Measuring Memory Reactivation With Functional MRI: Implications for Psychological Theory
Benjamin J. Levy and Anthony D. Wagner
Perspectives on Psychological Science 2013 8: 72
DOI: 10.1177/1745691612469031

The online version of this article can be found at:
http://pps.sagepub.com/content/8/1/72
Perspectives on Psychological Science
8(1) 72–78
© The Author(s) 2013
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1745691612469031
http://pps.sagepub.com

Cues in the environment frequently remind us of past experiences, which allows the past to inform current behavior. Memory reactivation can occur with or without the intention to remember and with or without awareness of the connection between the cues and the memories they activate. Reactivated memories frequently inform our current goals, but in some instances they may be irrelevant to, and may even distract us from, our goals. As such, memory reactivation plays a critical role in guiding behavior, as it influences the knowledge upon which decisions and actions are based.

Although memory reactivation is central to many psychological theories, it is often difficult to directly observe reactivation in behavior. Researchers can explicitly ask participants whether they are currently reminded of past experiences, but individuals may not always be aware when reactivation occurs. Moreover, asking participants to reflect on whether they are remembering may influence the frequency and type of reminders they experience. As an alternative approach, researchers may indirectly probe memory and then infer when reactivation has occurred by observing effects on behavior (e.g., M. C. Anderson & McCulloch, 1999). In many instances, however, the link between the observed behavior and memory reactivation is ambiguous.

Neuroimaging methods, such as functional MRI (fMRI), offer a novel approach to measuring memory reactivation that complements behavioral indices (e.g., Kahn, Davachi, & Wagner, 2004; Nyberg, Habib, McIntosh, & Tulving, 2000; Wheeler, Petersen, & Buckner, 2000; for review, see Danker & Anderson, 2010; Mitchell & Johnson, 2009). In some instances, fMRI can even provide powerful evidence of reminding independent of behavior, and fMRI can yield measures of memory reactivation at the individual trial level (e.g., Chadwick, Hassabis, Weiskopf, & Maguire, 2010; Johnson, McDuff, Rugg, & Norman, 2005; for review, see Rissman & Wagner, 2012). In this article, we review recent fMRI studies that demonstrate how neural measures of memory reactivation can be used to inform psychological theories about the representations and processes underlying generalization, forgetting, and consolidation (e.g., Kuhl, Rissman, Chun, & Wagner, 2011; Kuhl, Shah, DuBrow, Corresponding Author:
Benjamin J. Levy, Department of Psychology, Stanford University, Jordan Hall, Building 420, MC 2130, Stanford, CA 94305
E-mail: benlevy@stanford.edu
Integrative Encoding as a Mechanism for Generalizing Across Experiences

One situation where memory influences behavior is when we encounter a novel choice that requires that we generalize knowledge across multiple prior experiences to determine the appropriate decision or action. From one prominent view, such generalization occurs at the time of choice and is based on the retrieval of multiple discrete episodic memories, from which we logically infer the relationship between the stimuli and the decision that should be made (e.g., Eichenbaum, 2000; Kamaran & McClelland, in press). By contrast, an alternative mechanism for generalization is integrative encoding that occurs during the acquisition of the original experiences before the novel choice probe is ever encountered (e.g., Eichenbaum, 2000; O’Reilly & Rudy, 2000; Shohamy & Wagner, 2008). According to this account, we are reminded of similar prior events during encoding of subsequent related events, and when this reminding occurs, an integrated mnemonic representation is formed that spans past and present episodes. This integrated memory directly encodes the relations between stimuli from temporally distinct events, and this knowledge is retrieved later when the novel choice is encountered. A key question is whether memory-based inferences that support generalization can occur because of encoding of integrated cross-event representations or whether generalization exclusively occurs through the flexible expression of separable event memories at test.

Shohamy and Wagner (2008) provided support for integrative encoding, reporting fMRI evidence of reactivation during the learning phase of an acquired equivalence task (e.g., Myers et al., 2003). Participants initially learned associations between stimuli (e.g., A and X, B and X, A and Y) through trial-and-error feedback (see Fig. 1a). Because some of the associations overlapped (i.e., both A and B predict X), it is possible that participants reactivated other related associations while encoding a specific association (i.e., while encoding A–Y they may recall the A–X and B–X associations). To the extent that such reactivation occurs during learning, participants can form a representation that integrates information from past experiences with that present in the immediate environment. This integrated representation could then support subsequent decisions when the participant is faced with a novel choice that can be solved through generalization (i.e., generalizing the A–Y association to B–Y, because of A and B’s equivalence due to their common association with X).

Consistent with the integrative encoding account, Shohamy and Wagner (2008) observed that activity in the hippocampus increased across the phase when participants learned the original associations (see Fig. 1b). Because prior data indicate that repeated encoding of the same association results in decreasing hippocampal activation (e.g., Kohler, Danckert, Gati, & Menon, 2005; Zeineh, Engel, Thompson, & Bookheimer, 2003), whereas reactivation of past associations results in increased hippocampal activation (e.g., Eldridge, Knowlton, Furmanski, Bookheimer, & Engel, 2000), Shohamy and Wagner interpreted this hippocampal increase as a marker of memory reactivation during learning. Critically, the magnitude of this increase across learning predicted participants’ later generalization performance, suggesting that reactivation during learning supported subsequent memory-based inferences. These findings suggest that during subsequent events associative mechanisms in the hippocampus reactivate prior experiences allowing for knowledge integration across distinct episodes; the resulting integrative representations can support later generalization.

Of course, evidence of reactivation during encoding does not preclude a role for retrieval-based inference in generalization (e.g., Greene, Gross, Elsinger, & Rao, 2006; Heckers, Zalesak, Weiss, Ditman, & Titone, 2004; Preston, Shrager, Dudukovic, & Gabrieli, 2004). Using a single-shot learning paradigm, Zeithamova and Preston (2010) provided further evidence for integrative encoding—hippocampal activity during encoding of an event that overlapped with a prior experience (i.e., during B–C encoding, after having studied A–B) predicted later transfer during a novel choice (i.e., choosing C over D when presented with A)—and evidence consistent with a contribution of retrieval-based inference—participants who generalized well demonstrated greater test phase activation during generalization probes (A–C) than during directly learned probes (A–B). Thus, although computational modeling work suggests that reactivating event content from distinct memories at test is sufficient to explain generalization (Kamaran & McClelland, 2012), extant neuroimaging data clearly suggest that the ability to generalize from past experiences depends upon both integrative encoding arising from reactivation of past experiences during subsequent learning and retrieval-based inference. Moreover, to the extent that learning accompanies retrieval-based inferences, the mechanisms that give rise to such inferences might also provide a computational account of integrative encoding when combined with learning.

The Relationship Between Activation of Competing Memories and Forgetting

Sometimes memory reactivation can be counterproductive, particularly when an association that comes to mind is not relevant to current goals. In these situations, individuals suffer interference, which is thought to be a major cause of forgetting (M. C. Anderson, 2003; Crowder, 1976). According to many psychological theories, forgetting due to interference is a retrieval phenomenon in which stronger memories tend to come to mind readily and block retrieval of weaker memories (e.g., J. R. Anderson, 1983; McGeoch, 1942; Raajmakers & Shiffrin, 1981). An alternative theory, however, posits that interference elicits a secondary process (or set of processes)
Fig. 1. Evidence for integrative encoding from Shohamy and Wagner (2008). A: Behavioral paradigm for the acquired equivalence task. Participants learned face–scene associations through trial-and-error feedback (upper left). Critically, some cues shared a common associate during the learning phase (lower left). During a subsequent test (upper right), participants were tested on the trained pairs and on generalization trials. Generalization trials could be correctly responded to by way of two different mechanisms (lower right): during test, participants may retrieve multiple distinct trained associations and draw inferences from them about what response should be made; alternatively, during learning, prior associations may be reactivated and then integrated into a single cross-event representation that is then available to be retrieved when the generalization trials are encountered at test. B: Neuroimaging evidence for integrative encoding. Activity in the right and left hippocampus (top) increased across the learning phase. Across participants, the degree to which activation increased from early to late learning was predictive of performance (% correct) on the generalization trials (middle). A median split of the participants revealed that “good” generalizers showed increased activity from the early to late learning phase, whereas “poor” generalizers did not (bottom).
that counteract(s) interference by inhibiting competing memories (e.g., M. C. Anderson, 2003; Melton & Irwin, 1940; Osgood, 1949). Thus, from this perspective, it is the process of overcoming interference during a prior retrieval attempt that renders competing memories harder to recall later.

Considerable behavioral and neuroimaging evidence has been gathered in favor of the inhibitory account of forgetting (for reviews, see M. C. Anderson, 2003; Bäuml, Pastötter, & Hanslmayr, 2010; Levy & Anderson, 2002; Storm, 2011), but inhibitory forgetting effects are often relatively small and are not always replicated (e.g., Bulevich, Roediger, Balota, & Butler, 2006; Perfect et al., 2004; Williams & Zacks, 2001). Although some of this variability may be attributable to variations in cognitive control abilities across individuals (Levy & Anderson, 2008), there are also boundary conditions on when inhibition will occur (for reviews, see M. C. Anderson, 2003; Levy & Anderson, 2002). In particular, participants in interference paradigms are thought to sometimes reactivate “competing” memories along with the target memory, potentially forming an integrated representation that spans both events (M. C. Anderson & McCulloch, 1999). This integration would protect the “competing” memory, thus reducing the need for inhibitory control and reducing forgetting on later memory tests. Neural measures of reactivation may be able to provide evidence about when this type of integration occurs.

Although fMRI data clearly indicate that competitive retrieval is associated with greater engagement of cognitive control processes (e.g., Danker, Gunn, & Anderson, 2008; Kuhl, Dudukovic, Kahn, & Wagner, 2007; Wimber et al., 2008), the initial studies did not provide clear evidence about when competing memories were reactivated. Kuhl et al. (2011) recently reported evidence demonstrating this type of reactivation using distributed pattern analyses of fMRI data. In the earlier section on generalization, we discussed how researchers used the magnitude of activation within the hippocampus as a marker of reactivation, whereas in Kuhl et al. (2011) reactivation was assessed through the reinstatement of distributed patterns of activity across many voxels within the ventral temporal cortex. Specifically, in a retroactive interference paradigm in which participants learned overlapping associations (A–B and A–C pairs), pattern analyses revealed evidence that the competing associate was reactivated during attempts to retrieve the target associate (i.e., B was reactivated during attempts to retrieve the A–C association). Critically, greater evidence for reactivation of the earlier B memory was correlated with better memory for those items on a final posttest, which is consistent with the hypothesis that reactivation led to the formation of an integrated representation that protected earlier memories from forgetting. Similarly, Kuhl et al. (2010) found that hippocampally mediated reactivation of cortical patterns during the encoding of associations that overlapped with previously learned associations predicted better subsequent memory for the earlier associations (i.e., greater hippocampal activity during A–C encoding predicted better subsequent memory for the A–B association). Both of these findings suggest that reactivation can lead to integrated cross-event representations that protect earlier memories from forgetting.

Prior behavioral evidence suggests that encoding, unlike retrieval, does not induce competition and therefore does not trigger the need for interference-resolution processes (e.g., Bäuml, 1996; Bäuml & Hartinger, 2002; Ciranni & Shimamura, 1999). Therefore, it is not surprising that reactivation during encoding trials did not induce forgetting. It is less clear, however, whether reactivation during retrieval should always be beneficial for later remembering, as observed by Kuhl et al. (2011). It seems likely that, at least in some retrieval situations, reactivating competing memories should trigger the need for inhibitory control and thus greater forgetting. The relationship between reactivation and inhibition, however, is not necessarily straightforward (for discussions, see M. C. Anderson & Levy, 2011; Bäuml et al., 2010; Norman, Newman, & Detre, 2007), and this relationship may be obscured by integrative encoding mechanisms. Nevertheless, neural measures that can provide continuous indices of reactivation on a trial-by-trial basis hold great promise for observing competitive dynamics during retrieval and promise to provide critical leverage on testing theories about the relationship between reactivation during retrieval attempts and subsequent forgetting (see also Öztekin & Badre, 2011).

The Role of Memory Reactivation in Consolidation

Memory reactivation is thought to be critical for systems consolidation, wherein a memory that initially depends on hippocampal representation is gradually encoded in the neocortex (e.g., Squire & Alvarez, 1995). Research in rodents suggests that recent memories are reactivated during sleep (e.g., Foster & Wilson, 2006; Lee & Wilson, 2002; Wilson & McNaughton, 1994), and this reactivation is thought to reflect the hippocampus playing back recent experiences to the neocortex, hypothesized to be a critical process in consolidation (e.g., McClelland, McNaughton, & O’Reilly, 1995). Recent neuroimaging work in humans has sought to experimentally manipulate memory reactivation during sleep, examining effects at the neural and behavioral levels. For example, Rasch et al. (2007) had participants study object–location associations while being exposed to a specific odor, and they then reexposed the participants to this odor during slow-wave sleep. This reexposure led to increased hippocampal activity and improved subsequent memory for the object–location associations (relative to exposure to a control odor), suggesting that the odor triggered reactivation of associated events, putatively facilitating their consolidation and rendering them more readily recallable. Using scalp electroencephalography (EEG), Rudoy, Voss, Westerberg, and Paller (2009) expanded on this finding by using unique sound cues that were each associated with a particular object–location association. Exposure to a subset of the sounds during sleep led to better subsequent memory for the cued associations, and EEG responses to the
cues during sleep significantly predicted subsequent memory performance.

These two studies support the possibility that memory reactivation can be induced by environmental cues during sleep, thus contributing to memory consolidation and enhanced post-sleep memory. Neither of these studies, however, provides compelling evidence that directly links clear evidence for reactivation with improved subsequent memory. For example, these measures might simply reflect further cue encoding, which then allows the cue to act as a more potent reminder during the postsleep memory test. Future fMRI studies can borrow from the approaches described above to determine if and when offline memory reactivation occurs, be it during sleep or during awake, resting states (Tambini, Ketz, & Davelchi, 2010). Although the evidence in this section may be more speculative and less certain, this is clearly another area in which fMRI measures of reactivation hold considerable promise for informing psychological theory.

Conclusion

We have described how fMRI methods, based on either activity from regions of interest or from distributed patterns of activity, can be used to track memory reactivation as it occurs, even on a trial-by-trial basis. In this manner, fMRI provides unique leverage on psychological theory that is not afforded by behavioral measures, which do not allow investigators to observe when a past experience has been covertly reactivated. Moving forward, we anticipate that neuroimaging will be informative for numerous open questions, as many psychological theories invoke covert retrieval as a key mechanism. For example, environmental cues are thought to activate negative stereotype-consistent information in groups vulnerable to social identity threat (e.g., Murphy, Steele, & Gross, 2007). Similarly, cues are thought to trigger retrieval of threat-related information in depression, anxiety, and posttraumatic stress disorder (e.g., Mathews & MacLeod, 2005; McNally, Kaspi, Riemann, & Zeitlin, 1990). Researchers in these and other areas may be able to leverage fMRI measures of reactivation to determine when memories are retrieved, which specific cues trigger the strongest reactivation, and the psychological and behavioral consequences of such reactivation. In our view, fMRI measures of covert and overt reactivation hold great promise for testing psychological theories, illustrating the complementary power of functional neuroimaging as a fundamental tool for psychological science.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This research was supported by National Institute of Mental Health Grants R01-MH080309, R01-MH076932, and F32-MH079648.

References


