PARAMETER STABILITY IN THE TD MODEL FOR COMPLEX CR TOPOGRAPHIES

J. W. Moore*, J. S. Marks, V. E. Castagna, and R. J. Polewan

Neuroscience and Behavior Program, University of Massachusetts, Amherst, MA

642.2

INTRODUCTION

A long-time goal of quantitative models of classical conditioning has been to describe the topography of the CR, its temporal dynamics following presentation of a conditioned stimulus (CS). Several computational models have been put forward in recent years to this end. (For a recent review, see Brandon, Vogel, & Wagner, in press).

One of the more promising computational approaches to CR topography is Sutton and Barto's (1990) Temporal Difference Model with the Complete Serial Compound Implementation, henceforth referred to as the TD (CSC) model. The model assumes a 'spreading activation' mechanism for representing the elapsed time triggered by the onset of a CS. For classical eyeblink conditioning, Moore and Choi (1997) devised a scheme whereby the TD model might be implemented in the cerebellum, where the basic processes of eyeblink conditioning are thought to occur (Hesslow and Yeo, in press; Schreurs and Alkon, in press).

For the TD model to describe CR timing and topography in complex paradigms involving multiple CSs. Both CS onsets and offsets are assumed to trigger cascades of spreading activation. This spreading activation is mapped onto the serial components of the TD model. Each nominal CS, such as a tone and a light, initiates an independent cascade which sequentially activates the variables X_i in the model (i = 1,2,3,...). The duration of activation of a serial component need not be fixed or constant, but for simulation purposes we have often assumed a temporal grain of 10 ms. Hence, this is the assumed duration of activation of a serial component. When activated, X_i = 1; when inactivated (after 10 ms), X_i resets to a baseline of 0 as the next serial component, X_{i+1} is activated. Although X_i is no longer active and therefore no longer contributes to Y, the output or response, its connection to the output, V_i, remains eligible for modification over succeeding time-steps. Eligibility decays at a rate determined by delta (δ) in Equation 3 below. And, as mentioned, just as a nominal CS initiates a cascade of activation among serial components, so too does its offset. The two cascades are assumed to operate independently and in parallel. There are limits on how long these cascades might last, i.e., on the number of sequentially activated elements in each cascade. The only requirement is that these cascades span the CS–US intervals employed in training.

The TD model is governed by the equation:

$$V_i(t) = \beta[\lambda(t) - \gamma Y(t) - Y(t-1)] \times \alpha \bar{X}_i(t) \qquad (1)$$

where

$$Y(t) = \sum_i V_i(t) X_i(t) \qquad (2)$$

$\bar{X}_i$ refers to eligibility for modification. Subscript j indexes the set of CSs that are active at any timestep.

$$\bar{X}_i(t+1) = \bar{X}_i(t) + [X_i(t) - \bar{X}_i(t)] \qquad (3)$$

TD MODEL PARAMETERS

As can be seen from Equations 1-3, the TD model generates CRs as a function of training in terms of 4 free parameters. Each parameter has its own role and interpretation:

Alpha-Beta ($\alpha\beta$): A real-valued function of CS salience mapped onto the semi-closed unit interval (0, 1]. In the cerebellum, alpha has been associated with the magnitude of mossy-fiber input to the granule cell layer and the activation of parallel fiber input to Purkinje cells via AMPA other receptors.

Gamma (γ): In the TD model, gamma is the 'discount parameter.' It is a real-valued function mapped onto (0, 1] which represents capacity for S-R based higher-order conditioning among successive temporally distinct mossy-fiber inputs to the cerebellum within a trial or CS presentation. As can be seen in Figure 1, high values of gamma are associated with larger amplitude CRs.

Delta (δ): In the TD model, delta controls the decay of 'synaptic eligibility.' This is the time after activation by the CS that successive inputs to the cerebellum can be modified to control parameters of CR topography. Delta is mapped onto (0, 1]. Low values of delta result in greater temporal spread of learning after CS onset and a concomitant reduction of peak CR amplitude.

Lambda (λ): A real-valued increasing function of US effectiveness. It was held constant (lambda = 1.0) in the simulations of CR topographies described later on. In classical eyeblink conditioning, where the cerebellum is known to play a central role in learning and CR topography (Hesslow and Yeo, in press), lambda has been associated with climbing fiber activation Purkinje cells.

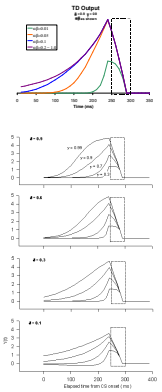

Figure 1 shows simulated CRs, Y(t), after 200 acquisition trials with a CS–US interval of 250 ms and with a 50-ms US (rectangles scaled for Y(t) with lambda = 1.0). The time step for these simulations was 10 milliseconds. Panel A shows the influence of alpha-beta on CR topography with constant values of gamma and delta. Notice that values of alpha-beta greater than 0.2 are essentially equivalent with gamma and delta equal to 0.9. Panels B-E are simulated CRs as joint functions of gamma and delta with alpha-beta constant at .05.

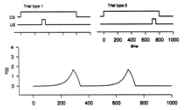

Figure 2 shows a simulation of bimodal CRs obtained for a random mixture of two CS–US intervals. There were 100 trials of each CS–US interval. The simulated CR is for a non-reinforced probe trial after 200 trials of training.

This poster describes experiments in rabbit eyeblink conditioning protocols designed to create complex bimodal CRs. One way to ensure complex CRs was to employ a mixed CS–US interval training protocol, the temporal uncertainty paradigm, together with a Kamin blocking design in which we assessed the effects on model parameters of pretraining to one CS at one CS–US interval on CRs generated in a second stage of compound conditioning with a random mixture of two CS–US intervals.

The question of interest concerns the stability of the model's parameters in classical eyeblink conditioning. That is, to what extent do best-fitting parameter values change as a function of amount of training and in relation to CSs of different modalities and susceptibility to blocking because of prior training in one CS?

METHODS

Subject and Apparatus: The subjects were 24 naive albino rabbits, weighing approximately 2 kg each at the start of the experiment. The animals were run two at a time in soundproofed file drawers while restrained in Plexiglas boxes. A laboratory computer controlled the presentation of CSs and the US. Conditioned responses were measured on non-reinforced test trials. Each subject was run for 2 sessions per day, 5 days a week for a total of 10 sessions per week. Each session lasted 30 minutes.

The light CS (CS1) was 800 ms in duration and was delivered by two 9 vdc incandescent lamps. The tone CS (CS2) was 70 dB SPL and 800 ms in duration. The US consisted of a periorbital electric shock of 0.2 mA and 50 ms in duration (5 dc pulses of 3.5 ms duration at 100 Hz).

Procedure and Design: Stage 1 consisted of 20 sessions spaced over 10 days. Only 1 Experimental and 1 Control animal were run together. Each session in stage 1 consisted of 60 trials of a delay-conditioning paradigm. On 54 trials CS1 was paired with the US occurring at 300 ms after CS onset and on 6 trials, presented randomly, the CS was presented alone. Sit Controls did not receive the CS or US presentations during stage 1.

Stage 2 consisted of 20 to 30 sessions spaced over 10 to 15 days. All rabbits received Stage-2 training. Each session consisted of 60 trials made up of 5 trial types. Trial type 1 (TL700) consisted of CS1 and CS2 presented concurrently for 800 ms with the US presented at 700 ms after CS onset. Trial type 2 (TL300) consisted of CS1 and CS2 presented concurrently for 800 ms with the US presented 300 ms after CS onset. CS1 and CS2 were presented concurrently on trial type 3 (TL-) in the absence of the US. CS2 was presented for 800 ms in the absence of the US on trial type 4 (T-), and CS1 was presented for 800 ms in the absence of the US on trial type 5 (L-). In both Stages 1 and 2, the intertrial interval (ITI) varied randomly from 25 to 35 seconds.

Experiment	CS Length	Stage 1	Stage 2
I	Long	L300	TL700, TL 300, TL-, L-, T-
II	Long	L700	TL700, TL 300, TL-, L-, T-
III	Long	L300	TL700, TL 300, TL-, L-, T-
IV	Short	L300	TL700, TL 300, TL-, L-, T-
IVb	Short	L300	TL700, TL 300, TL-, L-, T-
V	Short	L300	TL700, TL 300, TL-, L-, T-
VI	Short	T700	TL700, TL 300, TL-, L-, T-

(Above) A table showing the six experiments in this study. (Experiment IV was repeated due to data loss caused by a computer malfunction). Long CSs were 800 ms in length and short CSs were 300 ms. Stage 1 trial types are as follow: L300, light reinforced at 300 ms after CS onset; L700, light reinforced at 700 ms after CS onset; T700, tone reinforced at 700 ms after CS onset. In Stage 2 the trial types are as follow:TL700, tone and light reinforced at 700 ms after CS onset; TL300, tone and light reinforced at 300 ms after CS onset; TL-, compound (tone-light) probe trial; L-, light probe trial; and T-, tone probe trial.

Three parameters of the TD (CSC) model were fitted to CRs from individual rabbits for each of four blocks of 5 sessions of Stage-2 training, using 'brute force' searches with unconstrained variations of alpha, delta and gamma. Lambda was assigned a value of 1.0 for all simulations and parameter searches, and the simulation time step was 4 milliseconds. Simulations were confined to the single-CS test trials (L- and T-). The simulation for each block of 5 sessions were initialized with the associative values from the immediately preceding block of sessions. For Experimental rabbits, the values for the first phase of Stage-2 came from the last block of 5 Stage-1 sessions. For Control rabbits, Stage-2 values were initialized to 0. Although some subjects were trained with a 300-ms CS, none of the simulations assumed an active 'Offset Process' (see Moore and Choi, 1997). Best-fitting parameters were determined with a least-squares criterion. The best-fitting parameters for each rabbit and block of 5 sessions were subjected to ANOVA in order to assess the significance of variations in parameters as a function of experimental conditions and blocks of sessions of training.

RESULTS

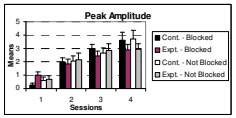

Figure 3 shows average peak CR amplitudes as a function Stage 2 sessions for Blocked and Not Blocked CS and for both Experimental and Control rabbits. The most salient feature of the figure is the progressive increase in peak CR amplitude across sessions for both Experimental and Control rabbits, indicative of CR acquisition.

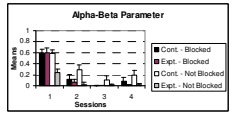

Figure 4 shows that best-fitting values of alpha-beta decreased progressively over sessions for all experimental treatments. A decrease in alpha-beta reflects a decline in of CS salience.

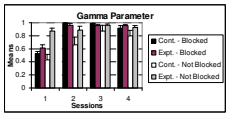

Figure 5 shows that best-fitting values of gamma increased from the first block of 5 sessions to the second, but remained relatively stable beyond that point. Gamma reflects the capacity of later serial components of a CS to reinforce (through higher-order conditioning) earlier serial components.

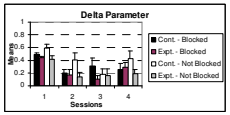

Figure 6 shows that best-fitting values of delta decreased early in training but showed some recovery later on. Delta reflects the capacity for early serial components of a CS to be reinforced by subsequent serial components.
Decreasing delta means that early serial components become progressively more susceptible to reinforcement by later serial components and the US. Later increases in delta may reflect 'inhibition of delay.'

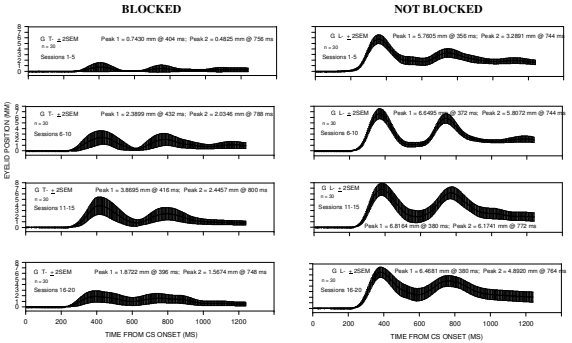

Figure 7 shows average CR waveforms to the Blocked and Not Blocked CSs over the four 5-session blocks of Stage-2 training for an individual rabbit in the Experimental Group. These data are typical recording in that they reflect the progressive decrease in alpha-beta and delta and corresponding increase in gamma. For the Blocked CS (tone), best-fitting values of alpha-beta across four blocks of sessions were .688, .004, .004, .016. The best-fitting values of gamma for these blocks of sessions were .496, .996, .996, .969. The best-fitting values of delta for these blocks of sessions were .400, .004, .004, .047. For the Not Blocked CS (pretrained light), best-fitting values of alpha-beta across four blocks of sessions were .012, .004, .004, .008. The best-fitting values of gamma for these blocks of sessions were .965, .996, .996, .980. The best-fitting values of delta for these blocks of sessions were .059, .004, .004, .023.

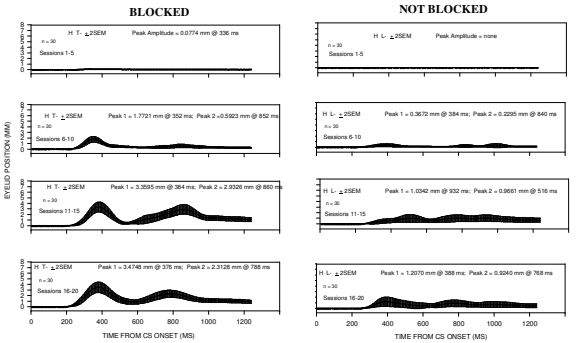

Figure 8 shows average CR waveforms to the Blocked (tone) and Not Blocked (light) CSs over the four 5-session blocks of Stage-2 training for the individual Control rabbit that was run at the same time (yoked) as the Experimental rabbit whose data appears in Figure 7. For the Blocked CS (tone), the best-fitting values of alpha-beta across four blocks of sessions were .844, .004, .012, .016. The best-fitting values of gamma for these blocks of sessions were .496, 1.0, 1.0, .988. The best-fitting values of delta for these blocks of sessions were .520, .465, .008, .020. For the Not Blocked CS (light), the best-fitting values of alpha-beta were .746, .641, .012, .008. The best-fitting values of gamma were .004, .445, 1.0, .984. The best-fitting values of delta for these blocks of sessions were .900, .492, .008, .031.

SUMMARY

Parameters of the TD (CSC) model that best fit complex eyeblink CRs do not remain invariant over training and experimental conditions. Rather, they were found to vary with amount of training and other experimental treatments.

Because of the potential for aligning TD (CSC) model parameters with computational compartments within the cerebellum and its cellular constituents, knowledge of variation in the model's parameters can point the way toward understanding how cellular processes controlling CS salience and higher-order conditioning operate on the scale of the hundreds of milliseconds that constitute a typical trial of eyeblink conditioning.

To date, neither Gamma nor Delta have been aligned with cerebellar physiology (Schreurs and Alkon, 2001), but recent computational models of cellular signaling in the cerebellum provide a number of plausible scenarios. (See Kuroda, Schweighofer, & Kawato, 2001).

REFERENCES

Brandon, S., Vogel, E. H., & Wagner, A. R. (In press). Computational theories if classical conditioning. In J. W. Moore (Ed), A neuroscientist's guide to classical conditioning. New York, Springer-Verlag.

Hesslow, G. & Yeo, C. H. (In press). The Functional Anatomy of Skeletal Conditioning. Computational theories if classical conditioning. In J. W. Moore (Ed), A neuroscientist's guide to classical conditioning. New York, Springer-Verlag.

Kuroda, S., Schweighofer, N. & Kawato, M. (2001). Exploration of signal transduction pathways in cerebellar long-term depression by kinetic simulation. Journal of Neuroscience, 21, 5693-5702.

Moore, J.W. & Choi, J.S. (1997). Conditioned response timing and integration in the cerebellum. Learning and Memory, 4, 116-129.

Schreurs, B.G. & Alkon, D.L. (In press). Cellular mechanisms of classical conditioning. Computational theories if classical conditioning. In J. W. Moore (Ed), A neuroscientist's guide to classical conditioning. New York, Springer-Verlag.

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

Website: http://moorelab.sbs.umass.edu **Supported by: *NIMH MH578936* and *NSF ECS9980062***