features must be remembered for each object. Memory & Cognition*. http://dx.doi.org/10.3758/s13421-013-0333-6. E-publication ahead of print.

Olsson, H., & Poom, L. (2005). Visual memory needs categories. *PNAS*, *102*, 8776–8780.

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*, 369–378.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition*, *17*, 398–422.

Petrusic, W. M., Baranski, J. V., & Kennedy, R. (1998). Similarity comparisons with remembered and perceived magnitudes: memory psychophysics and fundamental measurement. *Memory & Cognition*, *26*, 1041–1055.

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

Pratte, M. S., Ling, S., Swisher, J. D., & Tong, F. (2013). How attention extracts objects from noise. *Journal of Neurophysiology*, *119*, 1346–1356.

Rostron, A. B. (1974). Brief auditory storage: some further observations. *Acta Psychologica*, *38*, 471–482.

Rouder, J. N., Morey, R. D., Cowan, N., Zwilling, C. E., Morey, C. C., & Pratte, M. S. (2008). An assessment of fixed-capacity models of visual working memory. *Proceedings of the National Academy of Sciences (PNAS)*, *105*, 5975–5979.

Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working-memory capacity in the change-detection paradigm. *Psychonomic Bulletin & Review*, *18*, 324–330.

Rowe, E. J., Philipchalk, R. P., & Cake, L. J. (1974). Short-term memory for sounds and words. *Journal of Experimental Psychology*, *102*, 1140–1142.

Rowe, E. J., & Rowe, W. G. (1976). Stimulus suffix effects with speech and nonspeech sounds. *Memory & Cognition*, *4*, 128–131.

Saults, J. S., & Cowan, N. (2007). A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory. *Journal of Experimental Psychology: General*, *136*, 663–684.

Serences, J. T., Ester, E., Vogel, E., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, *20*, 207–214.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends in Cognitive Sciences*, *9*, 16–20.

Sligte, I. G., Scholtel, H. S., & Lamme, V. A. F. (2008). Are there multiple visual short-term memory stores? *PLOS One*, *3*(2), 1–9.

Sokolov, E. N. (1963). *Perception and the conditioned reflex*. NY: Pergamon Press.

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

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652–654.

Sullivan, E., & Turvey, M. (1972). Short-term retention of tactile information. *Quarterly Journal of Experimental Psychology*, *24*, 253–261.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, *428*, 751–754.

Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin*, *20*, 12–16.

Treisman, M., & Rostron, A. B. (1972). Brief auditory storage: a modification of Sperling's paradigm. *Acta Psychologica*, *36*, 161–170.

Turvey, M. T. (1973). On peripheral and central processes in vision: inferences from an information processing analysis of masking with patterned stimuli. *Psychological Review*, *80*, 1–52.

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

van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*, *109*, 8780–8785.

Verrillo, R. T., & Gescheider, G. A. (1975). Enhancement and summation in the perception of two successive vibrotactile stimuli. *Perception & Psychophysics*, *18*, 128–136.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1436–1451.

Winkler, I., & Cowan, N. (2005). From sensory to long term memory: evidence from auditory memory reactivation studies. *Experimental Psychology*, *52*, 3–20.

Winkler, I., Schröger, E., & Cowan, N. (2001). The role of large-scale memory organization in the mismatch negativity event-related brain potential. *Journal of Cognitive Neuroscience*, *13*, 59–71.

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., et al. (1993). Modulation of early sensory processing in human auditory cortex. *Proceedings of the National Academy of Science*, *90*, 8722–8726.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*, 23–35.

Zhang, W., & Luck, S. J. (2011). The number and quality of representations in working memory. *Psychological Science*, *22*, 1434–1441.