Cognitive Neuroscience of Declarative and Nondeclarative Memory

Paul J. Reber
Department of Psychology
Northwestern University
2029 Sheridan Road
Evanston, IL 60208
847 457 1624
The study of memory within cognitive neuroscience is an attempt to synthesize an account of both the fundamental mnemonic component processes and the neural basis of these processes. This approach generally takes a very broad definition of memory to include any changes within the brain that reflect long-term storage of new information. This broad definition encompasses traditional definitions of memory, the conscious acquisition and recollection of facts and events, but also brings a range of additional phenomena into the memory domain, e.g. skill learning, perceptual learning, eyeblink conditioning, and even plasticity in the aplysia gill withdrawal reflex (e.g., Milner, Squire & Kandel 1998). All of these types of memory depend on changes (plasticity) within identifiable brain regions and networks. From this perspective, the fundamental questions of memory are to identify the mechanisms of plasticity, the representations of stored information, and the methods by which these changes influence subsequent behavior.

The most well-studied and complex memory system of the brain is based on the medial temporal lobe (MTL), a neural system that contains the hippocampus and surrounding cortical areas. It is this system which, when lesioned in the patient H.M. led to profound selective loss of the ability to acquire new memories for facts and events (Scoville & Milner, 1957). This phenomenon of anterograde amnesia resulting from MTL damage has been extensively studied since the first report of patient H.M. (Squire 1992). In addition, studies of these patients have identified a number of memory phenomena that are not affected by damage to the MTL. To provide a taxonomy for organizing the types of memory, the set of related memory phenomena that depend crucially on the MTL memory system are termed declarative memory while memory that does not depend on the MTL is termed nondeclarative memory.

**Declarative Memory**

Declarative memory refers to the acquisition and retrieval of facts, events and episodes. A nearly synonymous term, explicit memory, emphasizes the fact that these types of memory are available to awareness, that is, we consciously recall these memories during retrieval. The bulk of what is thought of as the ordinary, everyday operation of memory depends on declarative memory. The conventional use of the term memory generally refers to the retrieval process (e.g., remembering or failing to remember), but before memories can be retrieved, they have to be acquired (stored) and this process depends heavily on the intact function of the MTL. Patients with damage to the MTL exhibit anterograde amnesia, a loss of ability to acquire new memories for facts and events. In many of these patients, retrieval of remote memories is intact. If you converse with a patient with anterograde amnesia, they will often stories of their youth, which sometimes leads their family members to suggest that their “memory seems fine.” However, if you meet them again the next day, they will generally not recognize you and often will tell you again the same story they told you the previous day. The preservation of older memories is not absolute, but follows a temporal gradient such that recently acquired memories (e.g., recently experienced before the event that caused MTL damage and amnesia) are typically lost or impaired. As memories are assessed further backwards in time, the number and quality of memories increases in these patients. This phenomenon of temporally graded retrograde amnesia has been carefully studied in animal models of amnesia and indicates that long-term memories undergo a consolidation
process that depends on the MTL (Milner, Squire & Kandel, 1998). Damage to the MTL disrupts this process and in addition to interfering with acquisition of new memories, recently acquired memories are also lost.

The nature and neurobiology of the consolidation process is an area of active research and debate. Theories for how the MTL accomplishes consolidation range from endogenous processes (e.g., occurring during sleep; Ellenbogen, Payne & Stickgold, 2006) to retrieval-triggered strengthening to the creation of multiple parallel memory traces (Moscovitch et al., 2005a). In humans, retrograde amnesia can extend years or decades into the past, indicating that the consolidation process is very gradual (Rempel-Clower et al. 1996). Because of this, theories of consolidation become intertwined with questions about the nature of long-term memories. Episodic memory of prior events is sometimes described as mental time-travel, an internal transportation to something like a re-experiencing of a prior event. However, much of our long-term declarative memory is semantic facts that are retrieved and used without the experience of mental time travel. The multiple-trace theory model of consolidation suggests that retrieval of episodic memories is always supported by the MTL and that only retrieval of remote semantic memories can be fully independent of the MTL (Moscovitch et al 2005b). The “standard model” of consolidation is that both episodic and semantic memories initially depend on the MTL and are gradually consolidated by similar processes to become independent. While these theories make important distinctions about the representation of long-term memory, it is difficult to distinguish between them experimentally. Some severely amnesic patients appear to have intact remote episodic memories (Bayley et al 2003, 2005) but it is impossible to determine with certainty if they have the subjective experience of mental time travel when the episodes are recalled.

This question of the neural basis of long-term memory representation and retrieval is mirrored in experimental studies that have contrasted remembering with a strong “feeling of knowing” and retrieval that is based on familiarity. In this domain, two competing hypotheses about memory traces have been proposed to describe the neural basis of memory retrieval. The two-process model (Yonelinas, 2002) proposes that a feeling of familiarity emerges from cortical regions of the MTL (but not the hippocampus proper) as a memory trace is retrieved. However, a second component, specifically dependent on the hippocampus is crucial to having a strong feeling of knowing (and may also be related to the mental time travel associated with episodic retrieval). An alternate trace-strength model (Wixted, 2007) suggests that activity increases monotonically with the strength of the memory across the MTL during retrieval and there is no necessarily special role for the hippocampus.

Studies attempting to reconcile these competing hypotheses have often looked at the contributions to memory function from the prefrontal cortex. While lesion studies show that the MTL is critical for memory function, even the earliest functional neuroimaging studies found strongly correlated activity in prefrontal cortex (PFC) during memory tasks (e.g., Squire et al. 1992). The role of the PFC in memory is generally hypothesized to be via meta-mnemonic processes that guide search, retrieval and encoding strategies. Regions within the PFC have been found to show increased activity during source memory recollection (e.g., Rugg et al. 1999; Buckner 2002), retrieval monitoring (e.g., Dobbins et al. 2002), and encoding effort (Reber et al. 2002). While the specific cognitive functions associated with regions within the PFC is not yet fully
understood, it is generally hypothesized that the PFC modulates MTL function so that the MTL can support the different functions of encoding, retrieval and consolidation within a single neural system.

While the MTL is frequently the focus of neuroscientific studies of long-term memory, it is clear that it acts in conjunction with many other brain regions. The lateral temporal cortex is hypothesized to support the representation of consolidated memories as they become independent of the MTL. Recent neuroimaging studies have also regularly observed memory-related changes in activity in posterior parietal areas, suggesting that these regions may play an important supporting role in memory function (Wagner et al. 2005). A key idea is that memories are represented as information represented across cortical regions that are bound together through the MTL (Paller, 2002). This idea is supported by the analysis of the computational problem posed by rapid acquisition of complex memory traces, which suggests that this type of neural circuitry would allow for one-trial learning without interference with prior knowledge (McClelland et al. 1995). While other areas support the MTL, declarative memory is essentially defined by the dependence on the MTL for memory formation. The memory processes that have been found not to depend on the MTL and the associated neural circuits are collectively termed nondeclarative memory.

Nondeclarative Memory

While the bulk of what is generally thought of as the everyday operation of memory involves declarative memory, i.e., conscious recollection of facts and events, a number of tasks have been identified that indicate the existence of memory function outside the MTL. This type of memory typically operates outside awareness and is often referred to as implicit memory. The terms nondeclarative and implicit memory are nearly (but not completely) synonymous. Examples of this type of memory will be reviewed here from a neuroanatomical perspective as nondeclarative memory. In different paradigms, nondeclarative memory phenomena have been shown to depend on a number of different brain regions, indicating this type of memory is not a single coherent system, but is rather a collection of phenomena with different neural substrates. Nondeclarative memory includes perceptual priming, conceptual priming, habit learning, skill learning, motor sequence learning, certain kinds of category learning and some forms of classical conditioning.

An early demonstration of the phenomenon of priming is seen in the stem-completion test. Participants read a list of words including the word *window*. At test, no reference is made to the study list but 3-letter stems are given such as *win__* and participants are asked to complete the stem with the first word that comes to mind. The probability of completing the stem with a word from the study list like *window*, is much higher than if the word had not been studied. This “popping to mind” phenomenon occurs when participants are not directed to the study list and even on occasions where participants are explicitly told not to complete the stem with a word from the list (Jacoby, 1991). The critical neuroscientific observation is that patients with memory dysfunction due to damage to the MTL exhibit normal priming levels even when their declarative memory for the studied words is remarkably impaired (Hamman & Squire, 1997; Stark & Squire, 2000). This dissociation indicates that the changes in the brain that occurred during the study phase that leads words to pop to mind later are not dependent on the
normal functioning of the MTL. There must therefore be neural plasticity elsewhere in the brain supporting changes that give rise to the priming phenomenon.

Evidence for the location of the neural basis of changes supporting the phenomenon of priming comes primarily from studies of functional neuroimaging. A second encounter with a repeated word (image or face) typically evokes a smaller response in visual cortical areas associated with sensory processing of the stimulus than a prior presentation (Schacter & Buckner, 1998). This “repetition suppression” effect is thought to reflect a change in the state of sensory cortex that reflects a nondeclarative memory of the first presentation (Wiggs & Martin, 1998). Similar effects have also been documented in prefrontal cortex for semantic priming (e.g., Macotta & Buckner, 2004). The changes that support the repetition suppression effect are thought to be local, unavailable to awareness and to occur independently of the MTL.

While nondeclarative memory is a heterogeneous collection of phenomena, each type of nondeclarative memory is hypothesized to operate by similar principles: depending on local changes to a circumscribed brain region, the representation of these changes is unavailable to awareness, and the plasticity underlying the changes does not depend on the MTL. Examples of nondeclarative memory following this form are several forms of conditioning that have been well studied in experimental animals: delay eyeblink conditioning (depending on the cerebellum), fear conditioning (depending on the amygdala) and the gill withdrawal reflex of the aplysia meets these criteria (Milner, Squire, & Kandel, 1998).

A number of more complex forms of nondeclarative memory have been studied in humans that are expressed as skill, habit and category learning. Collectively, these studies demonstrate that the plasticity mechanisms outside the MTL are capable of establishing representations that can be complex and abstract even though they influence behavior without awareness. Studies of more complex forms of nondeclarative memory have been defined by a collection of specific tasks that are amenable to demonstrations of learning without awareness. A consistent challenge of this type of research is that when healthy subjects are able to consciously deduce the structure behind the experimental paradigm, they may use declarative memory to support performance. For this reason, tasks that have been effective at showing learning without awareness and preserved learning in amnesic patients have generally become frequently investigated model tasks.

A well-studied, simple task that has been very well studied as a model of nondeclarative memory is the Serial Reaction Time (SRT) task. First reported by Nissen & Bullemer (1987) the task requires learning a sequence of motor responses to visual cues. Typically, four cue locations are shown on a screen with four possible keypress (button) responses aligned underneath. Participants wait for the appearance of the cue and simply press the button beneath the cue as quickly as possible. After cue offset, it reappears in another location, the participant makes the corresponding response and this continues for several hundred trials. Participants are not told that the cue order follows a predictable structure, usually a repeating sequence of 10-12 locations. Response times become much more rapid with practice and knowledge of the repeating sequence is assessed by removing the sequence and observing a slightly slowdown in average response time. Knowledge of the repeating sequence occurs even when participants are unaware of the existence of the sequence and learning occurs at a normal rate in patients with memory disorders (Nissen & Bullemer, 1987; Reber & Squire, 1994). Functional
neuroimaging has implicated the basal ganglia (particularly the putamen) as contributing to perceptual-motor sequence learning in this task (e.g., Destrebecqz et al. 2005), which is consistent with the finding that patients with basal ganglia dysfunction due to Parkinson’s disease are impaired at SRT learning (Siegert et al. 2006). This form of nondeclarative memory is hypothesized to depend on changes in cortico-striatal circuits involved in motor response planning.

The phenomenon of Artificial Grammar Learning (AGL) was first reported by A. Reber (1967) and involves learning sequences of arbitrary symbols. Participants are shown nonsense strings of letters (e.g., PQXVT) during a study phase in which they are asked to memorize, copy or simply observe. No mention is made of the fact that a complex set of rules is used to create the letter sequences (the rules are generally represented as a finite state machine, i.e., an “artificial grammar”). After the study phase and a delay, participants are told that the prior strings were constructed according to rules and are then shown new strings and asked to judge which of the new strings correctly follow the same rules. Participants typically report no knowledge of the rules (but see Dulany, Carson & Dewey, 1985) but are able to make “grammaticality” judgments at above chance rates. While there has been controversy over the ability to assess a complete lack of rule knowledge, amnesic patients have been shown to learn the AGL task at a normal rate (Knowlton, Ramus & Squire, 1992) as have patients with Alzheimer’s disease (Reber, Martinez & Weintraub, 2003). The neural substrate of AGL is not yet known as different functional neuroimaging studies have implicated differing regions: the posterior parietal association cortex (Skosnik et al. 2002) or the basal ganglia (Lieberman et al. 2004; but see Reber & Squire, 1999).

A variety of category learning tasks have also been shown to depend on nondeclarative memory. The dot-pattern classification task described by Posner & Keele (1968) appears to be learned normally by amnesic patients (Knowlton & Squire, 1993) and is associated with changes in visual cortical areas (Reber et al. 2003). A task known as the Weather Prediction task, sometimes referred to as probabilistic classification, requires assigning groups of cues to two different categories and appears to be at least initially learned normally by amnesic patients (Knowlton, Squire & Gluck, 1994). Learning this task is impaired in patients with Parkinson’s disease (Knowlton, Mangels & Squire, 1996) and has been associated with increased activity in the basal ganglia (Poldrack et al. 2001). While questions have been raised about the degree to which this task depends selectively on nondeclarative memory (e.g., after an initial period of learning, patients do not exhibit learning at the same rate as controls), a convergent line of evidence supporting the idea of a cortico-striatal category learning system has emerged (e.g., Nomura et al. 2007). The second line of evidence uses simple visual stimuli (sine wave gratings) organized into categories by a decision bound in a two-dimensional stimulus space (Ashby & Ell, 2001). Category structures that do not lend themselves to conscious deduction of a verbalizable rule appear to depend on cortico-striatal circuits connecting the posterior caudate to extra-striate visual areas.

The examples here are not an exhaustive list of the types of memory function that reflect nondeclarative memory. Skill such as reading mirror-reversed text (Cohen & Squire 1980) and a number of different forms of priming (Schacter & Buckner, 1998) have been studied in patients and with functional neuroimaging. The wide variety of tasks and brain regions implicated suggests that nondeclarative memory is not a memory
system per se, but rather reflects a general principle of inherent plasticity in neural circuits that can support certain types of learning.

Understanding the principles behind nondeclarative memory may help resolve the question of the memory processes that are involved in tasks where the terms implicit and nondeclarative are not synonymous. For example, Chun & Phelps (1999) reported a type of priming within visual search that operates outside of awareness (implicit) but appears to depend on the MTL (declarative). This type of memory may also be related to the phenomenon of priming for new associations (Graf & Schacter, 1985) which also is also impaired in patients with MTL damage (Shimamura & Squire, 1989), but appears to operate outside awareness in healthy participants. These tasks suggest that some principles of nondeclarative memory may operate even on declarative memory representations. To accommodate these types of interactions among memory types, cognitive neuroscience theories of memories will eventually have to move beyond descriptions based solely on multiple systems.

**Conclusion**

Memory for the ordinary facts, episodes and events of our lives is supported by the operation of a complex set of neural circuits depending critically on the MTL. These circuits are crucial for the acquisition of new memories, recognition and recall of recently acquired information, and a consolidation process by which this information gradually becomes independent of the MTL. Meta-memory processes are hypothesized to emerge from interactions between the PFC via modulatory effects on the MTL. This set of circuitry does not encompass all of the memory abilities of the brain. A large number of mnemonic processes depend on brain regions operating independently of the MTL. These memory phenomena typically operate outside of awareness and are observed as skill, habit and category learning in addition to priming and certain kinds of simple conditioning. Some of these phenomena depend on neural circuits such as cortico-striatal loops, while others depend on local changes within sensory cortex, the amygdala or the cerebellum.
References


