A Mechanism for Timing Conditioned Responses

John W. Moore
Computer Science Department
University of Massachusetts

Technical Report 92-3

This report is to be published in the proceedings of NATO-ARW *Time, Action and Cognition* held in St. Malo, France, October 22-25, 1991: F. Macar V. Pouthas, and W.J. Friedman (Eds.), Kluwer Acdemic Publishers, Dordrecht, The Netherlands. Preparation of this chapter and the author's research program were supported by grants from the Air Force Office of Scientific Research and the National Science Foundation. I thank Richard Ivry, Department of Psychology, University of California–Berkeley, and Chris Miall, Laboratory of Physiology, Oxford University, for their helpful critiques of an earlier draft.

A MECHANISM FOR TIMING CONDITIONED RESPONSES

J. W. MOORE
Departments of Psychology and Computer Science
University of Massachusetts—Amherst
Amherst, Massachusetts, 01003

ABSTRACT. Classical conditioning procedures instill knowledge about the temporal relationships between conditioned stimuli, which are regarded as predictive signals and triggers for action, and the unconditioned stimulus, the event to be timed. This knowledge is expressed in the temporal features of the conditioned response, which typically develop such that its peak amplitude occurs at times when the unconditioned stimulus is expected. A simple connectionist network, comprised of two neuron-like processing units, provides a mechanism that can acount for virtually all aspects of conditioned response timing. The unfolding of time from the onsets and offsets of events such as conditioned stimuli is represented by the propagation of activity along delay lines. Input to the two processing units from conditioned stimuli arise from collateral taps off of each sequential element of these delay lines.

1. Introduction

Psychologists interested in time, action, and cognition do not typically attend to research on classical conditioning (Block, 1990). One reason for this is that conditioned responses reside in the domain of learning, not cognition and perception. Classical conditioning does not engage the mind—it proceeds beneath the veneer of conscious awareness. Psychologists are interested in impressions and judgements about time. Conditioned responses are time veridical—they reflect real time rather than perceptually distorted impressions of elapsed time. These properties of conditioned responses make them attractive to neuroscientists and computational modellers interested in real-time motor control and its underlying neural mechanisms (Gabriel and Moore, 1990). This community of scholars regards the conditioned response as a microcosm for elucidating behavioral principles of wide generality and for discovering mechanisms for their expression.

This chapter summarizes some facts about the timing of conditioned responses and then presents a neural network model that can accommodate them. From now on, I use the following standard abbreviations: CS for conditioned stimulus, US for unconditioned stimulus, CR for conditioned response, and UR for unconditioned response.

1.1. Conditioning and Cognition

Despite its traditional emphasis on the principles of behavioral learning, it is often convenient to express the outcome of classical conditioning in terms of acquired knowledge.[1] Perhaps the most basic thing learned in classical conditioning is that the CS predicts the US. Prediction is the key to understanding conditioning because a CS's capacity to control behavior depends on the degree to which it is a reliable and nonredundant signal that the US will occur. Subjects learn to ignore stimuli that are poor predictors of the US, and they learn to suppress CRs to stimuli (conditioned inhibitors) that predict the withholding of an otherwise anticipated US.

In addition to expressing knowledge, a CR often possesses the elements of skill. Its topographical features—latency, rise-time, peak amplitude—typically vary from one set of procedures to the next in such a way that they are appropriate to the 'task demands' imposed by training parameters (Levey and Martin, 1968). The main evidence for this adaptive character of the CR is that these topographical features vary systematically with the CS–US interval employed in training. (The CS–US interval is typically abbreviated ISI, for 'interstimulus interval'.) In particular, eye blink CRs are 'temporally adaptive.' Temporal adaptability simply means that the peak amplitudes of CRs occur within a restricted temporal window that also contains the US. In this sense, CRs reflect the knowledge that the US occurs at a specific time after the CS. Recent evidence for this assertion is reviewed later. For now, it suffices to say that subjects learn not only to expect the US in the presence of the CS, but also when to expect it.

1.2. Mechanisms of Knowledge Acquisition and Expression

Classical conditioning procedures establish knowledge about timing. What are the mechanisms that bring this about? What happens in the brain that might explain the temporally adaptive properties of a conditioned eye blink, for example? Learning theorists and neuroscientists alike believe that the knowledge instilled by conditioning, and the accompanying rules for generating CRs, arise from 'associative mechanisms.' These mechanisms are captured by mathematical models that take the form of rules for changing the strength of the 'synaptic connection weights' between representations of the CS and the CR. Such rules consist of two factors—one factor is the level of CS processing; the other factor is the level of US processing (Desmond, 1990; Dickinson and Mackintosh, 1978; Sutton and Barto, 1990; Rescorla, 1988).

---

[1] Saying that conditioning procedures produce knowledge does not contradict the earlier statement that CRs arise from unconscious processes (Kihlstrom, 1987).

2. CR Waveforms and CS–US Intervals

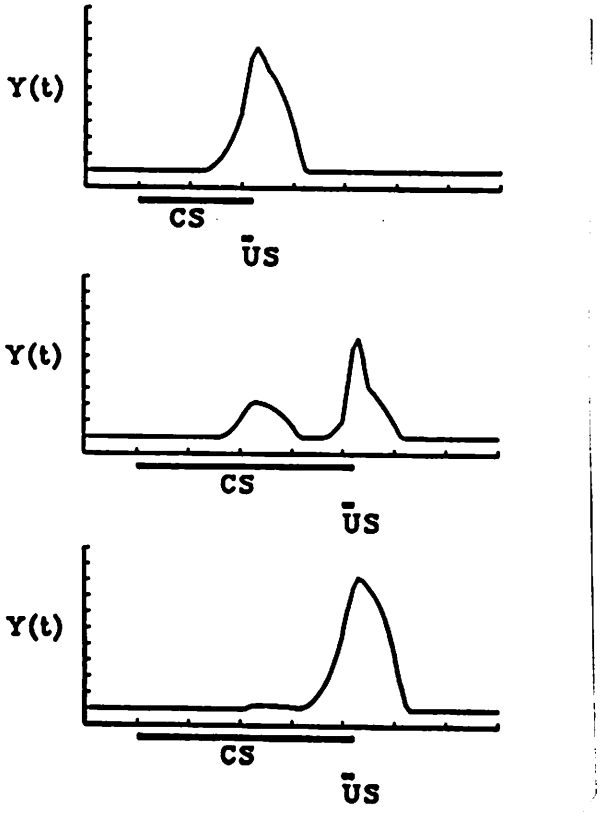

Figure 1. Effects of changing US timing on simulated CR topography, Y(t). Top: Topography after 25 short CS–US interval training trials in Stage 1. Middle: In Stage 2, the CS–US interval is lengthened. After 10 trials, the short-latency CR is somewhat diminished and the longer-latency CR begins to develop. Bottom: After 30 Stage 2 trials, tranformation of CR topography is nearly complete. Copywrite 1989, Springer-Verlag.

Conditioned response waveforms are functions of the CS–US interval used in training. Long CS–US intervals give rise to delayed CRs, which Pavlov attributed to an active inhibitory process which persists until the US is imminent ('inhibition of delay'). There is little evidence to support Pavlov's explanation of the delayed temporal placement of CRs. Instead of being due to inhibition, the phenomenon of delayed CR placement is more likely attributable to the fact that a CR generated by a long CS–US interval is a *different* response from that generated by a shorter CS–US interval.

That a CR generated with one CS–US interval can be regarded as being a different response from one generated with a different CS–US interval is attested to by studies in which subjects are trained with a short CS–US interval and then shifted to a longer CS–US interval. Typically, the CR in the temporal window defined by the short CS–US interval undergoes extinction while a new CR emerges in a window defined by the longer CS–US interval. It is *not* the case that the original CR migrates to the new temporal window. It remains within its original window but progressively loses amplitude while at the same time the new CR emerges in the time window defined by the new US locus. This is illustrated in Figure 1, which is a simulation from the VET neural network model discussed later on.

Additional evidence for the temporal specificity of CRs comes from a widely cited study of rabbit eye blink/nictitating membrane conditioning by Millenson, Kehoe, and Gormezano (1977). Their training protocol consisted of mixing two CS–US intervals, one of 400 msec and another of 700 msec. There were occasional 'probe trials' on which the CS was presented without the US for either 400 or 700 msec. On 400-msec probes, response topography showed a peak appropriate to the 400-msec CS–US interval. On 700-msec probes, response topography revealed two peaks—one appropriate for the shorter interval and another appropriate for the longer interval. This finding implies that subjects learned *If the CS extends beyond 400 msec, initiate another CR; if it does not, do nothing.* However one chooses to phrase it, the implication remains that subjects learned not one CR but two, one appropriate for each CS–US interval it experienced.

2.1. CR Waveforms and Trace Conditioning

The Millenson et al. (1977) experiment used a delay conditioning protocol, which is technical jargon for the fact that, on training trials, CS onset preceded the US and stayed 'on' until the US occurred, at which time it went 'off.' But what of trace conditioning protocols in which CS onset precedes the US but goes off beforehand? Where do subjects place their CRs: during the CS's 'on' phase or during its 'off' phase? As in the case of delay paradigms, CRs occur near the US, so CRs occur during the 'off' phase.

In trace conditioning there are two possible CS–US intervals, one defined by CS onset and the other defined by CS offset. Hence, two things might be learned: the US follows CS onset; the US follows CS offset. Which event, CS onset or offset, defines the temporal window in which to place CRs? The answer is *both*, although this is constrained by such details as the CS's duration and the time between its offset and the US, the so-called trace interval. Support for this conclusion comes from an experiment

by Desmond and Moore (1991a). They trained rabbits using a trace conditioning protocol in which the CS was a 150-msec tone followed 200 msec later by the US, giving a nominal CS–US interval of 350 msec. The US was a mild eye shock, and the CR was extension of the nictitating membrane. A second group of animals were trained in a delay conditioning procedure in which the tone was of 350 msec duration.

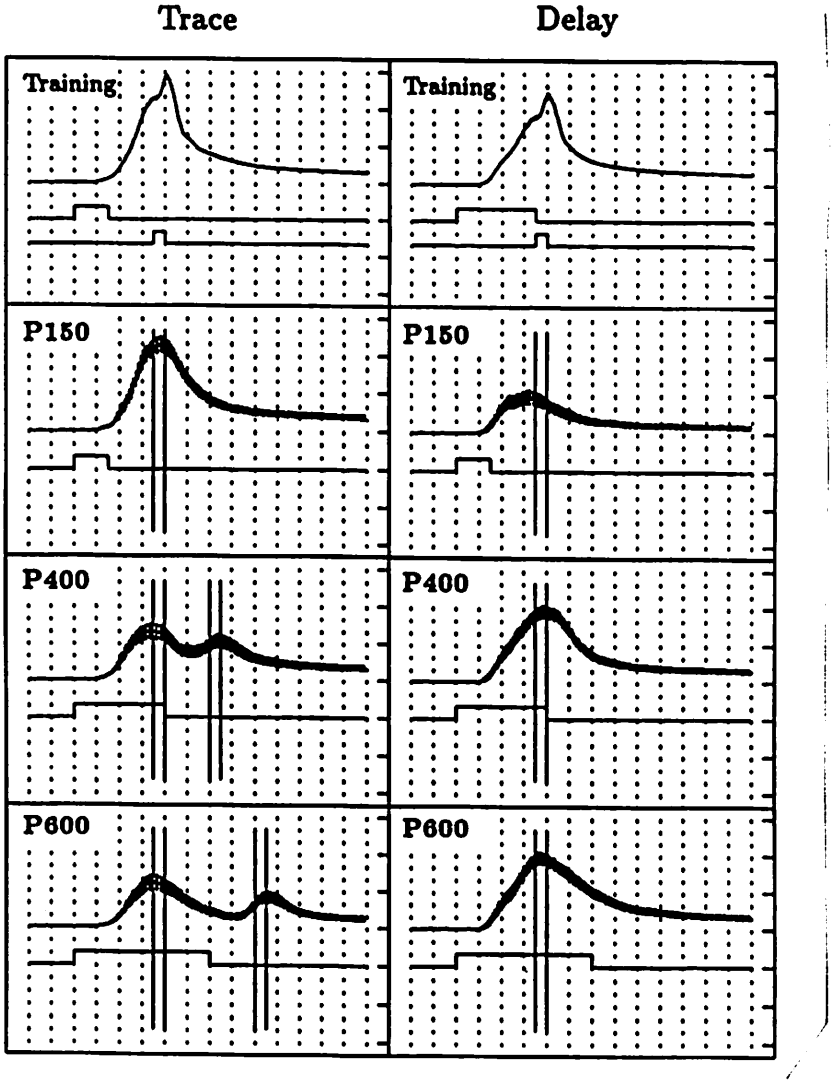

Figure 2. Average CR waveforms from training and probe trials for the Trace and Delay groups of Desmond and Moore (1991a). Fine-grain bars on waveforms during probe trials are standard error bars. Vertical cursors correspond to expected times of the US with respect to CS onset and offset. Copywrite 1991, Springer-Verlag.

After training, subjects in both the Trace and Delay groups were given probe trials without the US. These consisted of presentations of the tone for durations of 150, 400, and 600 msec. As Figure 2 shows, the two longer duration tones often resulted in bimodal (double peaked) CR waveforms in the Trace group, but not in the Delay group. The initial peak was located 350 msec after tone onset, within the temporal window defined by the interval between CS onset and the US. The second peak was located 200 after tone offset, within the temporal window defined by the interval between CS offset and the US. This second peak, though inappropriate for the nominal CS–US interval of 350 msec, is appropriate for the CS–US interval of 200 msec defined in terms of tone offset. Thus, for example, on a typical 600-msec probe trial, one peak appeared 350-msec after tone onset and the other appeared 800 msec after tone onset. Hence, the following knowledge was acquired: The US follows CS onset by 350 msec; The US follows CS offset by 200 msec. Let us turn now to considering how the 'motor program' *Initiate a CR such that peak amplitudes correspond to the times of the US* might be derived from this knowledge.

3. The VET Model

Desmond and Moore (1988) proposed a neural network model capable of simulating the features of CR timing described above (see also, Desmond, 1990; Moore, Desmond, and Berthier, 1989). We refer to this model by the mnemonic VET in order to emphasize its function of mapping associative *values* onto action based on *expectancies* about *timing*. The model assumes that CSs trigger propagated activity in the nervous system. In its simplest form, this activation can be represented by a delay line. A 'tap' or collateral from each element of the delay line encodes the time after the activation has been triggered by the CS. Each potential CS has its own set of delay line elements which are anatomically associated with its modality.[2] In addition to time-tagged stimulus elements, there are two processing units where learning occurs. One unit associates active stimulus elements with the US and passes this information to the other unit, which uses this information to generate appropriate (adaptively timed) CR waveforms. The assumed delay-line representation of time enables these two processing units, which are thought to reside within cerebellar cortex, to treat CS-initiated input as a sequence of discrete events.

In addition to the simulation shown in Figure 1 and the results of the Desmond and Moore (1991a) trace conditioning experiment (Figure 2), the VET model correctly predicts that, as in the case of delay conditioning, CR waveforms in trace conditioning

[2]Desmond (1990) describes an extension of the simple delay line representation of the CS into a planar array. The planar array representation can encompass stimulus generalization and discrimination.

peak at the point of US onset, that is, within the trace interval. The model also predicts that, with long CS–US intervals, CRs do not begin until the US is imminent, the phenomenon that Pavlov attributed to 'inhibition of delay.' Finally, the model predicts the outcome of experiments with multiple CS–US intervals such as the study by Millenson et al. (1977), which showed that training with randomly mixed trials having CS–US intervals of 200 and 700 milliseconds gives rise to CRs with two peaks, each centered at a time of US onset.

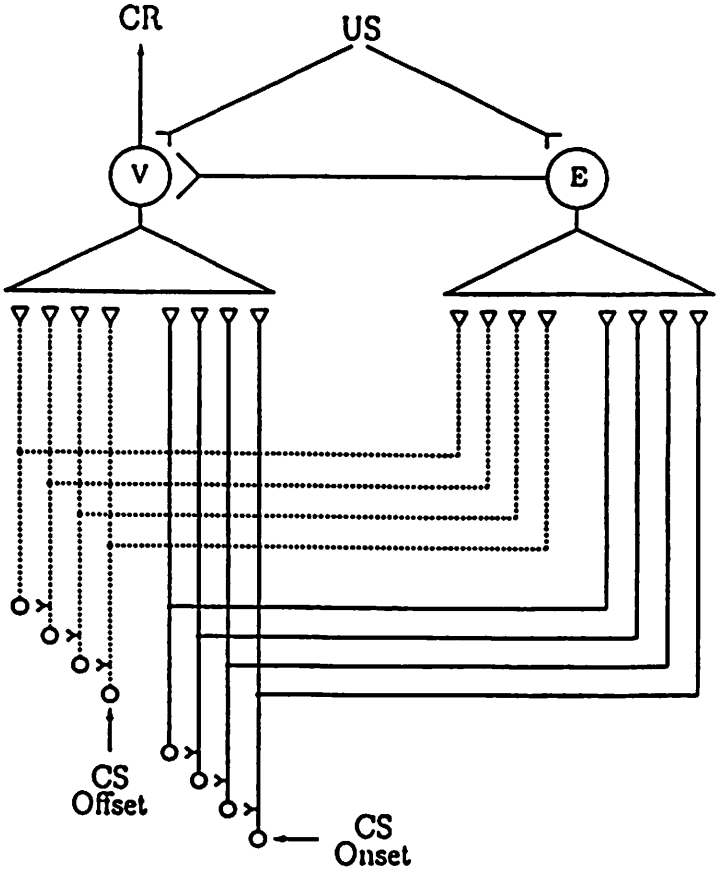

Figure 3. Diagram of the VET network. CS onset and CS offset are assumed to activate separate tapped delay lines that project to the V and E units, as explained in the text. Copywrite 1991, Springer-Verlag.

The structural components of the VET model are depicted in Figure 3. The two neuron-like processing units receive convergent input from CSs and the US. The V-

unit is the output device that generates CR topography. It has modifiable synaptic weights that are changed according to a competitive learning rule. Weight changes depend on local 'eligibility' factors, a global parameter dependent upon the ISI, and a reinforcement signal that reflects the expected time of occurrence of the US. The learning (weight update) rule contains two reinforcement factors: one is contributed directly by the US; the second is contributed by the E unit, which learns when the US occurs with respect to CS onsets and offsets. Both must exceed zero for weight changes on the V unit to occur.

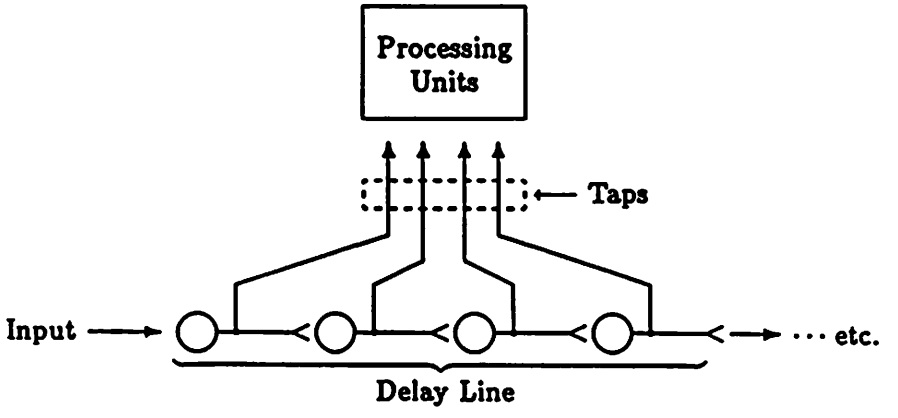

Figure 4. Basic tapped delay line. Injection of CS input begins sequential propagation of signal through a delay line. Each 'synapse' introduces a delay; the total delay from activation of the first element of the delay line to the last element is a direct function of the number of intervening sequential 'synapses' and the conduction speed of propagated activity in the cascade. Taps from the delay line elements send time-tagged information to higher order processing units. Copywrite 1988, Springer-Verlag.

Like the V unit, the E unit receives convergent input from CSs and the US, and it has modifiable synaptic weights that are changed according to a simple linear difference equation that includes local eligibility factors and the global ISI parameter. By providing a precisely timed positive signal to the V unit, the E unit prevents the eventual extinction of positive weights from input elements to the V unit, thereby permitting CR waveforms to anticipate the US. Without this mechanism, the output of the V unit would over the course of training be positive only within time steps that also contain the US.

Conditioned stimuli are provided with a temporal dimension through tapped delay lines (Figure 4) that encode, not only the source of the stimulus (e.g., a particular

component of a compound CS), but also the time since the stimulus began.

Another set of tapped delay lines encodes the time since the stimulus ceased. Thus, the model assumes the existence of separate and independent timed-tagged input elements for both stimulus onset and offset. Figure 5 illustrates how activation of delay line elements are triggered by CS onset and offset.

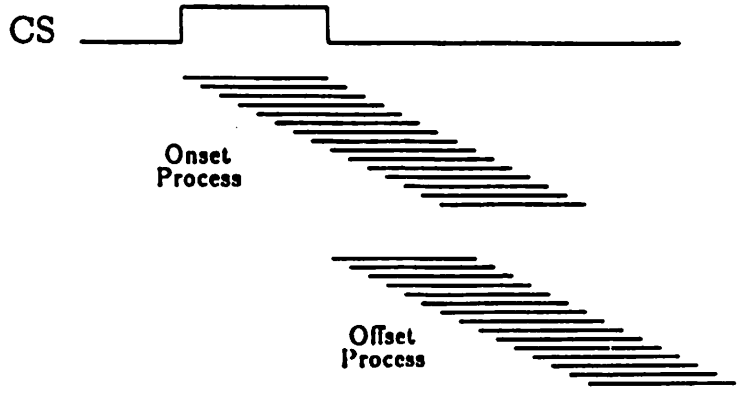

Figure 5. Onset and offset processes for a single CS. Time progresses from left to right. Each horizontal bar below the CS represents the activation time of an individual input element. The figure illustrates the overlapping activation times of individual elements in the onset and offset processes. This feature of the model allows for continous and 'smooth' ramping of the CR from a zero baseline position to its peak amplitude at the expected time of the US, as illustrated in Figures 1 and 2. Copywrite 1988, Springer-Verlag.

4. Brain Implementation of the VET Model

Because of their crucial involvement in eye blink conditioning, we sought to align the VET model with the cerebellum and associated brainstem structures (Moore et al., 1989). We hypothesize that E units, which learn when the US occurs, are Golgi cells and that V units, which are Purkinje cells, use this information to generate an output which ultimately produces a temporally adaptive CR. In brief, Golgi cells learn when USs occur in relation to CS onsets and offsets, and Purkinje cells learns how to generate appropriately timed CRs and their topographical features.

Where does the knowledge that *The CS predicts the US* arise, and what role does it play in the development of adaptively timed CRs? We have suggested that this knowledge comes about through simple Hebbian learning among brainstem neurons.

These neurons provide a 'coarsely coded' version of the CR—one lacking the temporal specificity of the real thing.[3] Their activation by the CS is conveyed to cerebellar cortex where it is fashioned into the appropriate 'finely coded' CR waveform. Specifically, the activation of the brainstem neurons instantiates *The CS predicts the US.* This activation is manifest as a burst of firing that persists for at least the duration of the CS–US interval. When projected to the granule cell layer of the cerebellum, it is intercepted by Golgi cells (E units) at the mossy fiber/granule cell interface.

The Golgi cells have learned to release their normal inhibitory hold on input from the brainstem neurons but only momentarily and at times relative to the CS when the US has occurred in the past. This release of inhibition permits the activation from the brainstem neurons to proceed via parallel fibers to the Purkinje cells (V units) where it can reinforce synaptic modifications of active inputs from the tapped delay line mechanism. In other words, Golgi cells encoding *The US follows the CS by x amount of time* interact with activation arising from the brainstem neurons in such a way as to provide the temporal specificity needed to instruct the Purkinje cells to *Initiate a CR such that the peak amplitude occurs at the time of the US.*

The delay lines illustrated in Figures 3 and 4 are not in the cerebellum but are extrinsic to this structure. Although their location has not been specified or experimentally determined, their existence ought to be evident in the firing of neurons that project to the cerebellum. The most likely place to find such evidence would be in the CS-evoked activity recorded from neurons of the pontine nuclei, which is a major source of mossy fiber input to the cerebellar cortex. This activity would tell us whether information is sent to the cerebellum in the manner implied by Figures 4 and 5. However, it would not inform us about which precerebellar structures are involved in its manifestation—a task that could be approached with fiber-tracing methods.

One candidate for tapped delay lines is the reticular core of the brain, as has been suggested by Scheibel and Scheibel (1967). Reticular formation neurons can fire sustained bursts to a CS (Richards, Ricciardi, and Moore, 1991), they provide a wide range of possible propagation speeds, which are subject to modulation by local and distal processes, and their axons show extensive collateralization, which could provide

[3]Desmond and Moore (1986) reported that the brainstem contains neurons that behave in the manner imagined in this scheme. For example, CS-evoked firing in some supratrigeminal reticular formation neurons predict CR amplitude, across a series of trials, but not CR latency. Some red nucleus neurons behave in this manner (Desmond and Moore, 1991b). By contrast, many cells in the deep cerebellar nuclei show CS-evoked firing patterns that are highly predictive of both CR amplitude and latency (Berthier, Barto, and Moore, 1991).

the taps depicted in Figure 4.

5. Acknowledgement

Preparation of this chapter and the author's research program were supported by grants from the Air Force Office of Scientific Research and the National Science Foundation. I thank Richard Ivry and Chris Miall for their helpful critiques of an earlier draft.

6. References

1. Berthier, N. E. Barto, A. G. and Moore, J. W. (1991) 'Linear systems analysis of the relationship between firing od deep cerebellar neurons and the classically conditioned nictitating membrane response in rabbits,' Biological Cybernetics 65, 99-105.

2. Block, R. A. (1990) Cognitive Models of Psychological Time, Erlbaum, Hillsdale, N. J.

3. Desmond, J. E. (1990) 'Temporally adaptive responses in neural models: The stimulus trace,' in M. Gabriel and J. Moore (eds.), Learning and computational neuroscience: Foundations of adaptive networks, MIT Press, Cambridge, MA., pp. 421-456.

4. Desmond, J. E. and Moore, J. W. (1988) 'Adaptive timing in neural networks: The conditioned response,' Biological Cybernetics 58, 405-415.

5. Desmond, J. E. and J. W. (1991a) 'Altering the synchrony of stimulus trace processes: Tests of a neural-network model,' Biological Cybernetics 65, 161-169.

6. Desmond, J. E. and Moore, J. W. (1986) 'Dorsolateral pontine tegmentum and the classically conditioned nicitating membrane response: analysis of CR-related single-unit activity,' Experimental Brain Research 65, 59-74.

7. Desmond, J. E. and Moore, J. W. (1991b) 'Single-unit activity during the classically conditioned rabbit nictitating membrane response,' Neuroscience Research 10, 260-279.

8. Dickinson, A. and Mackintosh, N. J. (1978) 'Classical conditioning in animals,' Annual Review of Psychology 29, 587-612.

9. Gabriel, M. and Moore, J. (1990) Learning and Computational Neuroscience: Foundations of Adaptive Networks, MIT Press, Cambridge, MA.

10. Gormezano, I. Prokasy, W. F. and Thompson, R. F. (eds.) (1987) Classical conditioning, 3rd edition, Erlbaum, Hillsdale, N. J.

11. Kihlstrom, J. F. (1987) 'The cognitive unconscious,' Science 237, 1445-1452.

12. Levey, A. B. and Martin, I. (1968) 'Shape of the conditioned eyelid response,' Psychological Review 75, 398-408.

13. Millenson, J. R. Kehoe, E. J. and Gormezano, I. (1977) 'Classical conditioning of the rabbit's nictitating membrane response under fixed and mixed CS–US intervals,' Learning and Motivation 8, 351-366.

14. Moore, J. W., Desmond, J. E. and Berthier, N. E. (1989) 'Adaptively timed conditioned responses and the cerebellum: A neural network approach,' Biological Cybernetics 62, 17-28.

15. Rescorla, R. A. (1988) 'Behavioral studies of Pavlovian conditioning,' Annual Review of Neuroscience 11, 329-352.

16. Richards, W. G. Ricciardi, T. N. and Moore, J. W. (1991) 'Activity of spinal trigeminal pars oralis and adjacent reticular formation units during differential conditioning of the rabbit nictitating membrane response,' Behavioural Brain Research 44, 195-204.

17. Scheibel, M. E. and Scheibel, A. B. (1967) 'Anatomical basis of attention mechanisms in vertebrate brains,' in G. C. Quarton, T. Melnechuk, and F. O. Schmitt (eds.), The neurosciences: A study program, Rockefeller University Press, New York, NY, pp. 577-602.

18. Sutton, R. S. and Barto, A. G. (1990) 'Time-derivative models of Pavlovian reinforcement,' in M. Gabriel and J. Moore (eds.), Learning and computational neuroscience: Foundations of adaptive networks, MIT Press, Cambridge, MA., pp. 497-537.