


8
Using Working Memory to Control Attentional Deployment to Items in Complex Scenes

Geoffrey F. Woodman and Steven J. Luck

When we look out on the world, we are often looking for something in particular. For example, when we approach a stop sign we are looking for cross-traffic. When we look out on the soccer field, we are looking for our children. When we open the refrigerator, we are looking for the milk to pour over the cereal. These acts of searching for a particular target require us to control attention in a top-down manner, as opposed to focusing processing in a bottom-up manner on the largest, brightest, or suddenly appearing objects. A dominant view in theories and empirical studies of attention is that top-down control over attention is made possible by holding representations of the searched for items in working memory (for reviews, see Olivers, Peters, Houtkamp, & Roelofsena, 2011; Soto, Hodsoll, Rottekin, & Humphreys, 2008; Woodman, Carlisle, & Reinhart, 2013). That is, when we want to find Alison, Carter, Henry, Hunter, Paige, or Sam on the soccer field, we maintain representations of them in visual working memory. This mechanistic explanation for how we implement top-down attentional control was proposed in the earliest theories of attention and has been a hot topic of debate in recent empirical studies.

Our goal in this chapter is to discuss the origins of this idea, review recent findings, and lay out the unresolved issues about how working memory is involved in controlling attention. We begin by discussing the historical origins of the idea that the representations in memory determine what attention selects.

Historical Context

The idea that attention and working memory are intimately intertwined is as old as the field of psychology itself. James (1890) proposed that our ability to select one of the multitude of possible inputs to process was as much determined by the active thoughts in our minds as the stimuli that impinge upon us: “While the object excites it [attention] from without, other brain-cells ... arouse it from within” (p. 441). This quote captures the idea that attentional selection depends on both the bottom-up input (i.e., “without”) and the top-down representations (i.e., “within”) in mind.
This general idea was expanded upon in considerable detail by Pillsbury (1908) in his book *Attention*. He proposed that attentional selection was largely determined by the nature of our memory representations. Like James, he discussed at length the idea that attention is controlled by the representations that we keep active in primary memory (i.e., short-term memory or working memory in modern terminology). Pillsbury (1908) provides us with this concise description of the top-down control of attention during search:

> It is much easier to see any lost article if you have a definite picture of what is sought. In fact, searching for anything consists ordinarily of nothing more than walking about the place where the object is supposed to be, with the idea of the object kept prominently in mind, and thereby standing ready to facilitate the entrance of the perception when it comes itself. (p. 36)

As we will discuss below, the idea that representations stored in working memory allow us to select matching perceptual inputs has dominated the thinking of cognitive scientists and theoreticians for at least the last century. To substantiate that statement, we need look no further than the first major theory of attention proposed in the middle of the last century. Broadbent (1957) proposed that attention works like a filter to determine which perceptual input channels receive the benefit of detailed analysis and can therefore be recognized. Stimuli presented in other input channels went unrecognized. His elegant mechanistic analogy for the operation of attention was a system of tubes that converged on a bottleneck governed by the attentional filter that worked like a flap to gate which inputs were allowed access to the bottleneck. However, Broadbent also proposed that this system had temporary storage of information in short-term memory via recirculation tubes that could keep inputs from an unselected channel (or tube) circulating through the system for a brief period. This allowed the model to account for findings from immediate memory tasks that were recent at that time (Brown, 1954).

The idea that memory representations are integral to the selection process carried out by attention was carried to its most extreme theoretical position by Norman (1968). This theory proposed that what we call attention is a consequence of having a limited number of representations active in memory. This perspective continues to be well-represented in current models of working memory (Cowan, 1998; Cowan, 2001; Oberauer, 2002). This brief historical review shows us how the idea that temporary memory (call it primary, short-term, or working memory) controls attention has a long, rich history that has shaped the development of theories since the dawn of psychology as a science of the mind.

**State-of-the-Art Review**

The first empirical tests of the proposal that working memory controls attention in the modern era of cognitive psychology were performed by Logan (1978, 1979). He used a dual-task procedure that would become standard in such studies decades later. The procedure involved filling working memory with information and then asking the subjects to perform another task before testing the subjects' memory at the end of the trial. The logic was simple: If working memory plays a critical role in the processing of information for the task performed during the memory retention interval, then performance of that task should be significantly impaired relative to a comparison condition in which the task was performed in isolation. Logan (1978) had subjects discriminate the identity of letters while maintaining digits in memory. He reported that subjects were slower to discriminate the target letters when the task was performed during the memory retention interval (i.e., general dual-task interference that changed the y-intercept), but that the set size effects were unchanged by the working memory load (i.e., the slopes of the reaction time [RT] functions were unchanged; Logan, 1979).

**Deploying Attention When Visual Working Memory is Full**

We were among the first group of researchers to adopt a dual-task procedure similar to that of Logan (1978) decades later (Downing, 2000; Downing & Dodds, 2004; Woodman, Vogel, & Luck, 2001). We describe our methods and findings from an example experiment to serve as a concrete illustration. The goal of this empirical work was to test the hypothesis that working memory is critical for the attentional selection of targets that we search for in complex scenes (Woodman et al., 2001; Woodman & Luck, 2010). The paradigm involved three different experimental conditions. In the search-alone condition, subjects performed a search task in isolation. In the memory-alone condition, they performed a working memory task in isolation. In the search-and-memory condition, they performed the search task during the retention interval of the working memory task. Figure 8.1 shows the memory-alone and search-and-memory conditions of Experiment 1 from Woodman and Luck (2010).

Although these experiments were published over 20 years after Logan's studies, the logic remained the same. If you fill visual working memory to capacity with the memory task, then working memory will be unavailable for use in the visual search. If working memory is needed for efficient search, this unavailability of visual working memory should reduce the efficiency of visual search (which is typically assessed by means of the slope of the function relating RT to the set size of the search array). Alternatively, subjects might clear out visual working memory to make it available for the search task. If this were the case, then we should see that they perform extremely poorly on the memory task when performed concurrently with the search task.

Across many experiments of this sort, we found no evidence for either of these predictions (Woodman & Luck, 2010; Woodman et al., 2001). As illustrated in figure 8.2, we found that the slopes of the search functions were not different whether search was performed in isolation (i.e., search alone) or during the retention interval of the memory task (i.e., search and memory). The slope of this function is a standard
Figure 8.1
An example of the sequence of stimuli presented during the dual-task condition of Experiment 1 of Woodman and Luck (2010). The letters to the left indicate that subjects began each trial by repeating the articulatory-suppression load. There were 500 ms intervals between each stimulus frame in which the screen was blank except for the central fixation point. Reprinted with permission of Psychology Press, Taylor & Francis Group.

Figure 8.2
The results from Experiment 1 of Woodman and Luck (2010). (A) The visual search RTs from the dual-task condition (dashed lines, empty symbols) and the search-alone condition (solid lines, filled symbols). (B) The change-detection accuracy in the memory-alone condition, left, and the dual-task condition as a function of set size or the presentation of a blank, right. The error bars in this and the subsequent figure show the 95% within-subjects confidence intervals (Lofquist & Lofquist, 1988). SS, set size. Reprinted with permission of Psychology Press, Taylor & Francis Group.

measure of the efficiency of the search process because it shows how well the visual system can cope with increased demands. In contrast, the y-intercept reflects the processes that occur prior to the search process (e.g., preattentive sensory analysis) and after the search process (e.g., response selection and execution). We consistently observed an increase in the y-intercept when search was performed during the memory task, but no change in slope. Thus, the presence of a visual working memory load did not slow the search process but only slowed the processes that precede or follow the search process.

In addition, performance of the memory task showed a consistent but small loss of information by visual working memory when search was performed during the memory task. This loss appears to be a function of presenting any stimulus during the memory retention interval. That is, this drop of between one half and one object's worth of information from visual working memory is observed even when a completely task-irrelevant stimulus is presented (Quinn & McConnell, 1996; Woodman et al., 2001), but not when no search array is presented although expected by the subjects (see the right bar in figure 8.2b). It may reflect an obligatory storage of the target, or perhaps any object for which a saccade has been planned, in working memory (Hollingworth, Richard, & Luck, 2008).

The findings of these dual-task experiments indicate that storage of information in visual working memory is not necessary for the efficient processing of objects in complex scenes. Recently, interest in this question has increased with further evidence from neuroimaging, event-related potentials (ERPs), and behavior. ERP and behavioral findings have been used to argue that items need to pass through visual working memory when interitem similarity is high within a complex scene (D. E. Anderson, Vogel, & Awh, 2013). In addition, neuroimaging and behavioral findings have suggested that novel nontargets require active suppression during search (Seidl, Peelen, & Kastner, 2012), a process that appears to require access to visual working memory (Risman,
Gazzaley, & D'Esposito, 2009). We will return to this topic of the active suppression of
distractors later in our discussion.

Although this issue continues to be debated, it is clear that attention can efficiently
select target objects in a complex scene when other objects are being maintained in
visual working memory. These findings indicate that when attention is used to search
for and find targets in a complex scene, all the information in the scene does not
need to pass through the limited-capacity bottleneck of visual working memory. For
example, nontarget items that are selected by attention during search are unlikely to be
represented in visual working memory.

In the visual search tasks used in these experiments, subjects searched for the same
target on every trial. This is a common laboratory practice in visual search studies,
but it does not reflect the way visual working memory is typically used in real-world
tasks. In the real world, we typically search for different targets at different moments.
For example, when making cookies, one might first search for the butter, then the
brown sugar, and then the vanilla. In this kind of situation, it seems much more likely
that working memory would be necessary to guide attention efficiently. To test this
hypothesis, we conducted a follow-up experiment in which the identity of the target
was cued on each trial (Woodman, Luck, & Schall, 2007). Again, subjects performed
in memory-alone, search-alone, and memory-and-search conditions. We found that
the search slope was much higher in the memory-and-search condition than in the
search-alone condition. Thus, there are situations in which working memory plays an
absolutely essential role in visual search. However, working memory does not play this
role when the target remains constant from trial to trial.

Are Attentional Templates Stored in Visual Working Memory?
The dual-task approach used in the experiments described so far is a fairly blunt
instrument for examining the relationship between attention and working memory.
Although it is useful to determine that attention can still be efficiently deployed when
visual working memory is full, this does not rule out the possibility that visual working
memory enables the deployment of attention in a way that does not require a large
amount of storage. Specifically, a number of theories of attention have proposed that
the role played by visual working memory typically only requires the storage of one
representation.

Many theories of attention propose that visual working memory is used to store
a representation of the searched-for target when we process a complex scene (e.g.,
Bundesen, 1990; Bundesen, Habekost, & Kyllingsbaek, 2005; Estes & Duncan,
1993; Navalpakkam & Itti, 2005). The simple idea is the top-down control of attention
is implemented by holding a representation of the target in visual working memory.
For example, when we want to find Sam on the soccer field, we retrieve a representa-
tion of Sam from long-term memory, and we maintain that representation in working
memory. In doing so, the representation in visual working memory feeds back to the
neurons in visual cortex that code for brown hair, brown eyes, the motion of his gait,
the red of his shirt and shoes, and so forth. This proposal could explain why the experi-
ments described above found that search displaced a minimal amount of information
from visual working memory (i.e., one-half to 1 object's worth of information).

The theoretical proposal that visual working memory is the source of top-down
attentional control has garnered considerable empirical attention. Many researchers
hypothesized that if visual working memory representations are used to control atten-
tion, then maintaining a certain representation in visual working memory should
be sufficient to bias attention to similar objects when they appear in our visual field
(Carlisle & Woodman, 2011a, 2011b; Downing, 2000; Downing & Dodds, 2004; Hol-
lingworth & Luck, 2009; Houtkamp & Roelfsema, 2006; Olivers, Meijer, & Theeuwes,
2006; Olivers et al., 2011; Soto, Heinke, Humphreys, & Blanco, 2005; Soto et al.,
2008; Woodman & Chun, 2006; Woodman & Luck, 2007). In figure 8.3A (plate 1), we show
the paradigm that we used in one of our experiments to test this proposal (Woodman
& Luck, 2007). The logic is that if we maintain a representation of a red object in visual
working memory, then a red object in our visual field should receive the benefit of
attention. Most of these experiments involve presenting visual search arrays that either
do or do not have a critical item that matches the representation in visual working
memory. When the memory-matching item is a distractor (i.e., a nontarget object) in
the array, it should attract attention, and this should slow the deployment of attention
to the actual target object (compared to when no memory-matching distractor is present).

Figure 8.3B (plate 1) shows the pattern of results obtained from one of these experi-
ments. We found that detection of the visual search target was not significantly slowed
by the presence of a memory-matching distractor, provided subjects knew that the tar-
get object in the visual search array would never match the item in memory. However,
when subjects knew that there was a chance that the target could match the memory
item, we found that memory-matching distractors did slow search. We concluded that
holding an object in working memory does not cause an obligatory capture of attention
by objects that match the features of that object. However, cognitive control pro-
cesses can allow people to use working memory to guide search when this is adaptive.
However, the story did not end with these observations.

Our initial study suggested that people used the contents of working memory to
deploy attention in the most adaptive manner. However, other researchers obtained
findings suggesting that subjects attend to memory-matching items under all condi-
tions, even when doing so has no benefit for search performance and can only impair
search for the target object (e.g., Soto et al., 2005; Soto et al., 2008). Previously, we
proposed that such deployments of attention to items matching representations in
working memory might be made strategically to improve the memory representation
memory (Raye, Mitchell, Reeder, Greene, & Johnson, 2008; Yi, Turk-Browne, Chun, & Johnson, 2008). This idea provides a logical explanation for why subjects would attend to memory-matching items despite these items' being in opposition to efficient search performance.

The weakness of this perceptual-resampling account is that these strategic attentional deployments do not appear consistent with all deployments of attention to memory-matching items (Balani, Soto, & Humphreys, 2010). For example, in the study of Balani and colleagues (2010), RTs during search were consistent with subjects' attending to memory-matching items; however, subjects were no more accurate for items that reappeared in the visual search arrays than items that did not. As we discuss next, the mixture of results and explanations that we have just reviewed led some researchers to step back and reconsider the theoretical proposals tested with these methods.

The idea that we automatically attend to any input matching a representation in visual working memory is a reasonable extension of the proposal that top-down attentional control comes from representations in visual working memory (Soto et al., 2008). However, alternatives have been proposed where the representations in visual working memory do not directly interact with perceptual selection of inputs without an additional step. In the proposed framework of Olivers and colleagues (Olivers et al., 2011), one visual working memory representation can be in an active state, and only this one active representation controls attention (see also Downing & Dodds, 2004). However, more recent research shows that it is possible to maintain at least two representations in visual working memory that can simultaneously control attention (Beck, Hollingworth, & Luck, 2012).

A slightly different proposal comes from the theory of visual attention framework (Bundesen, 1990; Bundesen & Habekost, 2008; Bundesen et al., 2005). These theorists have proposed that an executive-control process intervenes between storage in visual working memory and perceptual attention that enables a memory representation to influence attentional selection (Bundesen & Habekost, 2008).

Is the Deployment of Attention to Memory-Matching Items Being Actively Suppressed?

Overall, studies have obtained widely divergent results when asking whether attention is automatically attracted to items that match the information currently being stored in working memory (Carlisle & Woodman, 2011a, 2013; Olivers, 2009). This is analogous to the divergent results that have been obtained for decades in studies asking whether salient stimuli automatically capture attention (Egeth & Yantis, 1997; Folk, Remington, & Johnston, 1992; Theeuwes, 1992; Van der Stigchel et al., 2009). We have recently proposed a signal suppression hypothesis that can potentially explain many of the divergent results in both of these domains (Sawaki & Luck, 2010, 2014). Specifically, this hypothesis proposes that physically salient objects and objects that match the
contents of working memory automatically generate an attentional priority signal (an attend-to-me signal). However, this signal can be squashed by a top-down suppression process before attention is actually captured.

Evidence for this hypothesis comes from ERP recordings, in which an N2pc ERP component is observed when attention is allocated to a stimulus and a Pd ERP component is observed when a stimulus is suppressed. Under conditions that tend to cause behavioral evidence of attentional capture by salient stimuli, previous studies found that a salient but irrelevant stimulus elicits an N2pc component, consistent with capture of attention by the stimulus (e.g., Sawa & Luck, 2010; Sawa & Luck, 2014). However, under conditions that minimize behavioral measures of capture, the salient stimulus does not elicit an N2pc component and instead elicits a Pd component (Ransley & Klie, 2008; Hickey, McDonald, & Theeuwes, 2006) consistent with an active suppression process. Similarly, when a task-irrelevant stimulus array is presented during the delay interval of a visual working memory task, a Pd component is observed when an object in this task-irrelevant array matches the contents of working memory (Sawa & Luck, 2011). We therefore suggest that at least some of the differences across studies in the involuntary capture of attention by both salient stimuli and memory-matching stimuli reflect differences in the extent to which subjects employ the top-down suppression process reflected by the Pd component to squash the priority signal created by these stimuli. When this suppression process is used consistently, no capture is observed. However, when this suppression process is not used on a proportion of trials, some evidence for capture may be observed.

The active suppression mechanism that we just discussed is likely to be particularly important in several real-world settings. Primates often use eye-movement behavior in socially important ways. For example, among many species of primates, directing gaze to the eyes of an individual that is higher in the social hierarchy is seen as a challenge to that individual’s dominance. Clearly, it is adaptive to avoid overt eye movements to such potentially threatening targets to avoid unnecessary social conflict. Because there is such a tight link between covert and overt selection (Hoffman & Subramaniam, 1995), this is likely to involve the suppression of covert processing of such threat stimuli to prevent shifts of gaze, and has the added benefit of suppressing processing of the threatening gaze, reducing the emotional stress that accompanies the processing of such information.

Directly Measuring the Attentional Templates in the Brain
Most of the research described so far used a dual-task approach to look at interactions between attention and working memory. However, in some of our recent work we wanted to take the more basic approach of having subjects perform a visual search task while we eavesdropped on the normal operation of visual working memory with ERPs. Our approach was based on the cued visual search paradigm used by Chelazzi

and colleagues in their recordings of single-unit activity from the ventral visual cortex of monkeys (Chelazzi, Duncan, Miller, & Desimone, 1998; Chelazzi, Miller, Duncan, & Desimone, 1993, 2001). On each trial, the monkey was shown a cue stimulus that defined the target for an upcoming visual search task. It was observed that the neurons that coded for the target in the inferotemporal cortex showed an elevated firing rate during the period between the offset of the cue and the onset of the search array. This was interpreted as evidence that the monkeys were actively maintaining a representation of the cue in working memory during this interval, which then biased attention toward the target in the search array. This is like the experiment described earlier (Woodman et al., 2007) in which we cued the target on each trial rather than remaining constant across trials. A memory load interfered with search when the target was cued on each trial, implying that working memory plays a key role in visual search under these conditions.

We have recently used this trial-by-trial cuing approach with ERP recordings in human subjects to provide direct evidence that the cue representation was being held in visual working memory. This research exploited an ER component known as the contralateral-delay activity (or CDA), a sustained negativity that is observed during the delay interval in working memory paradigms. The CDA is contralateral to the location in the visual field where the to-be-remembered objects were initially presented. The amplitude of the CDA reflects the number of objects maintained in visual working memory and is highly correlated with behavioral measures of visual working memory capacity (Vogel & Machizawa, 2004; Woodman & Vogel, 2008). The CDA has a different time course and voltage distribution over electrodes compared to other lateralized components related to attention and response preparation (McCormick, Machizawa, & Vogel, 2007). We have also observed this ER component in macaque monkeys and have shown that it is driven by sustained memory-related activity in prefrontal structures (Reinhart et al., 2012). As we discuss next, the CDA can be harnessed to study the involvement of visual working memory in controlling visual attention during visual search.

Figure 8.4 (plate 2) shows the CDA while subjects performed a visual search task in which the identity of the target was cued at the beginning of each trial and changed from trial to trial (Carlisle, Arita, Pardo, & Woodman, 2011; Woodman & Arita, 2011). To provide a balanced sensory input, the cue stimulus included a red item and a green item, and subjects were supposed to search for the shape indicated by the red cue and ignore the green cue (or vice versa). In the interval between the target cue offset and the onset of the search array, we found a sustained negative potential contralateral to the relevant cue item. The CDA amplitude in this interval, before search began, predicted the speed and accuracy of the search response (Carlisle et al., 2011; Reinhart & Woodman, 2014; Woodman & Arita, 2011). Then, we showed that the CDA in this interval was twice as large when two relevant cues were present (indicating two possible target
Figure 8.4 (plate 2)
The stimuli and event-related potential findings from Woodman and Arita (2011). Example of the stimulus sequence (left) and the grand average waveforms from electrodes TS/6, contralateral (red) and ipsilateral (black) to the location of the cue on each trial (right). The gray region shows the epoch in which the significant contralateral-delay activity (CDA) was measured, and the inset shows voltage distribution. The amplitude of the CDA predicted the accuracy of the subsequent search across subjects (p < .05). Adapted with permission from the Association for Psychological Science and Blackwell Publishing.

Figure 8.5
The grand average event-related potential results, time-locked to the cue presentation from consecutive groups of trials. The plot shows the contralateral-delay activity (CDA) amplitude across consecutive trials with the same search target. The gray line shows the power-function fit, and the error bars represent ±1 S.E.M. Adapted with permission from the Society for Neuroscience.

Items that might occur in the search array). These findings provide strong evidence that visual working memory representations were used to control the deployment of visual attention when the task involves looking for a different target on each trial (Carlisle et al., 2011). However, our previous results showing no interference between working memory and visual search in dual-task experiments suggests that working memory plays little or no role in search when the target identity remains constant on every trial.

The idea is fairly simple: When we are frantically looking for our keys in our kitchen on our way to work, it is best to hold a representation of those keys in working memory. However, what happens as you look for those keys in room after room? Perhaps longer term memory systems can take over the job of controlling attention in this situation, freeing up working memory for other tasks.

To test this hypothesis, we conducted experiments in which the same target was cued across three to seven consecutive trials, and then the cue changed to specify a different target (Carlisle et al., 2011; Reinhart & Woodman, 2014; Woodman et al., 2013). We predicted that the amplitude of the CDA would decrease after several presentations of the same target, and this is what we found. As shown in figure 8.5, CDA amplitude decreased over successive presentations, suggesting the target was no longer represented in visual working memory and that some longer term memory system took over the guidance of attention. Moreover, when subjects searched for the same target shape for the entire experiment, there was no CDA following the target cue (Carlisle et al., 2011). These findings provide direct evidence of a transition from reliance on visual working memory representations to reliance on longer term memory representations in the top-down control of visual attention. As we will discuss further below, these findings also provide strong converging evidence for the idea that the development of automaticity involves a transition from working memory to long-term memory, which has been a cornerstone of many theories of learning and skill acquisition (J. R. Anderson, 1982, 2000; Logan, 1988, 2002; Rickard, 1997) and supported by inferences from behavioral studies (e.g., Woodman et al., 2007). This recent work shows how our measurements of ERPs can be used to more clearly define the roles of various memory representations in controlling attention and shape future theories of attention.

In the study that we just described, it was inferred that long-term memory representations accumulated to enable the takeover of attentional control from visual working memory. Although this is a reasonable inference, we would be far more confident in this conclusion if we could watch this information accumulation in long-term memory. To do this, several recent studies (Reinhart & Woodman, 2014; Woodman et al., 2013) measured a different ERP component, P170, that indexes long-term memory. Voss and colleagues (Voss, Schendan, & Paller, 2010) reported that the amplitude of this frontal
that the research we just reviewed demonstrates the inadequacy of these classic statements about the nature of the memory representations that control attention as we search for objects in our environment. Clearly, long-term memory representations are at least equally important in the top-down control of perceptual attention. However, if we return to Pilisbry (1908), we see that he also foreshadowed the idea that the representations in long-term memory are important for the selection of perceptual inputs by attention.

Not only do ideas that are actually present in consciousness at the moment have an influence in determining what impressions shall become conscious, but other experiences, which are much more remote in time and not in consciousness at the moment, also play a part. We can trace these other conditions backward in time, and as they become farther distant they also become more general and harder to trace as individual influence because combined with others in a total complex. (p. 30)

This elegant quote indicates that our recent ERP observations were predicted over a century before they were made. In addition, the above quote foreshadows one of the next obstacles that we will face as we accept the reality that both working memory and long-term memory combine to provide the top-down control of attention. That is, if all of our background knowledge could be contributing to what attention selects at any given moment (e.g., Moore, Laiti, & Chelazzi, 2003), then how do we study the deployment of attention without understanding the intricate dynamics between the representations in long-term memory?

Integration

The laboratory experiments that we discussed in this chapter have important implications for how we operate in our environment outside the laboratory. We began by discussing how representations in memory allow us to find specific objects in our environment. Whether they are our children, food, or threats to our safety, we appear to store representations of what we are looking for in visual working memory when we enter a new environmental context or search for a new target. For example, if we are building a sandwich (e.g., Hayhoe, Shrivastava, Muczek, & Peiz, 2003), then we need to switch from looking for the bread, to looking for the jelly, the knife, and then the peanut butter. Or if we are working on our car, we need to switch from looking for the bolts to remove to finding the socket needed to remove those bolts. It seems that many real-world tasks like these involve rapidly switching between looking for different specific targets to accomplish the complex task we are performing. The laboratory experiments described in this chapter clearly show that these switches between targets require target representations to be shuttled in and out of visual working memory to perform these real-world tasks. However, the involvement of long-term memory representations in well-practiced skills suggests that this is not the complete story.
As cognitive scientists, we often use driving as an example of a real-world task in which we need to attend to task-relevant information that is vital for our survival and that of others on the roadway (e.g., Kang, Hong, Blake, & Woodman, 2011; Woodman & Vogel, 2005). However, in this task it is likely that long-term memory plays a dominant role in allowing us to search for multiple possible targets at the same time with high efficiency (Wolfe, 2012). In these tasks, visual working memory and long-term memory are likely to work in concert to allow us to attend to possible targets and to be particularly sensitive to certain likely targets (Reinhart & Woodman, 2014). For example, in a school zone we might hold a representation of a child pedestrian in working memory while long-term memory continues to bias attention to trucks, icy patches on the road, distracted drivers, and the flash of red brake lights, leaving the limited capacity of working memory to handle the representation of the child entering the roadway. Consistent with the goal of this book to integrate findings from the laboratory with real-world processing demands, we believe that our ability to use multiple memory representations to simultaneously set attentional weights is critical in understanding the nature of attentional control in any setting (Woodman et al., 2013).

The setting of attentional weights using memory representations brings us to an issue that is often neglected in the discussion of attentional templates, but is an issue in the construction of models of attention. Specifically, setting attentional weights is not likely to be a question of just determining the relative importance of a variety of targets. Instead, it is seems apparent that learning what is irrelevant is also important. With our limited-capacity processing of visual information, it is important to not waste processing capacity on irrelevant information. In the context of our driving example, if we are going to be able to process task-relevant information like brake lights or children entering the roadway, then we need to avoid wasting resources on the trees, the clouds drifting by, and the attractive dashboard in our vehicle. This may sound obvious, but the idea that attention needs to be tuned to avoid certain nontargets is not always explicitly dealt with in models of attention (e.g., Bundesen & Halbeekos, 2008).

A growing body of evidence in the laboratory that appears to show how we come to ignore the flashing light on our dashboard or the pop-up banner on our desktop. Vecera, Cosman, Vatterott, and Eoper (2014) recently showed that initially task-irrelevant objects that are unique in an array of items capture attention in the laboratory. However, as subjects in these laboratory experiments search for items of a specific color, only items that match that color come to capture attention. Returning to our driving example, this could explain why the onset of a red light on our instrument cluster captures attention because it matches the onset of red brake lights that we are vigilant for while driving. In contrast, the onset of a green light on our dashboard, matching a green traffic light that requires no change in behavior, is not processed to the same degree. At the far extreme, there are the stimuli in our field of view that are completely task irrelevant and would be maladaptive to waste our limited-capacity processing on. For example, if we were to process the vents on our dashboard with an attentional weight above zero, then this means that these stationary, completely irrelevant stimuli in the environment would be robbing task-relevant inputs of our full focus of attention. Instead, the work we have reviewed above indicates that we can set attentional weights to objects such as the vents on the dashboard to levels below zero, suppressing attentional selection of these items completely and setting positive attentional weights for objects that are task relevant. The growing body of work that we have discussed points to long-term memory as the source of the representations that help determine which inputs receive negative versus positive attentional weights (Bundesen & Halbeekos, 2008; Vecera et al., 2014).

Future Directions

We believe that the next major challenge that we face is how to understand the control mechanisms that allow attention to be guided by working memory representations at certain times and by long-term memory representations at others. Moreover, there are many different types of long-term memories that might compete for the control of attention (e.g., episodic memory, priming, operant conditioning, and classical conditioning). This important topic of future study has been identified by recent theoretical proposals (Olivers et al., 2011) as a great need for our understanding of the top-down control of attention. At Hillsbury (1968) described above, understanding attention could require us to simply solve the problem of understanding memory, a problem of sufficient scale that it could become demoralizing to future attention researchers. Instead of walking away from the entire endeavor, we will review three approaches to understanding the integration of memory and attention that we believe are particularly promising.

The first theoretical perspective that we highlight proposes that attention and memory are not separate cognitive mechanisms at all. As we described briefly in a previous section, this perspective has a long tradition. Norman (1968) proposed the first major model built on the idea that attention and memory are really parts of the same cognitive system. This theoretical perspective has the strength of naturally explaining a tight relationship between attention and different forms of memory. Because working memory representations are just the representations in long-term memory that are within the focus of attention (by some accounts), this perspective can fairly easily accommodate the idea that either these working memory representations or other long-term memory can influence which inputs are selected without proposing different mechanisms for the guidance of attention by working memory and long-term memory. However, with this perspective it is less clear how attention can be used to select certain perceptual inputs while simultaneously maintaining other working memory representations, such as in the dual-task experiments that we reviewed above. This
Controlling Attention with Memory

The framework could be modified by the addendum that one working memory representation is special in being able to guide attention (Oliviers et al., 2011), but it remains to be seen whether this modification will be able to account for how attention is deployed as complex scenes are analyzed (e.g., Beck et al., 2012) and whether overlapping mechanisms of attention and working memory can exhibit sufficient flexibility to account for the dissociations that have been observed (e.g., Woodman & Luck, 2003).

The second theoretical perspective that could account for the use of both working memory and long-term memory representations to control attention is the multiple-component working memory model of Baddeley and colleagues (Baddeley, 1986, 2007; Baddeley & Hitch, 1974; Baddeley & Logie, 1999). The key component in this model that allows it to account for the findings we reviewed here is the central executive. This is an executive-control mechanism, or collection of mechanisms, that supervises the storage of information in working memory, long-term memory, and the selection of new perceptual inputs for further processing. This far-reaching mechanism obviously has the ability to account for how perceptual selection could initially be controlled by a representation in working memory and then later be controlled by representations that accumulate in long-term memory. The only weakness of this theoretical perspective is that it accounts for the dynamics of attentional control across task performance with a homunculus-like mechanism that monitors the contents of multiple memory stores and the demands of the task at hand. Although, as cognitive scientists, we have a natural tendency to favor simpler explanations over those that require such an intelligent agent (Attenweaver, 1960), it does not mean that this theoretical perspective will ultimately fail.

The final theoretical perspective that we believe may be particularly fruitful in understanding the nature of the memory representations that control attention comes from models of automaticity and skill acquisition. The earliest theories of the learning that underlies our ability to become proficient at new tasks proposed that we initially rely on declarative memory representations that we keep actively in mind. Then, as we become better at the task, we transition to relying on long-term memory representations that guide task performance without our actively maintaining them (Hillis & Posner, 1967). This general idea has been refined and used to account for our ability to become automatic at performing tasks such that our Information Processing capacity limits are drastically reduced or even eliminated (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). The most sophisticated of these models propose that improved processing is due to both working memory and long-term memory representations racing to guide processing (Logan, 1988, 2002) with the accumulation of long-term memory representations coming to dominate processing and selection with practice. Because this proposal hinges on the simultaneous contributions of both working memory and long-term memory representations, and because of its mathematical specification, we believe that this theoretical framework could potentially account for the body of evidence that we reviewed here, as well as providing novel predictions. However, as we await further modeling work, the jury is still out with regard to whether competing theories can account for the entirety of the body of evidence that we have reviewed here.

How memory representations provide top-down control over the deployment of attention is an active topic of research. Ideally, we are entering a new phase of inquiry where our experiments will be guided by contrasting predictions that allow us to distinguish between the competing theoretical explanations discussed above. We hope that the result will allow us to meet the challenge identified by Pillsbury (1908) to identify and trace the contributions of our memory representations to determining what information is selected by attention.

Box 8.1
Key Points
- The top-down control of attention is governed by representations in both working memory and long-term memory.
- Although the bulk of empirical and theoretical work has focused on how representations in working memory exert control, long-term memory representations appear vital in determining how attention is deployed in the laboratory and the real world.
- These memory representations appear to be used to determine which items are selected for preferential processing, as well as determining which items are actively ignored.

Box 8.2
Outstanding Issues
- Given that some models of working memory propose that attention is just the most activated portion of working memory, how do these models account for the ability of subjects to efficiently select certain inputs while maintaining other information in visual working memory?
- Some models of working memory propose that working memory consists of the focusing of attention on long-term memory representations. If this is true, then how does attention also operate to select sensory inputs for perceptual processing and for storage in working memory?
- What are the key factors that determine when long-term memory representations control attention?
- Given the evidence that long-term memory helps provide top-down control over attention, then can we understand attention without first understanding the encoding, maintenance, and retrieval of information from long-term memory?
- Can models account for the shift from initially using working memory representations to using long-term memory representations to control attention without positing an intelligent central executive that controls this shift?

- Here we focused on experiments that used visual search tasks in which attention needs to select targets in a field of distractors. It remains to be seen if the same dynamics of attentional control (i.e., shifting from working memory to long-term memory representations) occur in other laboratory paradigms, such as spatial cuing, go/no-go tasks, the n-back paradigm, and so forth.

Acknowledgments

G.E.W. was supported by grants to from the National Eye Institute (R01-EYO19882, P30-EYO08126, T32-EYO07135) and the National Science Foundation (BCS 09-57072). S.J.L. was supported by a grant from the National institute of Mental Health (R01-MH076226).

References


Controlling Attention with Memory


