Theoretical Issues in Inhibition: Insights From Research on Human Memory

Michael C. Anderson and Benjamin J. Levy

Several years ago, Benjamin J. Levy married the woman he had been dating for several years. On numerous occasions after the wedding, he mistakenly referred to his new wife as his “girlfriend.” During the years of courtship, he had become so accustomed to referring to her in this manner that the word just popped to mind when thinking of her. The recent marriage made that label problematic, however. His frequent references to his “girlfriend” left others to assume he had taken a mistress and was simply very candid about the fact. Although amusing at first, the humor of these mistakes was quickly lost on his wife. With effort and attention, however, he was able to override this well-practiced response and refer to her as his “wife.” This example illustrates a simple point: Often people are victims of an overly effective retrieval system.

Although this example is amusing, people often encounter reminders of things that they are much more motivated to avoid thinking about, such as a painful breakup, a particularly odious task, or a loved one who has passed away. In these circumstances, they may exert effort to prevent these memories from occupying their thoughts. How can one prevent intrusive, unwanted memories from coming to mind? Research from our laboratory suggests that people recruit executive control to override the retrieval of unwanted memories and that this cognitive act induces a lasting suppression of the unwanted memories, making them more difficult to recall later even when they want to return to them (for a review, see M. C. Anderson, 2003).

A core theme of this research is that people’s ability to control unwanted memories is directly analogous to their ability to control their overt behavior, a topic broadly studied in cognitive psychology and in cognitive neuroscience. In fact, avoiding unwanted memories resembles a classic situation that requires executive control, often referred to as response override. One must stop a prepotent response to a stimulus, either because that response must be withheld or because an alternative more weakly learned response to that stimulus is desired. The ability to stop prepotent responses is critical to the flexible control of behavior, whether the response is that of a baseball player stopping the
swing of the bat when the pitch is a ball, a husband avoiding embarrassment, or a person preventing an unpleasant memory from coming to mind. Without this ability, people would be slaves to their habits and reflexes.

Our suggestion, one in keeping with other research on executive control, is that people accomplish this control through inhibition of the prepotent response. When people are presented with reminders of unwanted memories, activation spreads from the cue to the traces stored in memory. If the dominant trace is not currently desired, either because the rememberer wishes to avoid thinking about it or because a more weakly associated trace is sought, then inhibition can be engaged to weaken the dominant memory, enabling him or her to stop retrieval or to have selective control over what is retrieved. We argue that these inhibitory mechanisms also have a lasting effect, leading to memory impairments for the avoided memories even later when people want to retrieve them.

In this chapter, we discuss evidence that uniquely supports this inhibitory control perspective in two response override situations in memory: the desire to stop retrieval and the need to selectively retrieve a memory. Then we describe a theoretical problem in the measurement of inhibition that is typically ignored in studies of inhibition. This issue, the correlated costs and benefits (CCB) problem, has extremely important consequences for the ability to adequately test theoretical models of inhibition and particularly for the ability to test inhibitory deficit theories concerning different populations of subjects. We argue that to make a strong claim in any study about the presence or absence of inhibition or about variations in the magnitude of inhibition as a function of condition or population, it is necessary to include an independent probe of the impaired items' accessibility. Without the independent probe, measurements of inhibition will suffer the CCB problem, precluding principled predictions about how behavioral effects should vary according to inhibitory theories. This problem is not at all unique to memory research, and we provide examples from other research domains discussed in this volume. By drawing attention to this issue, we hope to steer the field toward experiments that isolate the involvement of inhibitory mechanisms and that prevent unnecessary confusion and theoretical controversy in the literature on inhibitory processes.

**Selective Retrieval**

When recalling a memory, the desired trace is rarely the only memory related to the cues guiding retrieval. In fact, often the nontarget memories are more strongly associated with the cue than is the currently desired trace. In such situations, the associated traces compete for access to conscious awareness, necessitating some process to enable selective retrieval. In our framework, selective retrieval represents a paradigmatic case of response override, where one must select a weaker memory in the face of interference from one or more prepotent competitors. If stopping prepotent responses engages inhibition, the same mechanisms might also be engaged to stop prepotent memories from coming to mind, promoting selective retrieval. If so, perhaps inhibition will
induce long-lasting memory impairment for the competitors. Thus, the very act of remembering should cause forgetting of related memories.

This prediction has been tested in a procedure that we refer to as the “retrieval practice paradigm.” Subjects study lists of category–exemplar word pairs (e.g., fruits–orange, fruits–banana, drinks–Scotch) and then practice retrieving half of the studied items from half of the categories (fruits–or——), each of which is practiced three times. After a delay, subjects are asked to recall all of the previously studied exemplars. Not surprisingly, practiced items (orange) are facilitated relative to items from nonpracticed categories (Scotch), which serve as a baseline for how well items are recalled with no practice. More interesting, nonpracticed items from practiced categories (banana) are recalled less often than baseline items (see Figure 5.1). Thus, retrieving some items during retrieval practice leads to worse memory for related items on the final test. According to the inhibitory control hypothesis, these items are inappropriately activated during the retrieval practice phase and then are inhibited to promote successful retrieval of the desired response (orange). This phenomenon, known as retrieval-induced forgetting (RIF), has been replicated many times using a broad array of stimuli (for reviews, see M. C. Anderson, 2003; Levy & Anderson, 2002).

A brief survey of recent demonstrations of RIF illustrates the striking breadth of this phenomenon. When bilingual persons retrieve the nonnative word for a concept, the native language phonology of that word is inhibited (Levy, McVeigh, Marful, & Anderson, 2007). Retrieving multiplication facts
about a particular number (e.g., 7 × 7 = 49) impairs the ability to report unpracticed multiplications with that number (e.g., that 7 × 8 = 56; Phenix & Campbell, 2004). And retrieving individuating characteristics of a person from a stereotyped group causes inhibition of the target’s stereotypic traits (Dunn & Spellman, 2003). Furthermore, recent research demonstrates that whereas most RIF studies induce competition by episodic exposure to the competitors, semantically related items can be inhibited even when not studied (Johnson & Anderson, 2004; Starns & Hicks, 2004). Thus, a wealth of experimental results suggest that RIF is a general phenomenon that results in impairment whenever unwanted items intrude during retrieval.

One of the goals of this volume, however, is to consider whether inhibition is actually involved in “inhibitory” effects such as RIF. The basic finding of RIF described in the preceding paragraphs is compatible with several noninhibitory mechanisms and by itself is not compelling evidence that inhibition is involved. For example, the practiced items may be so strengthened by retrieval practice that they interfere during the final test, effectively blocking the subject from coming up with the correct response. This type of retrieval competition has a long history in formal models of memory retrieval (e.g., J. R. Anderson, 1983; Raaijmakers & Shiffrin, 1981), where the probability of recalling an item is predicted by the relative strength of the association between the cue and the target compared with the strength of the association between the cue and all the competitors. Other noninhibitory mechanisms can explain RIF as well. For example, subjects may simply unlearn the connection between fruits and banana, or practicing some exemplars may bias the representation of that category toward the practiced items (retrieving orange might bias the cue fruits toward citrus fruits). According to all of these noninhibitory explanations, one need not claim that any change is occurring to the item itself (for a review of these noninhibitory sources of impairment, see M. C. Anderson & Bjork, 1994). However, this is a core claim of the inhibitory control perspective: The actual unwanted item itself is being made less accessible.

Several properties of RIF uniquely support the involvement of inhibition. First, the inhibitory control perspective (M. C. Anderson, 2003) makes the unique claim that RIF should be observed regardless of which cue is used to test the memory. In other words, forgetting should be cue independent and should generalize to novel cues in the test phase rather than being specific to those used to perform retrieval practice (M. C. Anderson & Spellman, 1995). For example, recall of banana should be impaired not only when it is tested with the studied category (fruits) but also when it is tested with a novel, independent retrieval cue (monkey—b). Such cue-independent forgetting is difficult for noninhibitory mechanisms to explain, because they predict that impairment should be specific to the cues used during retrieval practice. For example, associative blocking cannot explain this type of impairment because the new independent cue (monkey) is unrelated to the strengthened exemplars (e.g., orange); thus, there is no reason why they should block retrieval of banana.

Cue-independent forgetting has been observed many times (e.g., M. C. Anderson & Bell, 2001; M. C. Anderson, Green, & McCulloch, 2000; M. C. Anderson & Spellman, 1995; Camp, Pecher, & Schmidt, 2005; MacLeod & Saunders, 2005; Radvansky, 1999), suggesting that the competing item itself
is inhibited. Other researchers have expanded on the notion of cue independence by arguing that if competitors are truly inhibited, then memory impairments should also occur on memory tests other than recall. RIF has now been found both on recognition tests (Hicks & Starns, 2004; Starns & Hicks, 2004) and on an implicit lexical decision test (Veling & van Knippenberg, 2004). The use of independent probes is the best tool available at present for distinguishing between interference and inhibition as mechanisms in long-term memory. Yet many researchers design studies of inhibition in memory without using independent probes. The problem is that one cannot isolate the role of inhibition in producing the memory impairment without a way of excluding contributions of noninhibitory factors. Lack of attention to this fact has led, we argue, to misleading conclusions about the nature of RIF and how it varies in different populations. We will return to this concept in greater detail later in this chapter.

Second, a hallmark of associative interference accounts of forgetting is that memory impairment should occur whenever competitors are strengthened. Several findings, however, have shown that strengthening practiced items does not always lead to impaired recall of competitors. For example, repeated study exposures to practiced items result in strengthening comparable to that produced by retrieval practice, yet this type of strengthening produces no forgetting of competing items provided that output interference at testing is controlled (M. C. Anderson, Bjork, & Bjork, 2000; Bäuml, 1996, 1997, 2002; Ciranni & Shimamura, 1999). If strengthening can occur without producing impairment of the related items, it suggests that retrieval practice, not strengthening, is what causes the impairment. Thus, RIF appears to be both recall specific and strength independent.

Finally, the amount of impairment depends on the extent to which the competitors interfere, a property referred to as interference dependence. For example, M. C. Anderson, Bjork, and Bjork (1994) observed more RIF when the competitors were high-frequency (e.g., orange) than lower frequency (e.g., kiwi) category members. Similarly, practicing retrieval of the meanings of asymmetric homographs results in significant impairment for the dominant meaning but not for the subordinate meaning (Shivde & Anderson, 2001). According to the associative blocking account, there is no reason to expect that the relative frequency of the unpracticed item should influence the degree to which the practiced item blocks retrieval during the final test. Each of these properties of RIF strongly implicates inhibition and provides serious challenges for any noninhibitory explanation. Taken together, these results support the inhibitory control perspective that selective retrieval is a special case of response override that results in lasting inhibition of the avoided memories.

**Stopping Retrieval**

The need to control behavior is not limited to selecting a nondominant response; it also sometimes is necessary to stop a behavior from occurring. A baseball player who makes the last-minute decision not to swing at a ball outside the strike zone is choosing to override a prepared response to swing. Sometimes a rememberer’s goal is to do something remarkably similar: He or she is not...
Figure 5.2. Final recall performance in the think/no-think (TNT) procedure. The graph shows the percentage of items correctly recalled on the final test as a function of whether subjects tried to recall the item (Respond), suppressed the item (Suppress), or had no reminders to the item (Baseline) during the TNT phase. The left side shows recall to the originally trained retrieval cue (i.e., the same probe); the right side shows recall to a novel, extralist category cue (i.e., the independent probe). The numbers are from a meta-analysis of 687 subjects run in the TNT paradigm in our lab.

trying to select a competing memory so much as trying simply to stop retrieval itself. When presented with a reminder of something upsetting (e.g., a photograph of a loved one who has recently died), one might desire to put those thoughts out of mind. Other times, one may simply want to stop a thought from popping into mind because one is trying to stay focused on a different thought. In fact, this type of control of internal thought is crucial for preventing one’s mind from wandering. Can inhibitory mechanisms be engaged to serve these goals?

Evidence for the ability to stop memory retrieval comes from the think/no think (TNT) paradigm developed by M. C. Anderson and Green (2001). Subjects learned cue–target pairs (e.g., ordeal–roach). They then were presented with the cue (e.g., ordeal) and asked either to think of the associated target word (e.g., roach) or to prevent that associated memory from coming to mind. After seeing each type of either cue as many as 16 times, subjects were asked to recall all of the previously studied target memories. When provided with the same cue (ordeal) and asked to recall the correct target memory (roach), subjects had more difficulty if they had previously avoided thinking about that memory than if they had encountered no reminders of it during the TNT phase (see Figure 5.2). This finding is counterintuitive, because one would expect that repeated reminders to an item should make it more accessible, not less.
As was the case with RIF, the basic TNT suppression effect is compatible with noninhibitory accounts. For example, subjects might generate alternative thoughts during the TNT phase for each of the cues that require suppression (e.g., some ordeal they experienced personally). Then, on the final test, these diversionary thoughts might come to mind, blocking retrieval of the target item. M. C. Anderson and Green (2001) ruled out this possibility by demonstrating memory impairment for the avoided items using novel, extralist retrieval cues (e.g., insect--r----- for roach) as independent probes. Associative interference from diversionary thoughts cannot explain independent-probe impairment, because the diversionary thoughts were generated unrelated to the originally studied cue. In their final experiment, M. C. Anderson and Green showed that simply asking subjects to avoid saying the response word does not result in the suppression effect, indicating that the attempt to prevent the unwanted memory from coming to mind is critical to producing the impairment. Taken together, these results indicate that the memory impairment arises from an inhibitory control mechanism that prevents the unwanted memory from entering awareness.

TNT suppression has now been replicated several times (M. C. Anderson et al., 2004; M. C. Anderson & Kuhl, 2007; M. C. Anderson, Reinholz, Kuhl, & Mayr, 2007; Bell & Anderson, 2007; Depue, Banich, & Curran, 2006; Hertel & Calcatera, 2005 [aided condition]; Hertel & Gerstle, 2003 [nondysphoric subjects on positive word pairs]; Hotta & Kawaguchi, 2006; Joormann, Hertel, Brozovich, & Gotlib, 2005; Kawaguchi & Hotta, 2006; Wessel, Wetzel, Jelicic, & Merckelbach, 2005; but for failure to find below-baseline impairment, see Bulevich, Roediger, Balota, & Butler, in press). Although M. C. Anderson and Green (2001) used neutral word pairs, several recent studies have investigated whether the valence of the unwanted thought influences suppression success. Three of these studies have found that suppression of negative stimuli leads to comparable (nonsignificantly increased) inhibition relative to either neutral stimuli (M. C. Anderson et al., 2007; Depue et al., 2006) or positive stimuli (Joormann et al., 2005). However, Hertel and Gerstle (2003) found that nondepressed subjects were unable to suppress negative adjective–noun pairs. It is unclear what produced the different pattern in Hertel and Gerstle’s study, but taken together these results suggest that the valence of the unwanted memory itself is not the primary determinant of success at suppression.

The memory impairment in TNT studies suggests that inhibitory control mechanisms may be recruited to prevent unwanted memories from coming to mind. This finding has obvious implications for when people wish to avoid persistent, intrusive thoughts. One approach might be simply to avoid reminders of those memories, but this strategy is often not practical. The foregoing results suggest that when presented with inescapable retrieval cues in more naturalistic settings, repeatedly avoiding memories may cause long-lasting impairments in recalling those memories. Thus, the TNT paradigm may be a useful laboratory model of the voluntary form of repression (suppression) proposed by Freud (1915/1963). If so, the TNT paradigm can be used as a tool for exploring clinical issues related to motivated forgetting.
Neural Substrates of Stopping Retrieval

Earlier in this chapter we outlined how stopping retrieval is analogous to stopping a motor response. A key difference between these two situations is that motor suppression is observable with the naked eye (e.g., one can see when a baseball player checks a swing), but there is no outwardly observable sign when someone chooses to stop a declarative memory. Despite the similarities between stopping actions and stopping thoughts outlined earlier, they could be accomplished in entirely different ways. One way of testing our idea about a fundamental similarity between these situations is by investigating whether these processes are produced by the same neural mechanisms. If stopping retrieval is really related to stopping a motor response, then a common underlying neural network should be involved in accomplishing both types of stopping.

Studies of motor response override have shown that a network of control-related regions, including the lateral prefrontal cortex, anterior cingulate cortex, lateral premotor cortex, and intraparietal sulcus, are recruited to stop motor responses (e.g., Garavan, Ross, Murphy, Roche, & Stein, 2002; Menon, Adleman, White, Glover, & Reiss, 2001). Thus, we predict that suppressing unwanted memories should also engage these regions. There should also be regions unique to memory control: those that are the target of suppression. Given the role of the hippocampus in conscious recollection of declarative memories (e.g., Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000; Squire, 1992) and the goal of suppressing conscious recollection in the TNT task, the hippocampus seems a likely region to target. M. C. Anderson et al. (2004) hypothesized that control-related regions (particularly lateral prefrontal regions) should be involved in disengaging hippocampal processes to prevent conscious recollection of the unwanted memories.

M. C. Anderson et al. (2004) used functional magnetic resonance imaging to identify the brain regions that support intentional memory suppression. Subjects were scanned while they participated in the TNT task. Again, subjects recalled significantly fewer suppression words than baseline words. M. C. Anderson et al. observed significantly more activation in control-related regions during the suppression trials than during the respond trials. Therefore, stopping retrieval is not simply a failure to engage retrieval processes: Instead, the heightened activation of these control regions during suppression trials suggests that subjects actively engage processes to prevent the unwanted memories from coming to mind. As predicted, the observed regions of increased activity overlapped highly with control-related regions involved in stopping motor responses: Common brain regions control stopping both unwanted memories and unwanted actions.

In addition to these control regions, M. C. Anderson et al. (2004) observed reduced hippocampal activity bilaterally on no-think relative to think trials. This difference could be attributed to increased activity during the think condition. However, the degree of hippocampal activity was correlated with behavioral memory inhibition, which is inconsistent with the difference on think and no-think trials being caused simply by heightened recollection in the think
condition. Instead, it suggests that subjects can strategically down-regulate the hippocampus to prevent conscious recollection.

These neuroimaging results suggest that subjects can prevent unwanted memories from coming to mind by the same neural mechanisms that are recruited to stop motor actions. Instead of targeting a motor response, these regions are recruited to suppress declarative memory representations. So despite the difficulty involved with observing the stopping of memory retrieval, there now is both behavioral and neuroanatomical evidence to support the idea that people can prevent unwanted memories from coming to mind. The findings thus help to specify a model of how motivated forgetting might occur.

The Correlated Costs and Benefits Problem

In addition to isolating the role of inhibition in memory retrieval, the foregoing studies also revealed key theoretical issues about how inhibition should be measured. In the second half of this chapter, we discuss an important issue that has wide implications for investigators interested in inhibition: the CCB problem. We argue that failure to attend to this issue has contributed to significant confusion in the literature, especially in work on individual differences in inhibitory function. Such confusion has hindered theoretical development and generated doubt about inhibition phenomena—doubt that may underlie some of the skepticism evinced by opponents of inhibitory theories. We offer a solution that has proved useful in research on memory inhibition and that could be adapted for other cognitive domains.

All investigators would agree that whatever cognitive mechanisms are proposed to explain a phenomenon ought to operate consistently: If a mechanism is engaged under one set of circumstances, it ought to be engaged in future situations that have similar circumstances. Surprisingly, this seemingly uncontroversial premise is routinely ignored in studies of inhibition. Studies that induce inhibition at Time T and then later measure the behavioral aftereffects of that inhibition at Time T + 1 frequently fail to consider the involvement of inhibition at the time of the measurement (T + 1), even though inhibition should contribute to performance during the measurement process itself. The CCB problem refers to the theoretically predicted difficulty in quantifying the amount of inhibition that has taken place at Time T by measuring inhibition using an assessment at Time T + 1 that itself requires the same inhibitory process.

The retrieval task depicted in Figure 5.3A illustrates this problem. In both situations, a cue is associated with two items in memory, and the task is to retrieve either Item A in Situation A or Item B in Situation B. According to the inhibitory control perspective, in Situation A if the cue activates both traces A and B, they should compete for access to awareness; thus, if Item B interferes with the retrieval of Item A, inhibition should suppress Item B, facilitating the retrieval of Item A. The lingering effects of inhibition on Item B should make it harder to recall on later occasions. The analogous pattern should hold
**Figure 5.3.** Illustration of the correlated costs and benefits problem. Panel A: In Situation A, a cue is associated with two responses. The inhibitory perspective predicts that practicing Item A in response to the cue should result in inhibition of Item B (provided that it interferes). Furthermore, in Situation B, practicing Item B should inhibit Item A, as Situations A and B are essentially equivalent. Panel B: These diagrams illustrate how Situations A and B apply to the retrieval practice paradigm. Practicing orange during the retrieval practice phase should inhibit banana. However, consistent application of the concept of inhibitory control at retrieval predicts that inhibition will also aid subjects in recalling banana at the time of test by inhibiting orange.

for Situation B. Inhibitory control should be involved in both of these equivalent retrieval situations.

Although this reasoning seems obvious, it has not been obvious to investigators studying the role of inhibition in memory retrieval. To illustrate this point, Figure 5.3B shows the different phases of the retrieval practice paradigm.
In the retrieval practice phase, subjects practice retrieving orange, suppressing the competitor banana and making it less accessible. The same items are on the final test, only now the goal is to retrieve the former competitor, which is now the target (e.g., banana). Clearly, consistency demands that the same inhibitory control mechanisms used to suppress banana in the retrieval practice phase ought to now be engaged to suppress orange, which has become the competitor. Indeed, inhibitory control should be more necessary during the test both because banana is inhibited (and harder to retrieve) and because orange is stronger from retrieval practice. Thus, the same inhibitory mechanisms that impaired banana during retrieval practice (Time T) should aid in the retrieval of banana by overcoming interference from the previously practiced item (orange) during the final test (Time T + 1). Importantly, the magnitude of both the cost and the benefit should be correlated, because they arise from the same underlying inhibition mechanism.

To see the impact of the CCB problem, RIF can serve as a model case. Suppose that researchers know, for each subject, how much of the RIF effect is produced by suppression during retrieval practice and how much is attributable to blocking on the final test. What would the predicted relation between these components and inhibitory control ability look like? How would they sum together to produce the aggregate RIF effect? Figure 5.4 plots the amount of RIF attributable to suppression, blocking, and both factors combined as a function of inhibitory control ability. Clearly, the suppression component should increase with inhibitory ability. So a patient with damage to the prefrontal cortex might be on the left side of this figure, unable to suppress competing items during retrieval practice, whereas a high-functioning young adult might be closer to the right side. If RIF were attributable solely to suppression induced during retrieval practice, then it clearly should increase with increasing inhibitory control ability.

Unfortunately, the CCB problem implies that the same inhibitory control ability that suppresses the competitor during retrieval practice (Time T) ought to assist in suppressing the practiced item during the final test (Time T + 1), reducing the blocking contribution to RIF. So as inhibitory control ability increases from left to right, the component of RIF attributable to blocking should diminish. Thus, on the left side of the figure, a hypothetical frontal patient, even one with absolutely no ability to suppress the competitor during retrieval practice, would be completely unable to suppress interference from the practiced items on the final test. To the extent that practiced items are strengthened, the frontal patient should experience exaggerated blocking from those strengthened items. Ignoring this fact would be quite curious given that perseverative behavior in memory tasks is largely what motivated the original inhibitory deficit hypothesis of frontal lobe dysfunction. By contrast, on the right side of the figure, the hypothetical young adult who effectively suppressed the competitor during practice would suffer little associative blocking on the final test. When the joint contributions of the costs (persisting suppression) and benefits (reduced blocking at T + 1) are summed, one can see that inhibitory control theory does not predict a clear relation between inhibitory control ability and the size of the RIF effect when subjects are tested with the same cue used to perform retrieval practice. Essentially, the costs and benefits should trade off, yielding
Figure 5.4. The correlated costs and benefits problem leads to unclear predictions on shared probe tests. Shown is the amount of retrieval-induced forgetting (RIF; numbers shown on the y-axis are arbitrary units and not amounts of specific data) observed on the final test that can be attributed to various causes as a function of inhibitory control ability. The solid black line shows the amount of RIF attributable to persisting suppression that was induced during retrieval practice, whereas the dotted line shows the amount attributable to blocking by practiced competitors on the final test. Subjects with more inhibitory control should suffer more persisting suppression, whereas subjects with less inhibitory control should experience more blocking. The dashed line represents the total effect, considering the combined influences of these factors. This total changes little with inhibitory control ability, illustrating how same-probe data may not clearly show differences between individuals with differing levels of inhibitory control ability.

Effects that appear the same behaviorally but that are generated by different underlying chains of events.

Although this analysis may seem straightforward, several RIF studies testing inhibitory deficit hypotheses have not considered the role of suppression on the final test. For example, studies by Moulin et al. (2002) and Conway and Fthenaki (2003) both measured RIF with the standard category cued recall test (Figure 5.5). Both studies contrasted a patient population thought to have a deficit in inhibitory control (either Alzheimer’s patients or frontal-lobe-damaged patients) against a control population. In both studies, RIF was observed in the deficit and control populations. If RIF is a pure measure of inhibition, then either these populations do not have deficits in inhibitory control or the inhibition measured by RIF is different from that supported by the prefrontal cortex.

These and other authors have used evidence such as this to argue that RIF reflects an automatic form of inhibition not mediated by the prefrontal cortex and not related to attentional control (Zellner & Bäuml, 2005). Yet if one invokes the inhibition process consistently when it is required, there should
be no clear relation between RIF and the deficit status of these populations as measured by the same-probe test. Put simply, the Alzheimer’s disease and frontal-lobe-damaged patients, even if they were completely lacking in inhibitory control, would be expected to show significant forgetting by virtue of their compromised ability to resolve interference at the time of the final test. This finding would make their RIF effect appear similar to control subjects with intact inhibitory function despite entirely different underlying causes.

Failing to consider the CCB problem can lead to theoretically significant conclusions about the nature of RIF or, alternatively, about the populations being studied—conclusions that, based on a more thorough consideration of CCB analysis, are not warranted and potentially very misleading. Evidence of this potential to mislead is provided in the next section, along with a solution to the problem.

**A Solution With Examples**

The key to solving the CCB problem is to redesign the final assessment of inhibition to eliminate or reduce the benefits of inhibition. Although illustrated with memory inhibition, the principle applies generally to all paradigms in which inhibition is measured. This approach requires devising a final T+1 assessment that minimizes heightened interference from practiced items, allowing the test to better reflect suppression that took place in the retrieval practice phase (Time T). Fortunately, this test already exists: the independent-probe method. Testing the inhibited item with a novel cue circumvents associative interference that might otherwise arise from the stronger practiced
items, permitting a measure of inhibition at Time T + 1 not contaminated by blocking. If this test truly minimizes blocking, then there should be less need to engage inhibitory control on the final test, reducing the benefits of that process. Thus, with the independent-probe method, the measured RIF should better reflect inhibition caused by retrieval practice (Time T), making performance on this test more sensitive to individual differences in inhibition capacity than performance on the same-probe test, in which costs and benefits are both expressed.

Although subjects in the TNT paradigm do not explicitly strengthen a competitor during the TNT phase (unlike in the retrieval practice paradigm), they may generate diversionary thoughts that become associated with the cue word, introducing a blocking component to the same-probe effect (see Hertel & Calcaterra, 2005). If so, independent-probe impairment should better track individual differences in inhibitory function. This predicted finding was obtained in three series of TNT experiments that examined the size of the forgetting effect as a function of working memory capacity (Bell & Anderson, 2007), age (M. C. Anderson et al., 2007), and trauma history (M. C. Anderson & Kuhl, 2007), all factors that might moderate how well someone can suppress distracting memories. In each series, inhibition was assessed both with the same test cue initially used to induce inhibition (same-probe condition) and with a novel cue (independent-probe condition), permitting greater isolation of the persisting inhibition component.

Figure 5.6 presents the results from the same-probe and the independent-probe tests. The data have been simplified by collapsing the individual experiments in each series, though the pattern was replicated in each individual study. These graphs reveal a pattern highly consistent with the CCB problem. In each case, the amount of impairment on the same-probe test is, at best, weakly consistent with an inhibitory deficit. However, in all cases, the impairment does not vary reliably with subject group. In contrast, in all three series, the independent-probe test revealed significant differences in inhibition as a function of subject group. In fact, the differences in inhibition are quite striking on the independent-probe measure, and the three-way interaction of inhibition, subject group, and test type is significant in each case. Thus, whether an inhibitory deficit was observed varied reliably as a function of test type in each series, exactly as predicted by the CCB analysis.

These findings illustrate the dangers of the CCB problem. If we had based our conclusions about the relative inhibitory abilities of these different populations on performance on the same-probe test, as did Moulin et al. (2002) and Conway and Fthenaki (2003), we might have concluded that these groups did not differ in inhibitory function and that inhibition is unrelated to working memory capacity or executive control. Yet we know from our neuroimaging work (M. C. Anderson et al., 2004) that the TNT task engages frontal cortical regions known to be involved in executive function. The independent-probe measure, by reducing the potential benefits of inhibition at testing, provided much more striking evidence for variation in inhibition, in line with prior theoretical expectations. Thus, this method provides a solution to the CCB problem that can prevent theoretical confusion.
Figure 5.6. The utility of the independent-probe method. Each graph plots difference scores (baseline-suppression recall) to indicate the amount of memory impairment. The columns show data from studies investigating the relation between inhibition and working memory, age, and trauma. We predicted greater inhibition with greater working memory capacity, younger age, and higher level of experience with trauma. All three studies showed the predicted pattern numerically in the same-probe condition, but the differences in inhibition across groups were far from reliable. In contrast, the independent-probe data clearly and significantly showed the predicted differences between the groups, and the three-way interaction of group by inhibition by test type was significant in each series of studies.

Generality of the Problem

Although we have used RIF and the TNT paradigm to illustrate the CCB problem, this problem applies broadly to a variety of paradigms in memory, attention, and language. In this section, we provide several examples of the problem in these settings so that investigators in those areas can better recognize how this problem applies to their research.

Episodic and Semantic Memory Paradigms

Like RIF, several memory paradigms have the characteristic that strengthening some associates of a cue impairs later recall of other associates of the cue, such as retroactive and proactive interference in the A-B, A-C design; output interference for items associated with a common experimental or contextual cue; and part-set cuing impairment. The standard practice of measuring impair-
ment by a shared cue at testing reflects the joint costs and benefits of inhibition. Semantic fluency tasks, in which subjects must produce as many exemplars of a category as they can within 1 minute, also suffer from this problem: As one generates new exemplars, other, yet-to-be-listed exemplars in semantic memory ought to be suppressed, yielding a cost to total fluency performance; on the other hand, that same inhibition process should reduce interference from already-generated exemplars, reducing perseveration and yielding a benefit to performance. Ironically, total fluency performance has often been taken as a measure of executive control ability, and in particular inhibitory control ability, when in fact no straightforward relation between total fluency and inhibitory function is predicted by inhibitory control theories.

Executive Control Paradigms

In the phenomenon of backward inhibition in task-set switching, subjects are presented with a stimulus on which they must perform a task as rapidly as possible (e.g., adding a pair of visually presented numbers). Other trials might cue subjects to perform a different task (e.g., subtraction) on the same class of stimuli. People generally are slower to perform the target task if they had just performed a different task, reflecting the time that it takes to switch task sets. It has been proposed that on switch trials, subjects engage inhibitory mechanisms to suppress the preceding task set (e.g., subtraction) so that a new set may be selected. For instance, Mayr and Keele (2000) found that if subjects performed Task A (e.g., subtraction) for one block of trials (Block 1) and then switched to Task B (e.g., addition) in the next block (Block 2), switching back to the switched-out-of task (e.g., addition) in the next block (Block 3) took significantly longer than switching to a new Task C (e.g., multiplication). Thus, there was some additional cost associated with having to switch back to a previously rejected set relative to a new set. Mayr and Keele interpreted this as evidence for inhibition in task-set switching and termed the phenomenon backward inhibition.

Backward inhibition has been used to study individual differences in inhibitory function in older adults (Mayr, 2001) and in frontal-lobe-damaged patients (Mayr, Diedrichsen, Ivry, & Keele, 2006) on the assumption that it provides a clean measure of the effects of inhibition. However, backward inhibition is subject to the CCB problem. On the one hand, the better subjects’ inhibitory control, the slower they should be to respond to the recently inhibited task set, because the effects of prior inhibition at Time T should persist to T + 1. On the other hand, when inhibition is measured on Block 3 (T + 1), subjects must also suppress interference from the recently engaged task set from Block 2 and retrieve the suppressed task set (evoked in part by the common stimulus—numbers). Thus, inhibitory control benefits the subject, helping to combat interference from the most recent set. Although it might seem that the experimental (Tasks A, B, and A) and control (Tasks A, B, and C) conditions are matched for the need to suppress Task B during the final test, this is not the case. In the experimental condition, the strength of Task B relative to the inhibited Task Set A should be greater than the strength of Task B relative to the new
Task Set C in the control condition, making inhibitory control more beneficial in the experimental condition. Thus, inhibition induces both costs and benefits that, when taken together, make it quite difficult to predict a clear relation between inhibitory function and the backward inhibition effect.

**Visual Selective Attention Paradigms**

Very similar arguments can be made about the negative priming phenomenon (for reviews, see Milliken, Joordens, Merikle, & Seiffert, 1998; Neill & Valdes, 1996; Tipper, 2001). In negative priming, the inhibition of a target on a preceding trial (Time T) is typically measured on a probe trial (Time T+1) that itself requires subjects to resolve competition. As in backward inhibition, this design provides the opportunity for inhibition to yield a greater measurement-epoch benefit in the experimental condition, in which the competition is between the suppressed target and a nonsuppressed distractor, compared with the control condition, in which the competition is between a nonsuppressed target and a nonsuppressed distractor. Again, the putative costs of inhibitory control are mixed with their benefits, yielding no clear predictions about how negative priming should vary as a function of inhibitory ability.

**Language Processing Paradigms**

The CCB problem also arises in language tasks that have been used to assess inhibition. One example is the role of inhibition in lexical ambiguity resolution (see Gorfein, 2001, for a collection of reviews). During sentence processing, interpreting a word that has multiple meanings requires that its contextually appropriate sense be accessed. It is generally thought that competition between the word's multiple meanings must be resolved, and this function has been attributed to inhibition by some investigators (Gernsbacher & Faust, 1991; Simpson & Kang, 1994). In these studies, subjects viewed trials composed of a briefly presented prime word followed by a probe letter string requiring a lexical decision. When the prime is a homograph (e.g., *bank*), reaction time to a related probe (e.g., *river*) is speeded. More interesting, when the following trial represents the same prime (*bank*) followed by a probe related to the previously rejected sense (e.g., *money*), probe lexical decision times are slower than if the first prime–probe trial had presented an entirely unrelated homograph (e.g., *arm shoulder*). The slowed response to the previously rejected sense has been taken as evidence that the meaning (*money*) was inhibited during the first prime–probe trial.

The assessment of inhibition in this paradigm again mixes the lingering effects of inhibition on the prime trial (Time T) with the need to overcome interference on the probe trial (Time T+1). When the probe trial re-presents the homograph cue word, the recently selected meaning grows very active in response, creating a situation analogous to associative blocking in RIF. Thus, processing of the probe on the second trial requires subjects not only to overcome the inhibition of the previously rejected sense but also to combat the heightened interference from the competing sense. Similar problems arise in numerous
other language processing tasks thought to involve inhibition, including anaphoric reference (Gernsbacher, 1989) and metaphor comprehension (Gernsbacher, Keysar, Robertson, & Werner, 2001). Thus, if one's goal is to use such behavioral effects to examine individual differences in inhibition or to study the theoretical conditions under which inhibition occurs, one needs a measure that does not conflate these sources of impairment (for examples of the independent-probe method in language processing, see Johnson & Anderson, 2004; Shivde & Anderson, 2001).

Conclusion

To the extent that inhibitory control has been assessed using measures that suffer from the CCB problem, the literature should be plagued with inconsistencies in the support of inhibitory deficit theories, generating reasonable doubt about the utility of these theories. We argue that such variability is not necessarily a sign of weakness of the theories but rather a sign of the poorly considered measures that have been used to assess inhibition. As a solution, we argue for the importance of developing testing methods for measuring inhibition that minimize the potential benefit of inhibition during the test epoch itself. By minimizing the benefits of inhibition, variation in inhibitory function can be measured more cleanly. On the basis of results from several studies, we argue that the independent-probe method provides one such example that has proved to be sensitive as a means of testing individual differences in this function. We believe that the logic underlying this method can be adapted to any domain in which inhibition might be studied.

The purpose of this volume is to review perspectives on the role of inhibition in cognition. On the one hand, many investigators have presupposed inhibition as part of a broader theory of cognition. The existence of inhibition seems both plausible and necessary: plausible, because many behavioral effects appear to arise from inhibition, and necessary, because some process for limiting the influence of distracting representations seems essential—a proposal supported by the widespread involvement of inhibition in the nervous system and by computational analyses indicating that inhibition is necessary for stability in neural networks (e.g., Easton & Gordon, 1984). On the other hand, skeptics of inhibition correctly emphasize that the mere presence of performance decrements does not by itself require such processes. Indeed, many behavioral phenomena labeled “inhibition” may be adequately explained without proposing an additional inhibition process. If so, parsimony dictates that the simpler theory be preferred, though it remains a point of debate whether such alternative noninhibitory theories are always simpler.

Our approach has been to develop clear standards to judge whether a behavioral deficit constitutes inhibition. In this chapter, we reviewed this work, which establishes functional properties of memory inhibition that uniquely favor the role of inhibitory mechanisms in memory retrieval. Collectively, these properties strongly support the view that inhibition overrides prepotent responses in memory. Thus, one of the key imperatives advocated by reasonable skeptics of inhibition—to go beyond the mere equating of a performance decre-
ment with inhibition—has been programatically addressed in the context of memory retrieval over the past decade. We do not imply that some noninhibitory mechanisms do not also contribute to apparent inhibitory effects in memory, only that inhibition clearly contributes to producing such effects.

In fact, the contributions of noninhibitory sources of impairment to apparent inhibitory phenomena motivated our discussion of the CCB problem. Put simply, measuring the behavioral consequences of an earlier act of inhibition on some target representation ought itself to engage the very same inhibitory processes to successfully process that inhibited representation. If so, the putative cost of inhibition ought to be mitigated by its putative benefit at the time of testing. When inhibitory functioning is impaired (because of either an inhibitory deficit or experimental manipulations), the costs should be reduced, but so too should the benefits, opening the door for noninhibitory components to contribute to an effect. Thus, to effectively measure inhibition, a test should minimize the benefits of inhibition. We have argued that the independent-probe method provides such a test and that its logic can be applied in any domain in which inhibition is studied. This proposal has been validated in three TNT studies that consistently showed that the independent-probe method provides a significantly more reliable index of inhibitory control deficits than do tests that suffer from the CCB problem.

We strongly emphasize that the CCB problem is not limited to memory retrieval: It applies broadly in domains such as executive control, visual attention, and language processing. We hope that recognition of this problem will encourage investigators to develop theoretically targeted methods of testing inhibition theories, enabling the field to advance beyond an indefinite cycle of assertion and skepticism that plagues the study of such processes. We believe that sufficient evidence exists for inhibitory processes in memory retrieval and that such processes are likely to contribute broadly to cognitive function, but reasonable doubt will remain as long as the field neglects the issue of measurement. Thus, our challenge to investigators of executive control, visual attention, and language processing is to develop solutions to the CCB problem to determine whether the properties of inhibition, identified in memory, also apply in other cognitive domains.

References


