IN SEARCH OF MEMORY TRACES

Richard F. Thompson
Neuroscience Program, University of Southern California, Los Angeles, California 90089-2520; email: thompson@neuro.usc.edu

Key Words learning, hippocampus, cerebellum, inactivation, localization

Abstract The key issue in analyzing brain substrates of memory is the nature of memory traces, how memories are formed, stored, and retrieved in the brain. In order to analyze mechanisms of memory formation it is first necessary to find the loci of memory storage, the classic problem of localization. Various approaches to this issue are reviewed. A particular strategy is proposed that involves a number of different techniques (electrophysiological recording, lesions, electrical stimulation, pathway tracing) to identify the essential memory trace circuit for a given form of learning and memory. The methods of reversible inactivation can be used to localize the memory traces within this circuit. Using classical conditioning of eye blink and other discrete responses as a model system, the essential memory trace circuit is identified, the basic memory trace is localized (to the cerebellum), and putative higher-order memory traces are characterized in the hippocampus.

CONTENTS

INTRODUCTION ..................................................... 1
Simplified Systems .................................................. 4
Strategies to Study Memory Formation ........................ 4
OUR MODEL SYSTEM: EYEBLINK CONDITIONING .......... 6
Motor Control ...................................................... 6
The Hippocampus .................................................. 8
The Essential Circuitry: The Cerebellum ....................... 11
CONCLUSION ....................................................... 16

INTRODUCTION

A major achievement of recent research on brain substrates of learning and memory, to which our work has contributed, is the recognition that there are several different forms or aspects of memory involving different brain systems (Figure 1). Squire (1992) has argued eloquently for the distinction between declarative and nondeclarative forms of memory, as has Schacter (1987). The distinction between episodic and semantic memory—“What did you have for breakfast?” versus “Where is the Eiffel Tower?”—has been stressed by Tulving (1985).
Figure 1  A current view of the various forms of learning and memory and their putative brain substrates.
Implicit or nondeclarative memory is very much a grab bag. In general, explicit memory involves awareness of the memory whereas implicit memory does not necessarily involve being aware of the memory (aware meaning verbal report). The schema of Figure 1 is of course oversimplified. When an organism learns something, a number of brain systems can become engaged. However, in most cases there is one critical brain system, which when damaged causes permanent impairment in the particular form of learning and memory.

Many readers will recall Lashley’s (1950) pessimistic conclusion that his series of experiments “has yielded a good bit of information about what and where the memory trace is not” in his famous article, “In Search of the Engram.” In all fairness to Lashley, the existence of several different forms of memory with differing neuronal substrates was not recognized at that time, nor were modern analytic techniques available then.

We adopt the following definitions from an earlier review (Lavond et al. 1993). The essential memory trace, together with its associated circuitry, is necessary and sufficient for the basic aspects of a given form of learning (e.g., acquisition and retention). Other brain structures may also form memory traces, defined loosely as learning-induced changes in neuronal activity, but these may not be essential for the basic learning. Thus, in eyeblink conditioning, long-lasting learning-induced increases in neuronal activity develop in the hippocampus (Berger et al. 1976), but the hippocampus is not necessary for basic delay associative learning and memory (Schmaltz & Theios 1972).

Aversive classical conditioning can modify the receptive-field properties of neurons in sensory systems. Particularly striking are the studies of Weinberger and associates (Edeline & Weinberger 1991, Weinberger et al. 1984) showing that neurons in secondary areas of auditory thalamus and cortex shift their best frequencies toward the frequency of the conditioning stimulus in aversive Pavlovian conditioning (see also Bao et al. 2003). The motor area of the cerebral cortex provides yet another example. Classical eyeblink conditioning results in marked and persisting increases in excitability of pyramidal neurons in motor cortex (Woody et al. 1984). However, the motor cortex is not necessary for either learning or memory of the conditioned eyeblink response (Ivkovich & Thompson 1997). These “higher-order” memory traces appear to play important roles in the adaptive behavior of the organism.

The obvious point here is that the organism can learn, remember, and perform the basic learned response following destruction of nonessential memory trace systems, but destruction of the essential memory trace system abolishes this ability. We draw a distinction between the essential memory trace and the essential memory trace circuit. The latter includes the necessary input and output circuits as well as the essential memory trace itself. These are all identified by lesions. But lesions, per se, cannot distinguish between the essential trace and the essential circuitry; lesions of either completely prevent and abolish the learned response.
Simplified Systems

How does one go about finding memory traces? One very productive approach is the use of simplified “model” systems, emphasized early by Thompson & Spencer (1966) (see also Kandel & Spencer 1968). Thus, vertebrate spinal reflexes and the gill-withdrawal reflex circuit in Aplysia have both served as very productive models to analyze processes of habituation and sensitization. In the case of habituation, at least, the memory trace is embedded within the reflex pathway under study so lesions abolish the reflex as well as the trace. For monosynaptic pathways in both Aplysia and spinal cord the mechanism of short-term habituation appears to be simply a decrease in the probability of transmitter release as a result of repeated activation, a presynaptic process of synaptic depression (Farel & Thompson 1976, Kandel 1975).

But even for a simple behavioral phenomenon like habituation, the neural substrate can prove to be complex, particularly for long-lasting habituation, e.g., between sessions as opposed to within session response decrements. In the Aplysia cellular monosynaptic system, Ezzeddine & Glanzman (2003) have shown that prolonged habituation depends on protein synthesis, protein phosphatase activity, and postsynaptic glutamate receptors. Indeed, their data indicate that stimuli inducing habituation of the gill-withdrawal reflex appear to activate sensitizing processes as well, as proposed in the “dual-process” theory of habituation by Groves & Thompson (1970).

Thus, much is still not known about the detailed mechanisms of memory traces in this simplest form of learning. By mechanisms we mean the physical-chemical substrate of memory store. Although there are many candidate mechanisms, as of this writing we do not yet know the detailed physical bases of memory storage for any form of learning and memory in the mammalian brain. But understanding the mechanisms of memory storage will not inform us of what the memories are. The actual memories are coded by the neuronal circuits that code, store, and retrieve the memories, what I term here the “essential circuits.”

Strategies to Study Memory Formation

There appear to be two general strategies for the study of memory storage. One approach is to pick a mechanism of neural plasticity such as long-term potentiation (LTP) in a structure known to be important for memory and attempt to show that it is a substrate for memories. Richard Morris and associates published a heroic effort to demonstrate that activity-dependent synaptic plasticity, particularly LTP (and LTD, long-term depression), “is both necessary and sufficient for the information storage underlying the type of memory mediated by the brain area in which that plasticity is observed” (Martin et al. 2000). They established several “formal” criteria by which to judge this hypothesis. We examine these criteria below but here we simply note their conclusion that “synaptic plasticity is necessary for learning and memory, but that little data currently supports the notion of sufficiency” (p. 649).
Actually, there are many forms of plasticity in the nervous system, synaptic plasticity being only one. There are also changes in neuron excitability due to factors intrinsic to the neurons, as in decreases in after-hyperpolarization, as well as structural changes in dendrites, formation of new neurons, and actions of glial cells. LTP has been the most intensely studied putative synaptic mechanism of memory storage, particularly in the hippocampus (e.g., Martin et al. 2000), but the evidence is still far from clear. There is a vast literature analyzing mechanisms of LTP but very little work attempting to show that it is a basis for the storage of memories (see Shors & Matzel 1997, Wilson & Tonegawa 1997).

The alternative strategy is to begin with a well-characterized form of learning and memory and determine the memory traces involved. In order to do this it is first necessary to find where in the brain the memories are stored, the classical problem of localization. Because learning involves changes in behavior as a result of exposure to stimuli that do not change, there must be alterations at some loci in the brain. A given form of learning may involve one or multiple loci of memory storage, but they must exist. We have used this strategy in our search for memory traces.

How does one go about finding a memory trace in the behaving mammal? Olds, Disterhoft, and associates wrestled with this issue in pioneering studies (Olds et al. 1972). Their basic approach was to record extracellular unit cluster discharges in many brain areas in freely moving rats, looking for increases in unit response on the conditioning day that were significantly greater than responses to the stimuli on the preceding pseudoconditioning day, i.e., a learning-induced increase in unit activity. The behavioral situation was simply an auditory conditioned stimulus (CS+) followed after 2 sec by a food pellet unconditioned stimulus (US) versus a CS− with no food pellet. The studies were very well controlled.

The initial criterion they set for identification of the memory trace was shortest latency learned responses:

The “earliest” conditioned brain responses (i.e., those appearing with the shortest latencies after application of the CS) might thus be considered to be “at the site” of conditioning, and other later conditioned brain responses might be considered to be secondary to them. Similarly (and hopefully validating this supposition) “early” new brain responses might be expected to emerge at the site of old responses indicating this site to be a junction point between old and new (Olds et al. 1972, p. 202).

In a most interesting thought piece, Michael Gabriel (1976) argued against the use of the shortest latency response as the criterion for identification of the memory trace. He proposed an alternative “bias” hypothesis. In essence, he suggested that a tonic flow of impulses originating from a distant neuronal locus biased the activity of the neurons being recorded as the shortest latency learned response. So the increase of the shortest latency unit response could be due to tonic actions on these neurons from a distant site where learning has resulted in a tonic increase in activity. Gabriel draws the following rather strong conclusion: “clearly the bias
hypothesis argues against the idea of mapping the brain for short-latency neuronal responses in order to localize engrams” (p. 279).

In a companion piece to the Olds et al. (1972) article, Disterhoft & Olds (1972) do consider the possibility of tonic influences, as Gabriel notes. They suggest an additional criterion, namely the earliest increase in neuronal activity that develops over the trials of training. I return to this issue below.

In the late 1960s and early 1970s we tried several different types of model systems to analyze neural substrates of mammalian associative learning and memory. Because of our success in using spinal reflexes for analysis of habituation we tackled classical conditioning of the flexion reflex in the acute spinal animal, with some success (Patterson et al. 1973). However, I concluded it was not a good model of associative learning in the behaving mammal.

OUR MODEL SYSTEM: EYEBLINK CONDITIONING

Thanks in part to efforts by Michael Patterson, then a postdoc in my lab who had obtained his Ph.D. with Isadore Gormezano, and because of the elegant and comprehensive behavioral studies by Gormezano, we adopted his preparation: classical conditioning of the nictitating membrane (NM) eyeblink response in the restrained rabbit. We have detailed the advantages of this preparation elsewhere (Thompson et al. 1972, Thompson et al. 1976). Perhaps the most important advantages are: (a) The unconditioned response (UR) provides an independent measure of performance against which to compare effects of variables acting on the conditioned response (CR), and (b) the behavioral conditioned response is robust, reliable, and discrete, and the exact amplitude-time course of the response is measurable. There were a number of other advantages as well: no sensitization, pseudoconditioning, or alpha responding; many trials required to learn; and extensive parometric data on the properties of the CR (see Gormezano et al. 1983, and the important theoretical analyses by Wagner and associates, e.g., Wagner & Donegan 1989).

Motor Control

We began our analysis with a focus on the motor nerves and nuclei that generated the reflex and learned eyeblink responses (Cegavske et al. 1976, Cegavske et al. 1979, Young et al. 1976). To oversimplify from our work and the work of others, the sixth nerve was critically important for the nictitating membrane (NM) response, but the third and fourth nerves were also involved and the seventh nerve was critical for closure of the external eyelid (obicularis oculi muscles). In terms of motor nuclei, the seventh nucleus controlled external eyelid closure and the accessory portion of the sixth nucleus innervated the retractor bulbous muscle, which retracted the eyeball and forced the nictitating membrane out across the front surface of the eye. The fourth and sixth nuclei acted synergistically with the accessory sixth and seventh. We showed with simultaneous recordings that NM extension and external eyelid closure [actually electromyographic (EMG)
activity of the obicularis oculi muscles] were essentially perfectly correlated over the course of learning (McCormick et al. 1982b). Indeed, obicularis oculi EMG appears to be the most robust and sensitive measure of learning (Lavond et al. 1990). Recordings from the relevant motor nuclei, particularly the sixth, accessory sixth, and seventh, showed essentially identical patterns of learning-induced increases in neuronal activity.

It is important to stress the fact that the pathways mediating the reflex eyeblink are in the brain stem and do not involve “higher” brain systems such as the cerebellum and hippocampus. Specifically, there are direct projections from the trigeminal nucleus (activated by stimulation of the cornea and surrounding tissues) to the relevant motor nuclei and indirect projections relaying via the brain stem reticular system to the motor nuclei (Hiraoka & Shimamura 1977).

The basic logic of our approach seemed reasonable—identify the critical motor neurons involved in control of the learned response and trace the essential circuitry backward to the source, the “engram.” The fact that several motor nuclei showed the same pattern of learning-induced activation argued against any plasticity unique to one motor nucleus (e.g., the accessory sixth) or to processes of plasticity at the level of the motor nuclei. Instead, it seemed most likely that the learning-induced response in the motor nuclei was driven from a common central source. The pattern of learning-induced increase in the activity of neurons in the motor nuclei in the CS period in paired trials and on CS-alone trials (i.e., the neuronal CR) was strikingly similar in form to the amplitude-time course of NM extension, external eyelid closure, and of course the EMG recorded from the obicularis oculi muscles. Indeed, an envelope of increased frequency of discharge of neurons in the motor nuclei preceded in time and closely predicted the form of the behavioral eyeblink-NM response (Figure 2). This close predictive parallel was most evident with unit cluster recordings but was also true for single unit recordings in the motor nuclei.

USE OF THE MOTOR NEURON NEURONAL MODEL. Our initial logic—to work backward from the motor nuclei—was not as simple as it might have seemed because of the very large number of central brain systems that project to the motor and premotor nuclei. So we adopted a different strategy. It was apparent that the amplitude-time course form of the conditioned eyeblink-NM response closely paralleled and followed the pattern of increased unit activity in the motor nuclei. So we focused on this behavioral model of the learning-induced increase in neural activity in the motor nuclei and used it as a template or model (Figure 2). The higher brain systems that act to generate and drive the learning-induced neuronal CR, the “model” in the motor nuclei, must show the same pattern of learning-induced activity as do the motor nuclei, and hence the amplitude-time course of the learned eyeblink-NM response. Thus, we searched through higher brain systems looking for learning-induced neuronal activity that correlated closely with and preceded in time the form of the learned eyeblink-NM response. In essence, we mapped virtually the entire brain of the rabbit in 1 mm steps, searching for neuronal models...
Figure 2  Histograms of unit cluster recordings from the ABD and simultaneous recording of the eyeblink response (NM, the “third eyelid”) in classical conditioning in the rabbit summed and averaged over eight trials before (A) and after (B) learning. First cursor onset of tone CS, second cursor onset of corneal airpuff US, trace duration 750 msec, upward movement of the NM trace indicates extension (eyelid closure). Note that the pattern of neural discharges precedes and models the amplitude-time course of the behavioral response. ABD, abducens motor nucleus; CS, conditioned stimulus; NM, nictitating membrane; US, unconditioned response. (From Cegavske et al. 1979.)

of the learned behavioral response (see, e.g., McCormick et al. 1983, Thompson et al. 1976). We did not search blindly but rather system-by-system.

The basic assumption is that the neuronal sites of memory trace formation will show a pattern of increased frequency of unit cluster and single-unit discharges that precedes and closely predicts the pattern of response in the motor nuclei and hence in the amplitude-time course of the behavioral CR. We felt this criterion was preferable to the earlier ideas of shortest latency response, earliest response over trials, or first site to show increased tonic responses. This assumption proved to be the key that permitted us to localize a memory trace.

The Hippocampus

Because of its involvement in memory, we began with the hippocampus (Berger et al. 1976). To our delight, both unit cluster recordings and single-unit responses recorded from CA1 and CA3 showed the requisite learning-induced responses, i.e., pattern of increased discharge frequency that preceded in time (within trials) and correlated virtually perfectly with the amplitude-time course of the behavioral CR (Figure 3).

We showed that this neuronal model of a memory is generated by identified pyramidal neurons (Berger & Thompson 1978b). Furthermore, the hippocampal unit response began to develop in the US period within just a few trials of training (a criterion by Olds et al. 1972) and moved into the CS period in close association...
with the development of the behavioral CR (Berger & Thompson 1978a). It seemed like a perfect candidate for the engram. In addition, electroencephalogram (EEG) activity recorded from the hippocampus at the beginning of training predicted the rate of learning (Berry & Thompson 1978). We characterized this learning-induced response in the hippocampus in some detail (see Berger et al. 1986). Unfortunately, as we noted above, animals could learn the basic delay-conditioned NM-eyeblink response following hippocampal lesions (Schmaltz & Theios 1972).

This apparent enigma illustrates a fundamental limitation of using neuronal recordings to localize memory traces. The fact that a neuronal model of the learned response occurs at a given locus does not necessarily mean it originates there. Indeed, the learning-induced neuronal model in the hippocampus, at least in the CS period, is abolished by appropriate cerebellum lesions (Clark et al. 1984). Actually, a number of loci in the brain exhibit the learning-induced neuronal model of the learned behavioral response. I return to this issue below.

The breakthrough insofar as the hippocampus is concerned came in a study in our laboratory by Paul Solomon and Donald Weisz using the trace-conditioning

---

**Figure 3** Behavioral eyeblink (nictitating membrane) and hippocampal unit response averaged over 100 trials in a well-trained rabbit. The unit response is of a single, identified pyramidal neuron in CA1. Upper trace, raw single 100-msec sweep to show the neuron response, middle trace averaged NM response, lower trace cumulated histogram of unit activity; cursors and durations in middle and lower traces as in Figure 2. Note that the pattern of response of this neuron both precedes and models the amplitude-time course of the behavioral response. (From Berger & Thompson 1978b.)
paradigm, where a 500-msec period of no stimuli separated CS offset from US onset (Solomon et al. 1986). In brief, large bilateral hippocampal lesions made before training markedly impaired subsequent learning of the trace (but not delay) eyeblink CR, a result replicated exactly by Moyer et al. (1990). If the hippocampal lesions are made immediately after learning, the trace CR is abolished (with no effect on the delay CR), but if the lesions are made a month after training the trace CR remains intact (Kim et al. 1995).

In sum, hippocampal lesions induced anterograde amnesia (impairment of postlesion learning) and marked but time-limited retrograde amnesia for the trace eyeblink CR. These are precisely the effects that such lesions have on declarative memory in humans (Squire 1987). These results suggest that trace eyeblink conditioning provides a simple model of hippocampal-dependent declarative memory, a possibility strongly supported by studies of humans with hippocampal-medial temporal lobe amnesia. Such patients are markedly impaired on acquisition of trace eyeblink conditioning if the trace interval is sufficiently long but not impaired at all on delay conditioning (see McGlinchey-Berroth et al. 1997).

Clark & Squire (1998, 1999) made the striking observation that awareness of the training contingencies in normal human subjects correlated highly with the degree of learning of trace eyeblink conditioning. Recall that awareness is a key defining property of declarative memory. They also showed that awareness played no role in delay conditioning, and that expectancy of US occurrence influenced trace but not delay conditioning (Clark & Squire 2000, Clark et al. 2001). They conclude that delay and trace conditioning are fundamentally different phenomena, with delay conditioning inducing nondeclarative or procedural memory and trace conditioning inducing declarative memory. Hence, trace eyeblink conditioning in rabbits would seem to provide a very elementary instantiation of declarative memory in animals.

These strikingly parallel results of hippocampal lesions in rabbits and humans for trace eyeblink conditioning suggest the possibility that a memory trace(s) may develop in the hippocampus in trace conditioning, but do not prove it. Recall that delay eyeblink conditioning results in the development of a pronounced neuronal model of the learned response in the hippocampus (as does trace training).

Several lines of evidence support the view that localized changes do develop in the hippocampus in eyeblink conditioning. Thus, the properties of increased neuronal activity are strikingly similar to the properties of LTP (Berger et al. 1986, Thompson 1997, Weisz et al. 1984), and induction of LTP in the hippocampus can enhance eyeblink learning (Berger 1984). There is also a dramatic learning-induced decrease in the after-hyperpolarization in pyramidal neurons in eyeblink conditioning, most evident in trace conditioning, which leads to increased excitability. This alteration is intrinsic to the neurons due to a decrease in one or more calcium-dependent outward potassium currents (Disterhoft et al. 1986, Disterhoft & McEchron 2000). There is also a dramatic increase in dendritic membrane-associated protein kinase C after eyeblink conditioning (Olds et al. 1989). After trace eyeblink conditioning there is a substantial increase in the number of multiple
synapse boutons in hippocampal neurons (Geinisman et al. 2001). Finally, there is a dramatic increase in the number of new neurons in the hippocampus in trace, but not delay, eyeblink conditioning (Gould et al. 1999, Shors et al. 2001).

These sets of structural and chemical changes are certainly consistent with the formation of higher-order memory traces in the hippocampus. But recall that cerebellar lesions can abolish both the behavioral learned response and the increased neuronal response in the CS period in the hippocampus in both delay and trace learning (see below and Clark et al. 1984, Woodruff-Pak et al. 1985). So doubts remain. Interestingly, the early (over trials)-developing increase in neuronal activity in the hippocampus in the US period is not completely abolished by cerebellar lesions (Clark et al. 1984). In my view, it is necessary to identify the entire essential circuitry from stimulus input to motor output in order to establish definitively that the essential memory trace, e.g., for trace conditioning, develops and is stored, albeit temporarily, in the hippocampus (see below). To date, not all the essential circuitry has been identified (see Thompson & Kim 1996). Yet another remaining puzzle is where the trace memories go after the hippocampus is no longer necessary.

The Essential Circuitry: The Cerebellum

So the hippocampus is not the locus of the essential memory trace for standard-delay classical eyeblink conditioning. In a detailed mapping study of the brain stem we identified several loci where neurons exhibited the requisite neuronal model of the CR—increase in frequency of discharge that preceded the onset of the US-UR and modeled the amplitude time source of the behavioral CR, just as did the neuronal response in the hippocampus (McCormick et al. 1983). The following brain areas showed this neuronal response predictive of the learned behavioral CR: the relevant motor nuclei (as noted above), a region in the vicinity of the fifth nucleus, various reticular regions, the cerebellar cortex (ansiform and anterior lobes), the cerebellar interpositus nucleus, the pontine nuclei, and the red nucleus. The response in the interpositus nucleus was particularly robust (Figure 4).

The fact that a neuronal model of the behavioral CR developed in the cerebellum was suggestive of a memory trace but we knew from our earlier work on the hippocampus that neuronal recordings, per se, could not identify the essential memory trace. We completed a series of lesion studies, initially involving large cerebellar lesions and later lesions limited to the interpositus nucleus (e.g., McCormick & Thompson 1984a,b; Clark et al. 1984), all of which completely prevented learning and completely and permanently abolished the CR with no effect on the UR (thus ruling out performance variables).

In a long series of studies completed while we were at Stanford, we identified virtually the entire essential memory trace circuit using electrophysiological recordings, electrical stimulation, lesions (aspiration, electrolytic, and chemical), and anatomical pathway tracing methods (see Chapman et al. 1988; Christian & Thompson 2003; McCormick et al. 1982a, 1985; Lavond et al. 1985; Mauk et al. 1986; Steinmetz et al. 1986, 1987; J.K. Thompson et al. 1985; Woodruff-Pak et al.
Figure 4 Neuronal model of learned eyeblink response recorded from the cerebellar interpositus nucleus. Each graph shows nictitating membrane movement (top trace) and a histogram of multiple unit activity (bottom trace) averaged and summed over 100 trials. Animals that received explicitly unpaired presentations of the conditioning stimuli do not develop altered patterns of neuronal activity (left column). Trained animals develop a neuronal model of the learned behavior. Total trace duration 750 msec. (From McCormick & Thompson 1984b.)

1985). In brief, the efferent CR pathway projects from the interpositus nucleus ipsilateral to the trained eye, via the superior cerebellar peduncle, to the contralateral magnocellular red nucleus and to the relevant motor nuclei ipsilateral to the trained eye. The cerebellar interpositus lesion abolition of the eyeblink CR is strictly ipsilateral and is due to damage to neuron somas, not fibers of passage. The inferior
olive-climbing fiber system appears to be the critical US reinforcing or teaching pathway and the pontine nuclei and mossy fiber system appears to convey the necessary CS information to the cerebellum.

A particularly satisfying aspect of our discovery of the essential role of the cerebellum in classical conditioning of discrete responses is the relevance of our work to the human condition. Irene Daum and associates, working in Germany, replicated in humans the fact that appropriate cerebellar damage completely prevents learning of the eyeblink CR (Daum et al. 1993), since replicated in many other studies. Christine Logan (a former graduate student of mine) and Scott Grafton, a neurologist then at the University of Southern California, completed an extensive positron emission tomography (PET) study of eyeblink conditioning in humans. They found significant activation in the cerebellar interpositus nucleus and several loci in cerebellar cortex, in close agreement with our recording studies in the rabbit cerebellum (Logan & Grafton 1995), again replicated in many other studies.

A point that is obvious from our studies but seems not to be widely understood is that our results on the role of the cerebellum in classical conditioning apply to the learning of any discrete movement: eyeblink, head turn, forelimb flexion, hindlimb flexion, etc. (see Shinkman et al. 1996). Eyeblink conditioning is simply a convenient response to measure. Our findings to date seem to have identified perhaps the critical function of the cerebellum, namely the learning of discrete skilled movements, a basic notion proposed in classic theories of the cerebellum as a learning machine (see Albus 1971, Eccles 1977, Ito 1984, Marr 1969). Indeed, our work constitutes a compelling verification of these theories.

Our data to this point strongly supported the hypothesis that the essential memory trace was formed and stored in the cerebellum, but did not prove it. Hence we adapted use of methods of reversible inactivation to localize the memory trace. The logic is straightforward. Having identified the essential memory trace circuit, reversibly inactivate each key locus in the circuit during training. If this completely prevents learning at a given locus, then this locus either conveys essential afferent information to the memory trace or is the site of the memory trace. However, if reversible inactivation of a given locus does not prevent learning at all, then this locus is efferent from the memory trace. But remember that inactivation of all these essential loci will completely prevent expression of the CR. We infused muscimol for inactivation of neuron cell bodies—it acts on gamma amino butyric acid (GABA<sub>A</sub>) receptors as an agonist and completely shuts down (hyperpolarizes) the neurons for several hours, after which they fully recover. To inactivate axons we infused tetrodotoxin (TTX) (see, e.g., Krupa et al. 1993; Krupa & Thompson 1995, 1997; Krupa et al. 1996; Thompson & Krupa 1994). David Lavond and his students completed parallel studies using reversible cooling (e.g., Clark et al. 1992, Clark & Lavond 1993).

Results are completely consistent and have since been replicated in other laboratories. In brief (see Figure 5) inactivation limited to the anterior interpositus nucleus completely prevented learning. After removal of inactivation animals showed no signs of learning; with subsequent training they learned normally as though they
Figure 5  Highly simplified schematic of the essential memory trace circuit for delay classical conditioning of the eyeblink response to illustrate the use of reversible inactivation to localize memory traces. Shading in a, b, and c indicate reversible inactivation of the key region of each structure during training using muscimol; d indicates inactivation of the axonal pathway exiting from the cerebellum, the superior cerebellar peduncle, using tetrodotoxin. Inactivation of the interpositus (c) completely prevents learning, but inactivation of the superior peduncle (d), the red nucleus (b), and the motor nuclei (a) does not prevent learning at all. (Modified from Thompson & Krupa 1994.)

had no prior training; there was no savings. In complete contrast, inactivation of the superior peduncle (the immediate efferent projection from the interpositus/dentate nuclei), the red nucleus, and the relevant motor nuclei, although completely preventing expression of the CR, did not prevent learning at all. After removal of the inactivation the animals had fully learned to asymptote. Thus, inactivation localized to the anterior interpositus nucleus completely prevented learning but inactivation of its immediate output pathway did not prevent learning at all. The fact that after interpositus inactivation there was no savings argues strongly that no part of the memory trace developed in structures afferent to the interpositus. Similarly, the fact that inactivation of structures efferent from the interpositus did not prevent learning at all argues that no part of the memory is formed in these structures.

I noted above that several loci in the brain exhibited the neuronal model of the learned behavioral CR, e.g., trigeminal nuclear region, pontine nuclei, etc. In a series of studies Lavond and associates (and we) showed that with inactivation of
the interpositus nucleus, all these models disappeared, i.e., they are all driven from the interpositus (Bao et al. 2000, Lavond & Cartford 2000).

Because muscimol acts only on neuron cell bodies and not on axons, the inactivation of the interpositus nucleus does not prevent normal activation of cerebellar cortex by the two major afferent systems, climbing fibers from the inferior olive and mossy fibers from many sources. The fact that no learning at all occurred with inactivation limited to the interpositus nucleus would seem to argue against formation of memory traces in the cerebellar cortex. However, this inactivation blocks the direct projections from the interpositus nucleus to the cerebellar cortex.

When we first discovered the essential role of the cerebellum in eyeblink conditioning I had hoped the memory traces would be formed and stored in the cerebellar cortex. The cellular machinery in the cortex is vast (each Purkinje neuron receives up to 200,000 synapses from granule neurons). However, we were never able to completely prevent or abolish the CR by lesions limited to the cerebellar cortex. Other workers, particularly Yeo and associates (Yeo et al. 1984), claimed to have done so. The problem is with the lesion method—it is impossible to remove all of the cerebellar cortex without damaging the critical region of the interpositus nucleus (see detailed discussion in Christian & Thompson 2003). However, we consistently found that with large cerebellar cortical lesions or reversible inactivation, learning was much slower and to a lesser degree and adaptive timing of the CR was lost. Normally the eyeblink closure CR peaks at the onset of the US; it is a maximally adaptive response. After large cortical lesions, the CR peaks earlier in time and is no longer adaptive (Garcia et al. 1999, McCormick & Thompson 1984b). So the cortex is critically important for normal adaptive learning and it seemed very likely that higher-order memory traces were established there. But again the lesion method is inconclusive.

Fortunately nature provided us with an ideal preparation, the Purkinje cell degeneration (pcd) mutant mouse. The brain of this mutant develops normally until about two weeks after birth. Then over the next several weeks, all Purkinje neurons in the cerebellar cortex die; as a result the animals have no functional cerebellar cortex. But the interpositus nucleus remains functional. Results were clear: The pcd mice were able to learn (L. Chen et al. 1996). They learned much more slowly and to a lesser degree than normal mice of the same strain (littermates), but they still showed significant and substantial learning. The CRs were shorter latency in the pcd mice. They also showed rapid extinction with CS-alone training. The possibility that the memory trace was somehow formed in loci other than the cerebellum was ruled out by lesioning the interpositus nucleus bilaterally in pcd mice before training. The lesioned pcd mice were unable to learn the conditioned eyeblink response (L. Chen et al. 1999).

These results were very similar to the effects of large cerebellar cortical lesions we had found earlier in rabbits. So the cortex is critically important for normal learning, and higher-order memory traces are likely formed there. There is considerable evidence supporting the view that a process of long-term potentiation
(LTD) at the parallel fiber-Purkinje neuron dendritic synapses, discovered by Ito (see 1984), plays a key role in initial plasticity (C. Chen & Thompson 1995, C. Chen et al. 1995, Shibuki et al. 1996, Kim & Thompson 1997). Indeed, Purkinje neurons show substantial changes in CS-evoked discharge frequency over the course of training, with many showing learning-induced decrease in discharge frequency, as could be expected if LTD developed (Christian et al. 2002, Christian & Thompson 2003). Other lines of evidence also support the view that memory traces are formed in cerebellar cortex (Cooke et al. 2004).

The evidence is now very strong that the basic essential memory trace is formed and stored in the anterior interpositus nucleus for classical conditioning of the eyeblink response. Indeed, the reversible inactivation studies provide the key evidence (Figure 5). It is clear that long-lasting changes in the critical neurons in the anterior interpositus nucleus do develop. Thus, infusion of protein synthesis inhibitors in this locus completely prevents learning (Bracha et al. 1998, G. Chen & Steinmetz 2000, Gomi et al. 1999). Further, training results in selective expression of a protein kinase in this locus, KKIAMRE, an enzyme for the cell division cycle (Gomi et al. 1999). Pharmacological isolation of the interpositus from the cerebellar cortex reveals clear learning-induced increases in excitability of interpositus neurons (Bao et al. 2002, Garcia & Mauk 1998). Perhaps most convincing, eyeblink conditioning results in a dramatic and highly significant increase in the number of excitatory synapses (but not inhibitory synapses) in the interpositus nucleus (Kleim et al. 2002).

CONCLUSION

The essential circuitry for classical conditioning of the eyeblink response is shown in Figure 6, along with the site of memory trace formation in the interpositus nucleus and a putative site of plasticity in the cerebellar cortical neurons. The nature of the memory is defined by this circuit; the circuit is the memory. The CS activates the sensory afferent pathways to the site(s) of trace storage in the cerebellum, which activates the efferent pathways to the motor nuclei and the learned behavior. The content of the memory, the conditioned eyeblink response, is completely defined and completely predictable from the essential circuit.

The formal criteria developed by Richard Morris and associates to demonstrate that a given set of phenomena establish what they term the “synaptic plasticity and memory” (SPM) hypothesis are as follows (Martin et al. 2000):

1. DETECTABILITY: If an animal displays memory of some previous experience, a change in synaptic efficacy should be detectable somewhere in its nervous system.

2. MIMICRY: Conversely, if it were possible to induce the same spatial pattern of synaptic weight changes artificially, the animal should display “apparent” memory for some past experience which did not in practice occur.
Figure 6  Simplified schematic (most interneurons omitted) of the putative essential circuitry for delay classical conditioning of eyeblink (and other discrete responses) learned with an aversive US. (The sensory and motor nuclei activated depend of course on the nature of the CS and US; the more central portions of the circuit appear to be general.) The reflex US-UR pathway involves direct and indirect projection from the trigeminal nucleus to the motor nuclei (for the eyeblink UR and CR, primarily accessory 6 and 7). The tone CS pathway projects from auditory nuclei to the pontine nuclei and to the cerebellum as mossy fibers. The US pathway includes projections from the trigeminal to the inferior olive and to the cerebellum as climbing fibers. The CR pathway projects from the interpositus to the red nucleus and on to premotor and motor nuclei. There is also a direct GABAergic inhibitory projection from the interpositus to the inferior olive. Solid cell bodies and bar terminals indicate inhibitory neurons; open cell bodies and fork terminals indicate excitatory neurons. Stars indicate sites of plasticity based on current evidence. See text for details. (From Christian & Thompson 2003.) CR, conditioned response; CS, conditioned stimulus; UR, unconditioned response; US, unconditioned stimulus.

3. ANTEROGRADE ALTERATION: Interventions that prevent the induction of synaptic weight changes during a learning experience should impair the animal’s memory of that experience.

4. RETROGRADE ALTERATION: Interventions that alter the spatial distribution of synaptic weights induced by a prior learning experience (see detectability) should alter the animal’s memory of that experience (p. 651).
I would expand the SPM notion to include nonsynaptic mechanisms of plasticity and memory as well, referring to all as the memory trace. I would also note that these criteria cannot be met unless the putative memory trace has been localized. Evidence to date on localization of the memory trace(s) for classical conditioning of eyeblink and other discrete responses would seem to satisfy Morris’ criteria.

1. DETECTABILITY: There is a dramatic increase in neuronal/synaptic efficacy in the cerebellum, both in the cortex of lobule HVI and in the anterior interpositus nucleus as a result of training (see, e.g., Shinkman et al. 1996, Bao et al. 2002). Indeed, long-term potentiation (LTP) has been reported to occur in the interpositus nucleus (Racine et al. 1986).

2. MIMICRY: Electrical microstimulation of the locus of the memory trace in the anterior interpositus nucleus can evoke the response to be learned before training (McCormick & Thompson 1984a, Chapman et al. 1988).

3. ANTEROGRADE ALTERATION: Infusion of muscimol into the site of the memory trace in the anterior interpositus nucleus completely prevents learning, i.e., induction of the memory trace.

4. RETROGRADE ALTERATION: Infusion of muscimol into the site of the memory trace in the anterior interpositus nucleus in a well-trained animal alters the synaptic weights (shuts them down) and abolishes the animal’s memory for that experience.

These of course are only opinions of what constitutes demonstration of a memory trace. In my view the key is that the essential circuit defines the memory, as I noted above. But we still do not know the detailed nature of the memory trace in the interpositus and how it is formed. It will be necessary to identify all the steps in the causal chain from initial activation of the neurons at the beginning of training to the final form of the memory trace, from the biochemical/genetic processes to the structural changes in the synapses and neurons that code the permanent memory trace.

ACKNOWLEDGMENTS

Work described in this paper was supported in part by National Science Foundation Grant IBN 92,15069, National Institutes of Aging Grant AG14751, a grant from the Sankyo Company, and funds from the University of Southern California.

The Annual Review of Psychology is online at http://psych.annualreviews.org

LITERATURE CITED

Berry SD, Thompson RF. 1978. Prediction of learning rate from the hippocampal EEG. *Science* 200:1298–300
Chapman PF, Steinmetz JE, Thompson RF. 1988. Classical conditioning does not occur when direct stimulation of the red nucleus or cerebellar nuclei is the unconditioned stimulus. *Brain Res.* 442:97–104
Clark RE, Lavond DG. 1993. Reversible lesions of the red nucleus during acquisition and retention of a classically conditioned...
behavior in rabbits. *Behav. Neurosci.* 107: 264–70
Hiraoka M, Shimamura M. 1977. Neural
mechanisms of the corneal blinking reflex in cats. *Brain Res.* 125:265–75
Thompson RF. 1997. Classical conditioning
has much to do with LTP. Behav. Brain Sci. 20:632–33
# CONTENTS

Frontispiece—Richard F. Thompson  

**PREFATORY**  

In Search of Memory Traces, Richard F. Thompson  

**DECISION MAKING**  

Indeterminacy in Brain and Behavior, Paul W. Glimcher  

**BRAIN IMAGING/COGNITIVE NEUROSCIENCE**  

Models of Brain Function in Neuroimaging, Karl J. Friston  

**MUSIC PERCEPTION**  

Brain Organization for Music Processing, Isabelle Peretz and Robert J. Zatorre  

**SOMESTHETIC AND VESTIBULAR SENSES**  

Vestibular, Proprioceptive, and Haptic Contributions to Spatial Orientation, James R. Lackner and Paul DiZio  

**CONCEPTS AND CATEGORIES**  

Human Category Learning, F. Gregory Ashby and W. Todd Maddox  

**ANIMAL LEARNING AND BEHAVIOR: CLASSICAL**  

Pavlovian Conditioning: A Functional Perspective, Michael Domjan  

**NEUROSCIENCE OF LEARNING**  

The Neuroscience of Mammalian Associative Learning, Michael S. Fanselow and Andrew M. Poulos  

**HUMAN DEVELOPMENT: EMOTIONAL, SOCIAL, AND PERSONALITY**  

Behavioral Inhibition: Linking Biology and Behavior Within a Developmental Framework, Nathan A. Fox, Heather A. Henderson, Peter J. Marshall, Kate E. Nichols, and Melissa A. Ghera  

**BIOLOGICAL AND GENETIC PROCESSES IN DEVELOPMENT**  

Human Development: Biological and Genetic Processes, Irving I. Gottesman and Daniel R. Hanson
CONTENTS

SPECIAL TOPICS IN PSYCHOPATHOLOGY
The Psychology and Neurobiology of Suicidal Behavior,
Thomas E. Joiner Jr., Jessica S. Brown, and LaRicka R. Wingate 287

DISORDERS OF CHILDHOOD
Autism in Infancy and Early Childhood, Fred Volkmar,
Kasia Chawarska, and Ami Klin 315

CHILD/FAMILY THERAPY
Youth Psychotherapy Outcome Research: A Review and Critique
of the Evidence Base, John R. Weisz, Amanda Jensen Doss,
and Kristin M. Hawley 337

ALTRUISM AND AGGRESSION
Prosocial Behavior: Multilevel Perspectives, Louis A. Penner,
John F. Dovidio, Jane A. Piliavin, and David A. Schroeder 365

INTERGROUP RELATIONS, STIGMA, STEREOTYPING,
PREJUDICE, DISCRIMINATION
The Social Psychology of Stigma, Brenda Major
and Laurie T. O’Brien 393

PERSONALITY PROCESSES
Personality Architecture: Within-Person Structures and Processes,
Daniel Cervone 423

PERSONALITY DEVELOPMENT: STABILITY AND CHANGE
Personality Development: Stability and Change, Avshalom Caspi,
Brent W. Roberts, and Rebecca L. Shiner 453

WORK MOTIVATION
Work Motivation Theory and Research at the Dawn of the Twenty-First
Century, Gary P. Latham and Craig C. Pinder 485

GROUPS AND TEAMS
Teams in Organizations: From Input-Process-Output Models to IMOI
Models, Daniel R. Ilgen, John R. Hollenbeck, Michael Johnson,
and Dustin Jundt 517

LEADERSHIP
Presidential Leadership, George R. Goethals 545

PERSONNEL EVALUATION AND COMPENSATION
Personnel Psychology: Performance Evaluation and Pay for Performance,
Sara L. Rynes, Barry Gerhart, and Laura Parks 571
CONTENTS

PSYCHOPHYSIOLOGICAL DISORDERS AND PSYCHOLOGICAL EFFECTS ON MEDICAL DISORDERS
  Psychological Approaches to Understanding and Treating Disease-Related Pain, Francis J. Keefe, Amy P. Abernethy, and Lisa C. Campbell 601

TIMELY TOPIC
  Psychological Evidence at the Dawn of the Law’s Scientific Age, David L. Faigman and John Monahan 631

INDEXES
  Subject Index 661
  Cumulative Index of Contributing Authors, Volumes 46–56 695
  Cumulative Index of Chapter Titles, Volumes 46–56 700

ERRATA
  An online log of corrections to Annual Review of Psychology chapters may be found at http://psych.annualreviews.org/errata.shtml