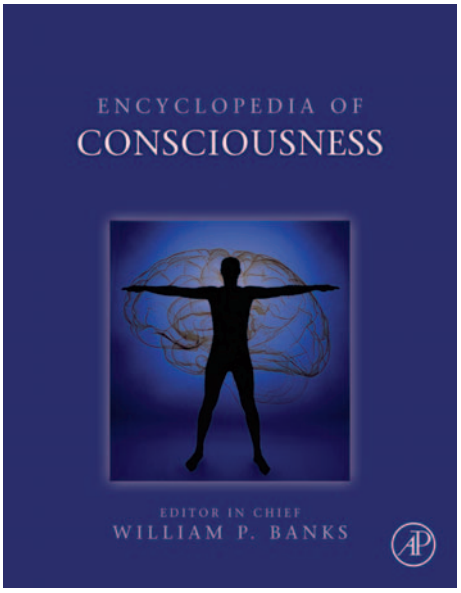
Provided for non-commercial research and educational use.
Not for reproduction, distribution or commercial use.

This article was originally published in the *Encyclopedia of Consciousness* published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

http://www.elsevier.com/locate/permissionusematerial

Levy B J and Anderson M C (2009), The Control of Mnemonic Awareness. In: William P. Banks, (Editor), Encyclopedia of Consciousness. volume 1, pp. 205-219. Oxford: Elsevier.

Author's personal copy

The Control of Mnemonic Awareness

B J Levy and M C Anderson, University of Oregon, Eugene, OR, USA

© 2009 Elsevier Inc. All rights reserved.

Glossary

Inhibition – Either the subtractive process by which the activation level of a mental representation is actively reduced or the resulting state of reduced activation.
Inhibitory control – The engagement of a controllable mechanism to lower the activity of a mental representation.
Response override – The stopping of a strongly habitual, prepotent response by control mechanisms.
Retrieval stopping – A specific instance of response override, where one must prevent a memory from entering conscious awareness by overriding the retrieval process itself.
Selective retrieval – Another instance of response override in memory that occurs when a cue is related to a strong, prepotent memory that is not currently desired. Remembering some other weaker memory in response to that cue requires the rememberer to override retrieval of the prepotent memory.

Introduction

Every person possesses a fundamentally private conscious awareness that defines their sense of their own existence. This sense includes awareness of perceptions from the surrounding environment, internal bodily states, as well as thoughts, ideas, and memories that may enter consciousness at any moment. Although awareness is often steered by stimuli around us and by our relatively automatic responses to these stimuli, consciousness can also be controlled. We can voluntarily bring to mind some past experience unbidden by any reminder, we can willfully change the direction of our thoughts, or we can focus awareness on a single idea to the exclusion of all else. Indeed, goal-directed cognition relies on the capacity to control awareness. Because this sense of awareness is central to who we are and because our capacity to control awareness is strongly connected to being goal-directed agents, scientific theories of consciousness need to explain how such control is achieved. What permits us to think about some things and not others? How do we regulate the focus of awareness?

In this article, we focus on a theoretical hypothesis about how this type of control is achieved: the response override hypothesis. According to this theory, people control awareness of memories and ideas by engaging executive control mechanisms that were originally developed to control overt motor action. In particular, controlling memory may be a special case of a general situation requiring executive control, referred to as response override. In these situations, one must stop a habitual response to a stimulus due to situational demands – an ability that is crucial for voluntary control (see **Figure 1**). For instance, after knocking over an object one might reflexively reach out to catch the item and stop its fall. If the falling object is a cactus, however, this otherwise useful perceptuomotor reflex must be overridden to prevent this pain-inducing response. This type of control is widely thought to be accomplished by inhibitory processes that suppress the inappropriate response. According to the response override hypothesis, this same inhibitory mechanism operates within the domain of memory to override the retrieval process, providing the mechanism that allows us to control the current contents of conscious awareness.

Two basic memory situations requiring response override have been identified and studied: the need for selection during retrieval and the need to stop retrieval itself. Selection is required when our goal is to recall an event or fact from long-term memory in the face of interference or distraction from

Encyclopedia of Consciousness (2009), vol. 1, pp. 205-219

Author's personal copy

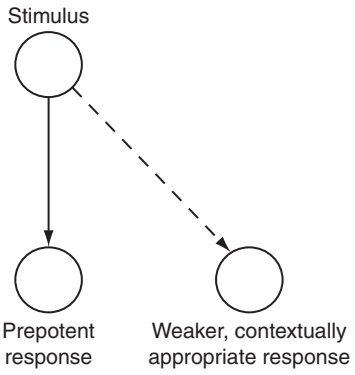


Figure 1 A typical response-override situation. In this figure, a stimulus is associated with two responses, one of which is stronger (prepotent) and the other weaker (indicated by the dotted line). Response override occurs whenever one needs to either select the weaker, but more contextually appropriate response, or to simply stop the prepotent response from occurring. Inhibitory control is thought to achieve response override by suppressing activation of the prepotent response. This basic situation describes many paradigms in research on executive control, including the Stroop and go/no-go tasks.

related traces. The need to stop retrieval arises when we confront a reminder and wish to prevent an associated memory from entering conscious awareness. Both of these processes are necessary for goal-directed cognition as they allow flexible control over whether retrieval is completed and, if so, which memory is retrieved. In both selection and stopping situations, attempts to limit awareness of activated and distracting memories impair memory for those traces later, highlighting an important connection between controlling consciousness and forgetting.

Lessons from Studying Perceptual Awareness as a Model System

The history of psychological research on consciousness has been primarily concerned with perceptual awareness. In these early studies, researchers attempted to isolate situations where a percept enters awareness in order to study the changes – both psychological and neurophysiological – that occur at the boundary between unconscious and conscious perception (see other sections of this volume). Even within perception, research on consciousness has focused primarily on visual perception. For example, studies of binocular rivalry have looked at how conscious awareness flips back and forth between two different stimuli that are independently presented to each eye. Pioneers in this area, such as Cristoph Koch, have argued that this focus has been necessary for initial progress because our understanding of the neurobiological systems underlying sensory processing, particularly vision, is quite advanced compared to those involved in higher-order cognition. Although this approach has been profitable in understanding how conscious perception arises, this research does not address much of what concerns our conscious experience. Our thoughts are not driven solely by external stimuli: they are also influenced by thoughts, ideas, and motivations. Additionally, work on perceptual awareness has focused more on the neural correlates of conscious states and less on how awareness can be controlled. One fundamental goal of theories of consciousness, however, should be to explain how control over consciousness is achieved. Building a more complete model of the internal conscious state requires understanding not only how perceptions become conscious but also how we regulate which memories and perceptions reach awareness and which do not.

Work on conscious perception provides several lessons that can be applied to the study of consciousness within memory. First, there are far more percepts than we can capture within focal attention at any moment. Our brains process much of this perceptual information to some degree, without many of these percepts ever reaching conscious awareness. Critically, attentional control processes are required to select a subset of these stimuli to be represented within consciousness. Similarly, there are far more memories and ideas represented internally than we can be currently aware of and many of these may be at least partially active at any given time. Thus, interactions with internal representations also must be governed by attentional mechanisms that select which memories enter awareness. Second, studies of conscious perception have relied heavily upon subjective reports. It is, in fact, critical for the study of consciousness as these are the most direct measure of consciousness we have. Establishing a mapping between neuronal activity and subjective

Author's personal copy

reports allows us to move toward physiological measures of consciousness that do not require subjective reports. Third, another critical tool utilized in the study of consciousness involves studying the same stimulus under different types of awareness, so that the constant features of the stimulus remain the same and only the phenomenal conscious awareness changes. These last two points have been influential in directing the research described here toward an explicit investigation of the conscious regulation of awareness in memory, as will be described later. Lastly, studies of conscious perception have emphasized the importance of identifying neural substrates involved in awareness. By identifying which brain regions underlie awareness, we move closer to understanding how consciousness arises. These insights and methods developed in research on perception are reflected in recent work on the control of mnemonic awareness, both in the context of selective retrieval and retrieval stopping.

Selective Retrieval

Our goals often require us to modify the current contents of awareness by redirecting attention to new information relevant to the current task. In the perceptual domain, this requires the selection of a particular aspect of our rich sensory input, to the exclusion of other inputs that may compete for the limited capacity for awareness. In the memory domain, we often need to bring to awareness some particular event or fact that is important for our immediate purposes. In the latter case, the key mental operation for achieving this alteration of the contents of awareness is memory retrieval, and, in particular, selective retrieval (the mnemonic equivalent of selective attention). During retrieval, we use cues relevant to our goals to guide our search for the desired content. Typically, though, these cues are associated with other representations in memory, in addition to the specific content we seek. In fact, often these related memories spring to mind more readily than the desired target. For example, trying to remember what you had for dinner last Tuesday might bring to mind other recent dinners or other evening plans. Similarly attempts to remember a new phone number after moving are often thwarted by the retrieval of the old, no longer relevant phone number from the prior residence. A long history of memory research suggests that when multiple memories are associated with the same cue they compete for access to conscious awareness during retrieval. This type of interference poses a significant problem for the effort to direct consciousness to the desired memory; it requires some form of control to override the nondesired memories. Recent research suggests that this competition for conscious awareness is resolved by inhibitory control processes that suppress distracting memories, similar to the involvement of inhibition in achieving perceptual selective attention. This weakening of related memories allows retrieval of the target, but at the cost of impairing future recall of the nonretrieved competitors. Studying the conditions under which this form of memory impairment occurs thus provides an important behavioral window into the use of inhibitory control mechanisms to manage the redirection of consciousness to new memorial content.

The role of inhibitory control processes in achieving selective memory retrieval has been studied by Michael Anderson and colleagues using a procedure known as the retrieval practice paradigm (see **Figure 2**). In a typical study, subjects study

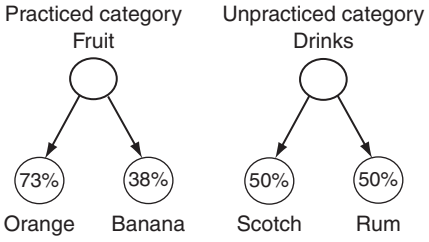


Figure 2 A standard categorical RIF study. Illustrated here are two items from each of two categories that subjects have studied (typically six items are studied from eight categories). In this example, subjects perform retrieval practice on 'Fruit-Orange,' but not on 'Fruit-Banana' (unpracticed competitor) or on any members from the 'Drinks' category (an unpracticed baseline category). The numbers show the percentage of items correctly recalled on the final cued-recall test. As shown here, retrieval practice facilitates recall of the practiced items relative to performance in baseline categories. RIF is reflected in the reduced recall of unpracticed members of the practiced category (Banana), relative to performance in baseline categories (Scotch and Rum).

Author's personal copy

lists of category-exemplar word pairs (e.g., Fruits-Orange, Fruits-Banana, Drinks-Scotch, Drinks-Vodka) and then practice retrieving some of these items from memory. Specifically, subjects practice half of the studied items from half of the studied categories (just 'Fruits-Orange' from the list above). The categories that are not practiced at all act as a control condition, so they can be used as an estimate of what baseline recall should be in the event that no items are practiced. After a delay, subjects are provided with all of the category cues again and they are asked to remember all of the items they studied earlier. As would be expected, the items that were repeatedly retrieved are recalled better on this final test than are the baseline items; thus, bringing past experiences into consciousness improves the ability to do this again, should the need arise. More interestingly, the unpracticed items from practiced categories (Fruits-Banana) are recalled more poorly than are baseline items. This intriguing finding suggests that when bringing a particular memory into consciousness, other memories that compete for the focus of awareness are inhibited. This finding, that retrieving specific targets induces forgetting of competing memories, has been termed retrieval-induced forgetting (RIF). RIF appears to be an instance in which the need to override automatic retrieval is solved by inhibitory control mechanisms that suppress the distracting content, supporting the effective redirection of consciousness.

Generality of Retrieval-Induced Forgetting

If RIF is a general consequence of attempts to control the redirection of consciousness to new memorial content, then it should be evident in any situation where we attempt to control retrieval. Indeed, RIF is not limited to category-exemplar pairs; rather, it appears to be a general phenomenon of importance to many everyday situations. Studies of learning, as would occur both in and outside the classroom, have shown that retrieving some facts about a topic impairs recall of other facts. For example, after reading textbook descriptions about two topics, being quizzed about some details of one topic causes forgetting of other details concerning that subject, but has no effect on the other studied topic. Similarly, recalling that 7 times 6 equals 42, makes it harder to remember that 7 times 9 equals 63. Beyond forgetting of simple facts, RIF also plays a role during the learning of a foreign language: when novice Spanish-speakers name pictures in Spanish, they subsequently experience difficulty generating the corresponding words in their native language.

Beyond these demonstrations of RIF in learning, this phenomenon is also of importance for other naturalistic situations. Studies of eyewitness memory show that interrogating subjects about specific details of a mock crime, in a manner consistent with actual police interviews, impairs memory for other noninterviewed details concerning the same crime. This suggests that the often numerous interviews performed by police and lawyers may have a profound influence on eyewitness memory for events. Recent work by Malcolm MacLeod and colleagues has even begun to explore whether RIF is responsible, at least partially, for the misinformation effect, originally pioneered by Elizabeth Loftus. Those classic studies showed that misleading information presented after an event, during an interview for example, is often mistakenly remembered as part of the original event. This new line of research suggests that the interview itself is actually critical for the misinformation effect, as the details that are weakened by RIF are the ones that are most vulnerable to subsequent misinformation.

The implications of RIF for social psychological phenomena have also been explored, with research showing that recalling some traits of a person makes it harder to remember that person's other personality traits. This phenomenon has even been extended to help understand how retrieving stereotypical information can lead one to forget stereotype-inconsistent, individuating features of an individual. Conversely, retrieving individuating features can cause forgetting of the stereotypical features. Recent studies also implicate RIF within autobiographical memory, suggesting that selective retrieval may play a role in shaping our own autobiographical history. The breadth of these findings indicate that selective retrieval is indeed involved in everyday cognition and has a profound impact not only on what is consciously remembered but also on what is excluded from awareness, both in the immediate term and in the long-term.

Author's personal copy

Evidence for Inhibition as the Mechanism that Produces RIF

To understand how consciousness is regulated during retrieval, it is critical to understand the precise mechanism by which selective retrieval is accomplished. Research suggests that RIF is produced by an inhibitory process that targets the competing memory trace itself. In contrast to this view, other researchers have proposed that noninhibitory mechanisms can produce the basic finding of RIF. For example, one noninhibitory account of RIF claims that during the final test phase, the items that were repeatedly practiced (Orange) are so dominant that they leap to awareness and block retrieval of the unpracticed competitors (Banana). By this account, the activation state of the competitor (Banana) has not changed. Instead, the rememberer perseverates on the practiced alternatives that have been strengthened, and this blocks retrieval of the competitor. Other noninhibitory accounts posit that the meaning of the retrieval cue is changed when it is used to practice a subset of its associates (one now thinks of citrus fruits when presented with 'Fruit' as a cue), rendering the cue useless as a means of retrieving noncitrus Fruits that had been previously studied. Importantly, these noninhibitory accounts all attribute the forgetting to some level (e.g., the cue, or the cue-target association) other than the forgotten item itself. By contrast, the inhibitory account makes the unique claim that the memory itself is being suppressed.

Several lines of evidence suggest, however, that selective retrieval inhibits competing items. First, RIF has been shown to be 'cue-independent,' as forgetting is observed even if the item is tested with a novel cue (e.g., Monkey-B_____ for Banana). According to the foregoing noninhibitory explanations, forgetting of competitors should only occur when the originally studied cue is used during the test. This should follow because the source of forgetting, according to those mechanisms, is specific to the original cue: the practiced response becomes so hyperaccessible given that cue, the meaning of the cue changes, or the link between the two items is unlearned. None of these explanations adequately explain why the competitor would be forgotten given an entirely novel cue. The inhibitory explanation, on the other hand, explicitly predicts that the item will be less accessible regardless of how it is tested. Building on this idea, in addition to being harder to recall, the competitors are also harder to recognize. Thus, it appears that the competitors have been reduced in activity.

Another property of RIF is that the forgetting suffered by competitors is not related to the strengthening of the target – a property known as 'strength-independence.' This means that strengthening target items, by itself does not cause forgetting of competitors. For example, if people are merely shown the category-exemplar pairs multiple times without having to retrieve them, similar strengthening is observed for the practiced items, but competitors are unimpaired. According to noninhibitory accounts, such as blocking, any form of strengthening should cause forgetting since strengthening of the practiced items is what leads them to block retrieval of the competitors. This finding demonstrates that RIF is 'retrieval-specific,' a property that is difficult to explain by most noninhibitory accounts. Lastly, the competitors that produce interference during the retrieval practice phase are inhibited more than ones that provide little interference. Thus, forgetting is 'interference-dependent,' suggesting that inhibition is engaged in response to interference from competing items. Again, this finding is difficult to explain by blocking, since the practiced items should block strong and weak competitors alike. Each of these properties strongly supports the claim that RIF is produced by inhibition, suggesting that inhibition plays a critical role in the way that conscious awareness is redirected to new traces in memory.

Neurobiological Basis of Selective Retrieval

Research has begun to explore the neurobiological underpinnings of inhibitory control during selective retrieval. As described earlier, inhibitory control is engaged during selective retrieval to prevent competitors from interfering with retrieval of the desired target. If successful, each successive retrieval practice should render competitors less interfering. Consistent with this, studies of RIF

Author's personal copy

using functional magnetic resonance imaging (fMRI) have revealed that performing retrieval practice engages both the ventrolateral prefrontal cortex (VLPFC) and the anterior cingulate cortex (ACC). The engagement of lateral PFC is consistent with prior research showing that the resolution of interference during selective retrieval from semantic memory also involves VLPFC. There is also neuropsychological evidence suggesting that patients with damage to lateral PFC experience difficulty resolving proactive interference. Similarly, the involvement of ACC is consistent with a broad range of findings that implicate that region in the detection of conflict. In the retrieval practice paradigm it appears that the competing memories trigger the need for top-down control, via the ACC, in order to resolve competition, which is then implemented by the engagement of inhibitory mechanisms mediated by the lateral prefrontal cortex. Supporting this idea, activity in these frontal regions declines across retrieval practice trials as the weakened competitors require less inhibitory control to be overridden. Critically, people who show greater decline in activity within these regions over trials show more memory inhibition. In addition, the people who show the highest degree of ACC activity during the initial trial are the ones most successful at suppressing. This suggests that subjects who experience the most competition initially are the ones who show the largest decline in lateral PFC activity across trials and the most forgetting.

Studies using EEG have suggested that selective retrieval is associated with a specific event-related potential (ERP) that indexes inhibitory control. In these studies, selective retrieval is contrasted with re-presentation of the studied stimuli, a condition that is known to not produce inhibition (as described earlier). Comparing activity in these two conditions yields an enhanced positive component in the selective retrieval condition over frontal electrode sites. Importantly, this enhanced activity is not due to strengthening of the practiced items, as these two conditions yield comparable facilitation; rather, this retrieval-specific component seems to index the inhibitory process that resolves interference. In fact, the magnitude of this component predicts how much forgetting the subject will experience. This finding suggests again that RIF is not produced by strengthening of practiced items, as is predicted by noninhibitory explanations. Rather, a specific inhibitory component is engaged to suppress the competitors rendering them less interfering. While localization of the source of ERP components is notoriously difficult, the frontal effect observed in these studies corresponds well with the fMRI findings on the importance of lateral PFC during selective retrieval. Thus, lateral PFC seems to subserve the selective filtering that controls which memories enter awareness, consistent with the view that response override mechanisms are central in the regulation of awareness.

Stopping retrieval

Whereas selective retrieval instigates the need to regulate which memories enter awareness, people generally do not form an explicit intention to down-regulate awareness in such situations. In other words, when attempting to remember some event or fact, remembering the target is the primary goal, and regulating interference occurs in support of this goal. However, sometimes stopping retrieval can itself be the person's primary goal. In these instances, we simply wish to stop retrieval from occurring. For example, when glimpsing an image of a loved one who has recently passed away we may marshal our efforts to stop painful thoughts of loss from coming into awareness. During the course of a typical workday we must frequently prevent distracting memories from involuntarily entering awareness and disrupting our current focus. These intrusive memories can be emotional in nature or simply consist of other activities or duties not related to our current goals. In extreme situations, survivors of abuse or combat veterans must exert this type of control in order to prevent traumatic memories from overwhelming their lives. In these instances, there is a clear conscious intention to prevent a memory from entering awareness. Clearly, such motivated retrieval stopping is a critical ability for daily mental functioning and for understanding how the content of consciousness is controlled.

Author's personal copy

This situation, overriding retrieval of an unwanted memory, has been studied using the Think/No-Think (TNT) paradigm. In these studies, subjects learn pairs of words (e.g., hug-rose, steam-train, broom-house) and are then asked to exert executive control over these memories. On some trials, in what is known as the 'Think' condition, people are asked to try to bring the target word to mind (when you see 'hug' think of 'rose'). For other 'No Think' trials, people are instructed to attend to the cue, but to willfully prevent the unwanted memory from entering consciousness (when you see 'steam' prevent the associated word from entering awareness). An additional set of cue words (e.g., broom) are not shown during this phase in order to provide a baseline measure of how accessible these pairs would be if they were neither retrieved nor suppressed after their initial learning. In the final phase a surprise memory test is given for all of the studied word pairs. Studying the memorial consequences of either thinking of a memory or excluding it from consciousness gives us an objective behavioral window into the mechanisms by which awareness is regulated.

People are, of course, better able to remember the words that they thought about compared to the baseline words, again affirming the idea that bringing memories into awareness improves one's ability to do so again later on. Evidence for inhibitory control arises from the finding that attempting to suppress awareness of response words during 'No Think' trials renders them harder to recall than the baseline items (see **Figure 3**). This below-baseline recall is present even when subjects are paid for correct answers or when they are misled into believing that the avoided words should be the easiest to remember. Thus, failure to recall does not reflect biases on the part of the person toward not reporting otherwise recallable 'No Think' items. Crucially, the impairment is not observed when the instructions are simply changed so that person only needs to withhold the vocal response, rather than avoid thinking about the memory. This indicates that the attempt to regulate conscious awareness is a critical and necessary component to produce this type of forgetting. Thus, regulating awareness is accomplished through inhibitory control of unwanted memories.

As was the case with RIF, forgetting in the TNT paradigm could be produced through noninhibitory mechanisms. For example, when presented with NT cues, subjects may simply generate alternative associations to distract themselves from

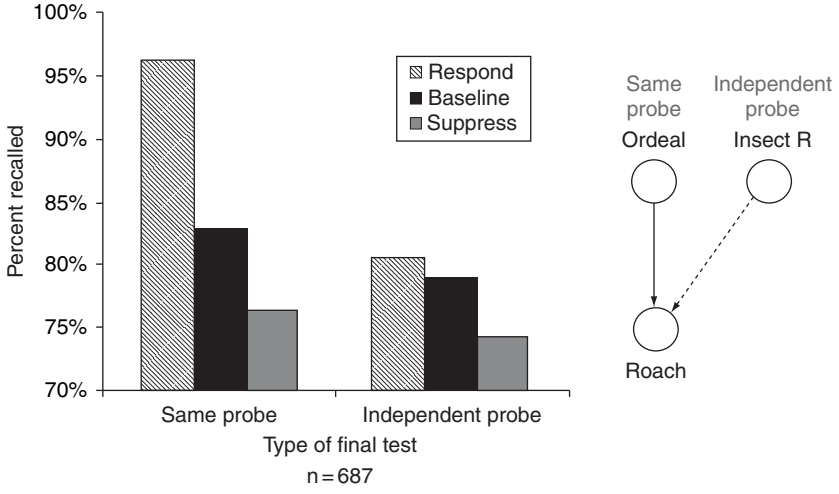


Figure 3 Final recall performance in the TNT procedure. The graph shows the percentage of items that subjects correctly recalled on the final test as a function of whether they tried to recall the item (Think), suppressed the item (No Think), or had no reminders to the item during the TNT phase (Baseline). The left side shows recall when tested with the originally trained retrieval cue (i.e., the Same Probe), whereas the right side shows recall when tested with a novel, extralist category cue (i.e., the Independent Probe). Reproduced from Anderson MC, and Levy BJ (2006) Encouraging the nascent cognitive neuroscience of repression. *Behavioral and Brain Sciences* 29: 511–513, with permission from Cambridge University Press.

Author's personal copy

thinking of the learned response. If true, this would mean that the unwanted memory was not so much intentionally pushed out of consciousness, as it was merely replaced by an alternate memory. If people accomplish the task by this form of thought substitution, then these alternative thoughts would become strengthened, potentially blocking retrieval of the response word during the final test phase. In order to rule out this blocking explanation, memory for the word pairs can also be tested with new categorical cues that subjects have not seen earlier in the experiment. Doing this yields similar forgetting, suggesting again that forgetting is 'cue-independent.' Other studies have also established that avoiding these memories also makes them harder to recognize, further confirming that these avoided memories have been inhibited.

Neurobiological Basis of Stopping Retrieval

Neuroimaging studies have found that attempts to stop retrieval are associated with increased activity within the lateral PFC, including both dorsolateral and ventrolateral regions (see **Figure 4**). Supporting the idea that lateral PFC is critical for suppression, individual differences in the magnitude of DLPFC activation are positively correlated with the amount of forgetting observed (see **Figure 5**). In addition to lateral PFC, suppression attempts activate a frontoparietal network of regions, including ACC, intraparietal sulcus, and the lateral premotor cortex, that is often observed in studies where subjects must prevent unwanted motor actions. The strong overlap between the network activated by retrieval stopping and motor stopping supports the claim that overriding is accomplished generally by a common system, regardless of whether the output being suppressed is motor or memorial in nature.

In addition to regions that are considered to be the source of the inhibitory signal, such as DLPFC, interest has also been taken in identifying the sites of inhibition: regions that are modulated by control. This goal dovetails with the approach described earlier within conscious perception, where an emphasis is placed on identifying candidate regions necessary for conscious awareness and on studying how these are influenced by tasks

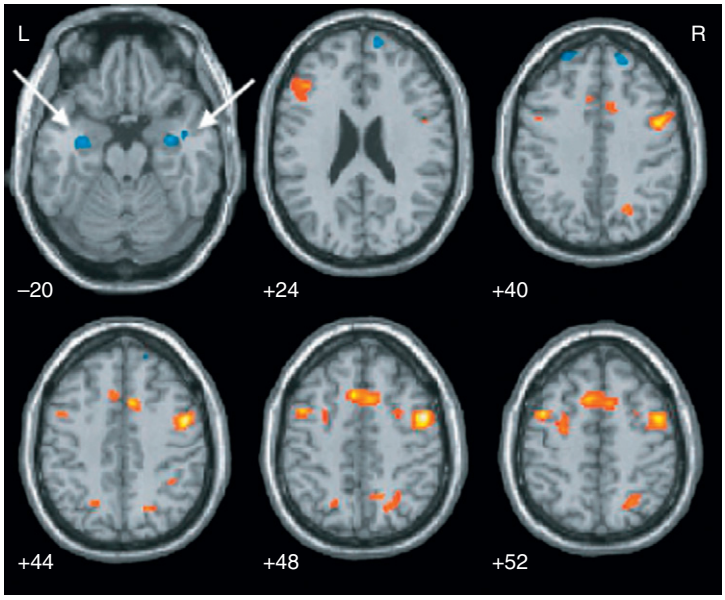
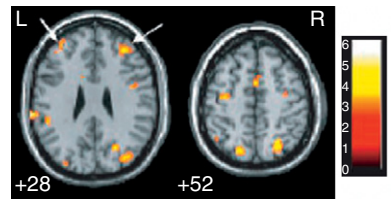
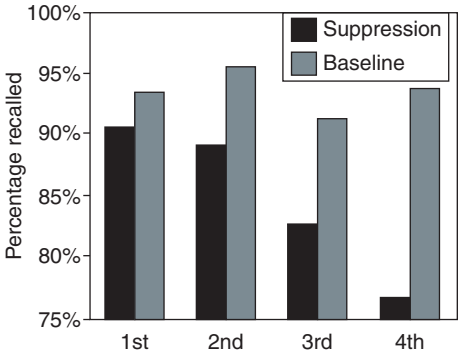


Figure 4 fMRI results from Anderson *et al*. Plotted above are the brain regions that significantly differed in activation between the Suppression trials and Respond trials during the TNT phase ($n = 24$). Areas in yellow were more active during Suppression trials than during Respond trials, whereas areas in blue were less active during Suppression ($p < 0.001$). The white arrows highlight the reduced hippocampal activation in the Suppression condition. From Anderson MC, Ochsner K, Kuhl B, Cooper J, Robertson E, Gabrieli SW, Glover G, and Gabrieli JDE (2004) Neural systems underlying the suppression of unwanted memories. *Science* 303: 232–235. Reprinted with permission from AAAS.

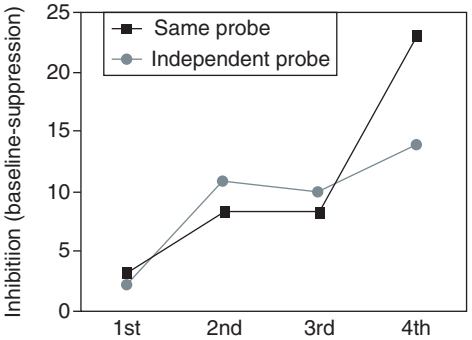
Author's personal copy



(a)



(b) DLPFC activation quartile (suppress > respond)



(c)

Figure 5 Successful recruitment of DLPFC predicts behavioral inhibition. (a) Shown here are the regions that correlate with the magnitude of the suppression effect observed on the final memory tests (the white arrows indicate the DLPFC). (b) Memory inhibition effects for four subject groups, differing in DLPFC activation. Subjects with greater DLPFC activity (on the right side) show reduced recall of No Think items, but do not differ from other subjects on their recall of Baseline items. (c) Magnitude of the suppression effect on both the same probe and independent probe tests for each DLPFC group. From Anderson MC, Ochsner K, Kuhl B, Cooper J, Robertson E, Gabrieli SW, Glover G, and Gabrieli JDE (2004) Neural systems underlying the suppression of unwanted memories. *Science* 303: 232–235. Reprinted with permission from AAAS.

that manipulate conscious awareness. Here, in contrast to motor stopping situations where the targets of inhibition would be motor systems, the likely candidate region would be one known to be involved in memory retrieval. It is well established that the medial temporal lobe, particularly the hippocampus, is involved in both the encoding of new experiences and the conscious retrieval of these memories later, especially for recently acquired memories. Moreover, it seems to play an especially critical role in memory for the type of richly detailed episodic memories that form the foundation of our conscious mental life. Interestingly, when subjects attempt to override retrieval, the hippocampus is down-regulated below the activity observed during trials where subjects perform retrieval. More impressively, it is also down-regulated below the level of activity observed during baseline periods where the subject has no task to perform other than to stare passively at a fixation cross. Thus, attempts to regulate conscious awareness by suppressing unwanted memories result in reduced hippocampal activation, produced by inhibitory control processes mediated by the frontoparietal regions described above. These findings support the view that response override systems mediated by the lateral prefrontal cortex can be targeted at structures involved in memory to prevent retrieval, consistent with the response override hypothesis.

Electrophysiological studies have also arrived at similar conclusions. Attempting to prevent an unwanted memory from entering awareness is associated with an early ERP component – the No-Think N2 – that arises over frontal sites and resembles, in topography and timing, the motor N2 typically observed in motor response suppression tasks such as the stop-signal and go/no-go paradigms. Importantly, when subjects are specifically instructed to suppress the unwanted memory directly whenever it comes to mind – and not simply generate distracting thoughts – the magnitude of this component predicts later memory inhibition effects, as would be expected based on the response override hypothesis of memory regulation. In addition to this early component, electrophysiological studies have also observed a late left parietal component that was specific to learned Think items. The timing and topography of this component are consistent with the widely studied parietal episodic memory (EM) effect, which has been linked to the subjective experience of consciously recollecting a past event. Crucially, this component is entirely absent during the No Think trials. Indeed, the EM

Author's personal copy

component was reduced in magnitude to the level observed for word pairs that were never learned. Thus, suppression completely eliminated this late retrieval-related component, suggesting that executive control can stop conscious recollection very effectively. Thus, ERP evidence is consistent with the model generated from fMRI studies suggesting that inhibitory processes mediated by lateral PFC accomplish control by actively downregulating the hippocampus and, as a consequence, conscious recollection.

Stopping Retrieval as a Laboratory Model of Freudian Suppression

The TNT paradigm provides a useful model for studying the psychological control process that Sigmund Freud referred to as suppression. According to Freud, suppression involved willfully preventing an idea from entering conscious awareness. This is precisely what subjects are asked to do in the TNT task, which suggests that results from those studies may be relevant to clinical issues related to suppression. Importantly, this new research provides empirical evidence about an issue that has largely been treated as scientifically intractable. While suppression has been difficult to study empirically, these new studies indicate that engaging suppression has a clear influence on subsequent access to these avoided memories.

Caution is necessary, however, in applying these laboratory studies to issues of clinical importance. The fact that a subject can forget a neutral word is not evidence that people can inhibit rich, episodic memories of traumatic experiences, as claimed by Freud and others. Recent studies have made inroads in establishing the viability of this paradigm for studying more naturalistic forgetting. In particular, evidence suggests that more complex memories of emotionally arousing events can also be inhibited. The original demonstrations of forgetting have now been replicated with emotionally negative memories even when the stimuli were naturalistic photographs (e.g., the scene of a major car accident), stimuli that are likely to be more vivid and, one might expect, harder to forget. In these studies, forgetting is typically as robust, if not more so, for negative stimuli as compared to neutral stimuli. Thus, even complex, vivid, emotional memories can be inhibited. This suggests that this psychological process has relevance for real-life memory suppression, which likely involves negatively charged memories. More research will be needed to further establish the ecological validity of this process. It is important to note, however, that the studies discussed here do not speak to the existence of unconscious repression, where unwanted thoughts are automatically pushed out of awareness. Indeed, the processes studied in this work are deliberate and intentional.

Individual Differences in Stopping Retrieval

There is considerable variability in how effective individuals are at recruiting inhibition to control which memories enter consciousness. Some people show dramatic forgetting of the avoided memories, while others actually seem to remember these items better despite their attempts to keep them out of mind. Much of this variability appears to be due to individual differences in executive control abilities. As described earlier, overriding retrieval of an unwanted memory is a specific example of a response override task, which is widely regarded as involving executive control. Therefore, it seems reasonable that variations in this ability would predict how successfully people inhibit unwanted memories. In support of this conclusion, neuroimaging studies have shown that the subjects who most strongly recruit dorsolateral PFC, a region thought to be critical for response override ability, show the most inhibition on the final test (see **Figure 5**). Similarly, complex working memory span tasks can be used as a proxy for executive control abilities, as these tasks require considerable executive control. Subjects load items into working memory and then respond to intervening items, which interfere with the maintained information and disrupt rehearsal. People with a high working memory span show large inhibition effects in the TNT procedure, whereas people with low working memory spans show facilitation of these avoided memories. Additionally, memory control ability differs across various populations known to vary in executive control. For example, older adults, who suffer a disproportionate loss of executive control abilities relative to other cognitive functions, show

Author's personal copy

difficulty inhibiting in the TNT paradigm. On the basis of this evidence, it seems that differences in memory suppression are at least partially attributable to differences in executive control more broadly. If so, these findings provide further support for a linkage between general response override mechanisms and the regulation of memorial awareness, and further point to clear individual differences in how effectively people regulate awareness.

Identifying conscious awareness of intrusive memories

The earlier sections describe a framework for understanding how inhibitory processes are engaged in order to maintain control over which memories enter awareness. Recent work has examined more explicitly how this type of control is related to consciousness. As described earlier, research on conscious perception has emphasized the use of subjective reports to distinguish between different conscious experiences of a stimulus. The same general approach can be taken within memory by having subjects report when a memory enters awareness. Recent research using the TNT paradigm has begun to employ subjective reports in order to address how conscious awareness of unwanted memories relates to the observed memory impairment. As in typical TNT studies, subjects learn word pairs and then later practice either thinking of the associated response or preventing it from coming to mind. The difference in these studies is that after each trial during the TNT phase, subjects make a subjective rating about whether or not they thought of the response word during the previous trial. Specifically, subjects report whether they 'never' thought of the response, thought of it 'briefly,' or thought about it 'often' during the time the cue word was on the screen. This procedure provides a means of distinguishing between trials where the subject is successful at suppressing awareness of the unwanted memory (i.e., they report never thinking about it) and trials where the unwanted memory 'intrudes' into conscious awareness, even if this awareness is fleeting.

This type of binary distinction may at first seem too gross, but, in essence, that is exactly what the study of consciousness requires. While the activation level of a given memory can be conceived of as a continuous measure, the distinction made in research on consciousness is between representations that have either passed threshold and entered awareness or that have not. It is this transition from a below-threshold to an above-threshold representation that is critical for identifying neural correlates of consciousness. Prior studies of consciousness in memory have generally lacked a proper way to identify when a memory transitions from being unconscious to entering conscious awareness and it has been difficult to induce a situation where a subject can monitor for their occurrence and be likely to actually experience a specific memory entering awareness.

As described earlier, forgetting in the TNT paradigm appears dependent on subjects attempting to exclude a memory from conscious awareness. Using subjective reports, however, provides a direct measure of the regulation of awareness. Using this approach has provided further evidence that attempts to regulate awareness produce behavioral inhibition on the final test. In these studies people report frequent intrusions initially as they struggle to successfully inhibit the unwanted memories. However, with practice, subjects down-regulate the frequency of intrusions, so that with practice intrusions become quite rare. Critically, people who are best able to down-regulate the frequency of intrusions are the same ones that show memory impairment on the final test. Thus the ability to regulate awareness and overcome intrusions with practice predicts the likelihood of forgetting those items on the final test, directly linking inhibitory control with the regulation of conscious awareness.

Neuroimaging evidence also supports a strong coupling between the down-regulation of intrusions and successful inhibition. In particular, overlapping regions within DLPFC predict both measures of inhibitory control ability – the down-regulation of conscious intrusions and inhibition as measured on the final memory test. Thus, these two measures show similar variability across individuals and, furthermore, this variability is produced by common neural substrates. Thus we have strong evidence suggesting that attempts to prevent a memory from entering awareness inhibit the avoided memory, making it less intrusive on subsequent trials and less memorable later even after suppression attempts have ceased.

Author's personal copy

Fascinatingly, when a person experiences a conscious intrusion of the unwanted memory, increased activation is observed in the lateral parietal cortex – a region that has previously been implicated in reflexive orienting to abrupt onsets in the perceptual world. In these studies, there is no abrupt perceptual onset, but the intrusion of the unwanted memory can be considered to be an abrupt internal onset that draws attention toward this newly activated memory. Thus, internally oriented attention can involuntarily focus attention on an abrupt onset in memory just as startling events in the perceptual environment can capture attention. Critically, this same region is not activated during Think trials, so it does not reflect retrieval itself or the representation of the memory within awareness. Rather, it reflects involuntary, reflexive retrieval that occurs when a memory pops into mind without an intention to retrieve it. This suggests that a common brain region may be engaged when a sudden perceptual event captures attention (and thus shifts awareness) as when a sudden mnemonic event diverts the focus of mnemonic awareness. This underscores potentially important commonalities in systems that challenge the effective control of perceptual and mnemonic awareness.

Controlling access to working memory

Another excellent example of common principles governing the regulation of perceptual and mnemonic awareness comes from research on the role of inhibitory control in regulating which aspects of perception gain access to working memory. In a task devised by Adam Gazzaley, subjects view pictures of faces and scenes and are asked to either attend to the faces, attend to the scenes, or to passively view the stimuli without attending specifically to either stimulus type. During each trial, subjects view a series of faces and scenes and then after a brief retention interval are asked to judge whether a test image was in the previous set. Neuroimaging studies using this task have focused on activity within the cortical regions that are specialized for processing each specific type of stimuli: the fusiform gyrus for face stimuli and the parahippocampal gyrus for scene stimuli. These studies find that attending to a stimulus class enhances activity within that region (i.e., attending to faces engages the fusiform gyrus) relative to when the same stimulus is passively viewed. In contrast, when subjects are asked to ignore a stimulus class, activity within the region devoted to processing that stimulus type is less active than when the stimuli are passively viewed. This suggests that ignoring a specific category of perceptual stimuli reduces activity within the brain region known to represent that stimulus. This clearly supports the idea that inhibitory control can be engaged in a top-down fashion to modulate representations. In this paradigm, older adults have difficulty with inhibiting processing of the currently irrelevant stimulus, while showing no impairment at content-selective enhancement. Like the TNT task, people in this task also regulate conscious awareness by selective attention mechanisms. This control appears to be implemented by the downward modulation of activation in the cortical region involved in processing the ignored stimulus, similar to the hippocampal modulation observed during suppression of episodic memories. Thus, inhibitory control may gate access to consciousness for both external, perceptual information and internal, memorial information.

Studies investigating thought suppression

A parallel body of research examines the ability to suppress unwanted thoughts, as opposed to episodic experiences. In these studies, pioneered by Daniel Wegner, people are asked to keep a specific thought (e.g., a white bear) out of mind for several minutes. During this delay, people are typically asked to continuously speak their thoughts aloud, with no specific instructions about what they should think about. They are further instructed that during this time they should monitor awareness for the presence of the unwanted thought and indicate whenever it comes to mind (e.g., by pressing a button or ringing a bell). Afterward, people engage in another think-aloud period where they are free to think any thoughts they wish. The typical finding from this paradigm is that the people who are asked to avoid a specific thought during the prior phase are more likely to think of the avoided thought. Furthermore, people often have difficulty keeping the thought out of mind during the suppression phase. From these results,

Author's personal copy

Wegner and colleagues concluded that attempting to avoid an unwanted thought results in it being 'ironically' more accessible later and, therefore, thought suppression is ultimately a futile endeavor. This paradigm has proven to be profitable in clinical research, particularly in relation to obsessive– compulsive disorder, depression, and anxiety.

These findings, however, appear to be at odds with the conclusions drawn from studies using the TNT procedure, namely, that people can successfully suppress unwanted memories. While these seem to be contradictory findings, it is also possible that they represent two different situations where people attempt to control awareness, but with differing results. While people are quite unsuccessful at suppressing in the White Bear studies, this does not mean that people lack the ability to regulate conscious awareness. As the TNT paradigm demonstrates, they can be effective at engaging inhibitory control to prevent an unwanted memory from entering consciousness under the right circumstances. Thus, both paradigms appear to capture situations where thought suppression is employed in naturalistic settings. Further research is needed to specify what factors determine success at regulating awareness.

Conclusion

Much of our conscious experience is driven by perceptual stimuli in the environment and by relatively automatic retrieval processes that respond to those environmental cues. Such automatic retrieval often enables us to retrieve appropriate behavior and interact with the world in an effortless manner. Many situations, however, require us to exert control over memory in order to behave flexibly. Moreover, many hallmarks of conscious awareness, such as our ability to adapt, to change our minds, and to think creatively require us to control retrieval and dictate which memories enter awareness.

In this article, we reviewed evidence for the idea that people exercise control of the contents of mnemonic awareness by engaging executive control processes that have developed in service of behavioral regulation. In particular, we suggest that controlling mnemonic awareness, at its most basic level, involves controlling the retrieval process, which may profitably be viewed as a special case of the broader problem of response override. The parallels between motor response override and mnemonic override may be observed at both the functional level and neurobiological levels: at the functional level, two main functions that engage inhibitory control over motor actions – selection and stopping – also engage inhibitory control in memory retrieval; at the neurobiological level, a common region for response override can be observed in the lateral prefrontal cortical regions that subserve motor inhibition and memory inhibition. A key difference, however, concerns the neural regions targeted by inhibitory control; for motor inhibition, motor cortical structures are affected, whereas for memory inhibition, brain systems involved in conscious recollection of the past (the hippocampus) or conscious perception of the present (the fusiform face area or the parahippocampal place area) are downregulated. Thus, evidence from the behavioral and the neural level point to the existence of mechanisms that actively diminish processing of that which we wish to exclude from awareness – mechanisms whose behavioral footprints may be observed in the later forgetting of those memories. If correct, this view suggests the intriguing principle that controlling the type of content we allow to enter awareness is a matter of controlling internal cognitive actions (like retrieval) that generate that content – an ability that is grounded in fundamental processes that have evolved in service of controlling what we do (and do not do) in the world around us.

See also: Consciousness and Memory in Amnesia; Memory: Errors, Constructive Processes, and Conscious Retrieval.

Suggested Readings

Anderson MC (2003) Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language* 49: 415–445.

Anderson MC and Green C (2001) Suppressing unwanted memories by executive control. *Nature* 410: 366–369.

Anderson MC and Levy BJ (2006) Encouraging the nascent cognitive neuroscience of repression. *Behavioral and Brain Sciences* 29: 511–513.

Anderson MC, Ochsner K, Kuhl B, *et al.* (2004) Neural systems underlying the suppression of unwanted memories. *Science* 303: 232–235.

Author's personal copy

Anderson MC and Spellman BA (1995) On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review* 102: 68–100.

Anderson MC and Weaver C (2008) Inhibitory control over action and memory. In: Squire L (ed.) *The Encyclopedia of Neuroscience.* http://www.elsevierdirect.com/brochures/ens/index.html. Elsevier.

Bergström ZM, Velmans M, De-Fockert J, and Richardson-Klavehn A (2007) ERP evidence for successful voluntary avoidance of conscious recollection. *Brain Research* 1151: 119–133.

Depue BE, Banich MT, and Curran T (2006) Suppression of emotional and non-emotional content in memory: Effects of repetition on cognitive control. *Psychological Science* 17: 441–447.

Depue BE, Curran T, and Banich MT (2007) Prefrontal regions orchestrate suppression of emotional memories via a two-phase process. *Science* 37: 215–219.

Gazzaley A, Cooney JW, McEvoy K, Knight RT, and D'Esposito M (2005) Top-down enhancement and suppression of the magnitude and speed of neural activity. *Journal of Cognitive Neuroscience* 17: 507–517.

Kuhl BA, Dudukovic NM, Kahn I, and Wagner AD (2007) Decreased demands on cognitive control following memory suppression reveal benefits of forgetting. *Nature Neuroscience* 10: 908–914.

Levy BJ and Anderson MC (2002) Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences* 6: 299–305.

Levy BJ and Anderson MC (2008) Individual differences in the suppression of unwanted memories: The executive deficit hypothesis. *Acta Psychologica* 127: 623–635.

Wenzlaff RM and Wegner DM (2000) Thought suppression. *Annual Review of Psychology* 51: 59.

Biographical Sketch

Dr. Benjamin Levy is a postdoctoral fellow at Stanford University. He received his PhD from the University of Oregon and his BS from the University of California, Davis.

Author's personal copy

Dr. Michael Anderson is a professor of cognitive neuroscience at the School of Psychology at the University of St Andrews, Scotland, where he is the director of the Memory Control Laboratory. He has previously served as a professor of cognitive neuroscience at the University of Oregon, has been a visiting scholar at Stanford University and the University of California, Berkeley, and has served on the editorial boards of the *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *Memory and Cognition*, and *Perspectives in Psychological Science*. He received his PhD in cognitive psychology and neuroscience from the University of California, Los Angeles.