

Gradus ad parnassum: Ascending strength gradients or descending memory traces?

Peter R. Killeen

Department of Psychology, Arizona State University, Tempe, AZ 85287-1104.
 killeen@asu.edu <http://tops.asu.edu/dresearch/blab.html>

Abstract: Decay gradients are usually drawn facing the wrong direction. Righting them emphasizes the role of stimuli that mark the response, and leads to different inferences concerning the factors controlling response–reinforcer associations. A simple model of the concatenation of stimulus traces provides some insight to the problems of impulse control relevant to ADHD.

The target article constitutes an important synthesis of behavioral and biological causal factors for ADHD. It, and the precommentary, offers the promising and provocative hypothesis that, *inter alia*, dopamine deficits shorten and steepen the delay of reinforcement gradient, a hypothesis that organizes many of the data. In this commentary, I suggest a clarification of that key hypothesis.

Gradients are often drawn as in Figure 1, top (see the target article’s Fig. 7 and Catania’s Fig. 2). Such representations too easily lead the eye, and then the mind, to see reinforcement acting backward in time. But that can only happen through a history of pair-

ing precursors with reinforcement, so that they become conditioned avatars of primary reinforcement. The delay gradient drawn as a fading trace of the response (the bottom panel of Fig. 1) gives a fairer picture of the process. It is not so much that a delayed reinforcer weakens over time as that the memory of the initiating response weakens, giving reinforcement less signal on which to operate among the buzz of other traces.

This is a difference that makes a difference. At a delay of 20 seconds, doubling the magnitude of reinforcement might improve conditioning; but the trace of the response is so weak compared to more recent stimuli and responses that much of that increased magnitude is more likely to benefit behavior other than the target response. Contrast this with operations that change the strength of the response trace. Doubling memorability at the time of the response will double memorability 20 seconds later. Even though the absolute increment at 20 sec will be much less than at 0 sec, all of it will be vested in the target response. Conversely, in situations where memorability of the response is degraded (Bottom curve, Fig. 1), the trace, and thus the reinforcer’s ability to strengthen the response, may fall below the noise level.

The literature supports this distinction. Lieberman et al. (1985) showed that the presence of a light flash after a response greatly enhanced acquisition. Williams (1999) showed that such marking was much more effective in the differential acquisition of a response than having the same marker signal onset of reinforcement – and thus act as a conditioned reinforcer. In fact, the conditioned reinforcer impeded conditioning. There are three morals to this story.

1. Marking a response when it is made can facilitate conditioning.

2. Bridging stimuli intended as conditioned reinforcers might actually shorten the reach of reinforcement rather than lengthen it, as desired for behavior therapy of ADHD.

3. Dopamine released at the time of reinforcement is more likely to strengthen consummatory rather than instrumental responding. However, the dopamine released when a response has stimulus concomitants – is marked – would strengthen instrumental conditioning. All forms of conditioning are enhanced in an aroused organism (Killeen 1975), perhaps as a result of sensitized response-dependent release of dopamine.

Popular models of self-control are also exemplified with backward gradients, and they support inferences of relevance to ADHD. Most organisms choose a larger or better reinforcer over a smaller or inferior reinforcer. When the larger reinforcer is sufficiently delayed, preference switches to the smaller, more immediate reinforcer. This might be construed as a rational choice by organisms that attribute higher value to the soon-small outcome; but, in the modern parlance, it is called a failure of self-control. For such a reversal of preference to the more immediate reward, gradients must not be parallel, thus ruling out the ideal (constant discount) exponential decay form of the gradient. But what controls the choice behavior? Certainly, neither the backward gradients nor precognition, which have similar ontological status, will do the job. Control by delayed reinforcers occurs either because the organism has a history of such a delay in the present context, or has been promised a delayed reward and infers its immediate value from personal histories of such delays. In both cases conditioned reinforcers – differential stimuli such as key lights or tones, or self-instructions to “keep the eyes on the prize” – may mediate the choice of the delay. Indeed, Williams’ (1999) data suggest that direct conditioning of choice response traces will be blocked by conditioned reinforcers as those emerge. If the conditioned reinforcer immediately follows the target response, the response will be strengthened; if it does not, conditioning of the response will be blocked.

The strength of the conditioned reinforcers may be calculated by decomposing the conditioning process into brief continual acts of attention to the stimuli (CSs) which fill the gap. Figure 2 shows that the saturation of memory by the CS is proportional to the integral of the delay gradient. But that representation of the CS ex-

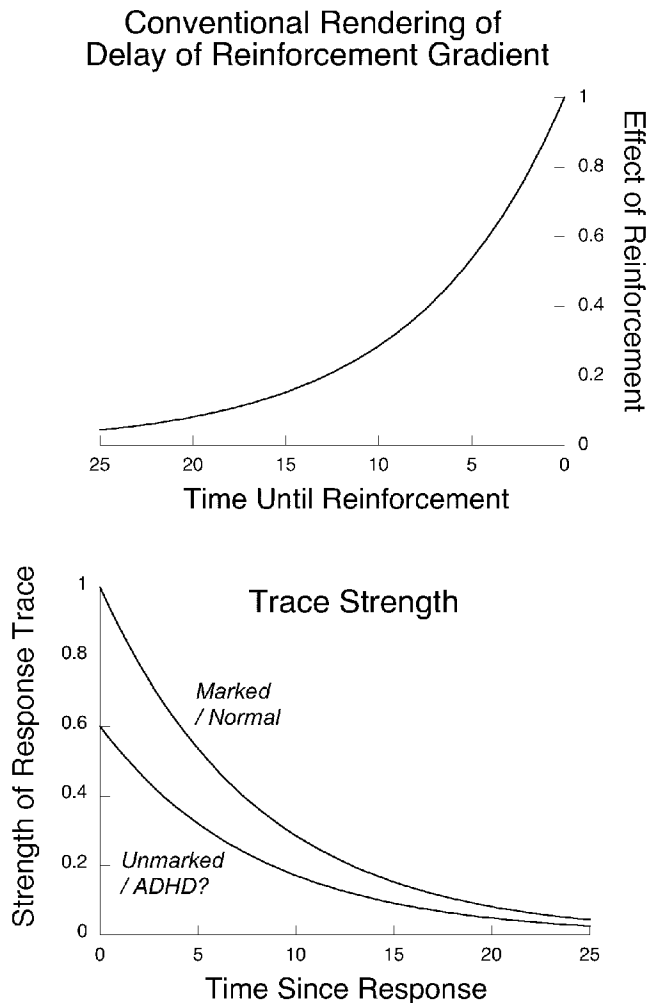


Figure 1 (Killeen). Decay of reinforcement gradients (top) are more properly called delay of reinforceability gradients (bottom). If memorability of the response is strengthened by marking, or weakened by conditions such as ADHD, the ability of a reinforcer to strengthen behavior is affected accordingly.

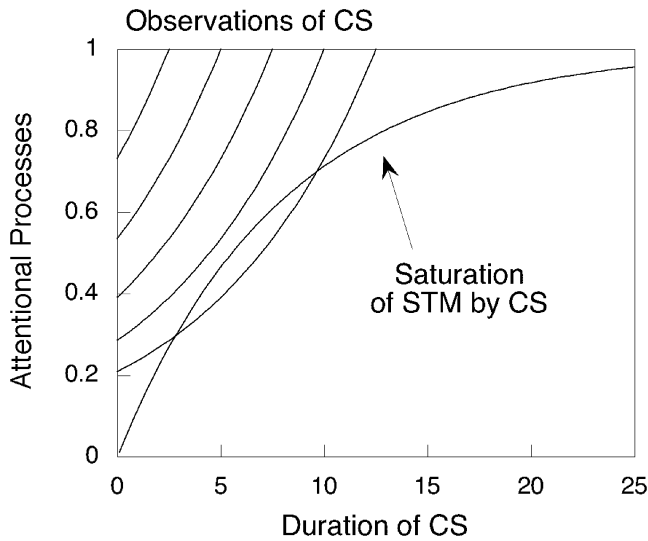


Figure 2 (Killeen). The CS is coupled to primary reinforcement by the decaying traces of memory of its elements at the time of reinforcement, some of which are shown at the left of the figure. The integral of these traces at the time of reinforcement (the *y*-axis) is given by the curve ascending to the right.

tends over a longer and longer interval as the delay to reinforcement, *td*, increases. The density of reinforcement in the presence of the memory of the CS may be calculated by dividing the saturation level by *td* (Killeen 2001a; 2001b). If the trace gradient is exponential with rate of decay of λ , then the strength of the conditioned reinforcer is given by either:

$$S = \frac{\int_0^t e^{-\lambda t} dt}{t_d} = \frac{1 - e^{-\lambda t}}{\lambda t_d} \quad (1)$$

or

$$S = \frac{\int_0^t \lambda e^{-\lambda t} dt}{t_d} = \frac{1 - e^{-\lambda t}}{t_d} \quad (2)$$

These two forms correspond to the two types of (reversed) traces shown in Catania's Figure 4. Equation 1 hinges the gradient at λ when *td* = 0: Variations in the rate of decay of the trace do not affect the strength at zero delay (see Fig. 3, open symbols). Equation 2 hinges it at λ to maintain a constant area under the curve. When these equations are embedded in a more fully articulated model, the presence or absence of the rate constant in the denominators is absorbed by other constants. But in cases where the rate parameter λ is itself under consideration, as in the target article, the differences are important. If individuals with ADHD have steepened gradients, Equation 1 predicts that at long delays conditioned reinforcers will be debased by the larger value of λ ; Equation 2 predicts that steepened gradients would have little differential effect at long delays, but would actually be beneficial at shorter delays due to the quicker saturation of memory (see Fig. 3, filled symbols). Individuals with ADHD have difficulty deferring gratification – difficulty in ordering their behavior with respect to delayed outcomes, despite an apparent general desire to do so – and no obvious advantage at short delays, suggesting that Equation 1 may be the correct form. Figure 4 applies Equation 1 to Catania's data, showing that it is not easily discriminated from the inverse "hyperbolic" gradient often used to fit such data.

Having established Equation 1, it may be developed to address the self-control paradigm – that is, changes in organisms' prefer-

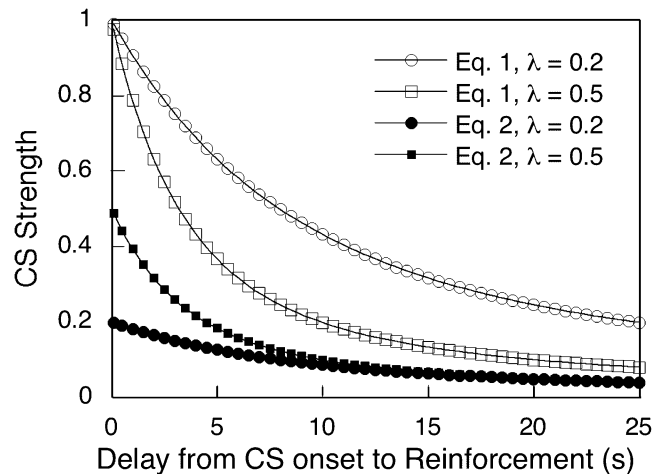


Figure 3 (Killeen). The contrasting predictions made by Equation 1 (open symbols) and Equation 2 (filled symbols) for moderate (circles, $\lambda = 0.2$) and fast (squares, $\lambda = 0.5$) gradients.

ence for the larger delayed reinforcer as the delay to that reinforcer increases. The proportional strength of CSs signalling different delays and amounts of reinforcement may be written as $P = v_1 S_1 / (v_1 S_1 + v_2 S_2)$, where v_i is a constant reflecting the value of the reinforcer, and S_i is the strength as inferred from Equation 1:

$$P = \frac{v_1(1 - e^{-\lambda t_1}) / t_1}{v_1(1 - e^{-\lambda t_1}) / t_1 + v_2(1 - e^{-\lambda t_2}) / t_2} \quad (3)$$

Because the rate constant cancels out of the denominators, the same prediction also follows from Equation 2. In the case where the delay to the small reinforcer is constant, the right addend in the denominator may be assigned a constant value, such as 1.0, giving:

$$P = \frac{v_1(1 - e^{-\lambda t_1})}{v_1(1 - e^{-\lambda t_1}) + t_1} \quad (4)$$

Equation 4, and the more general Equation 3, provides a map to the data of self-control experiments, parsing the effects into incentive value, or valence (v_i) and rate of gradient decay (λ). Equations

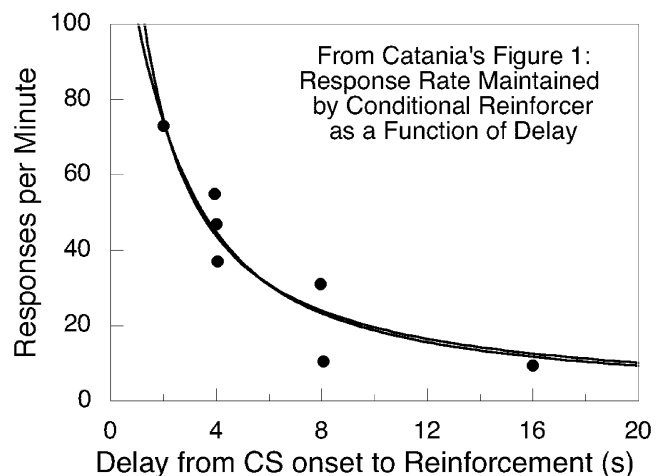


Figure 4 (Killeen). The decreasing efficacy of a conditioned reinforcer as a function of the delay it signals. One curve is proportional to Equation 1 ($\lambda = 0.79s - \lambda$), the other to an inverse function of delay ($\lambda + 1.23t - \lambda$).

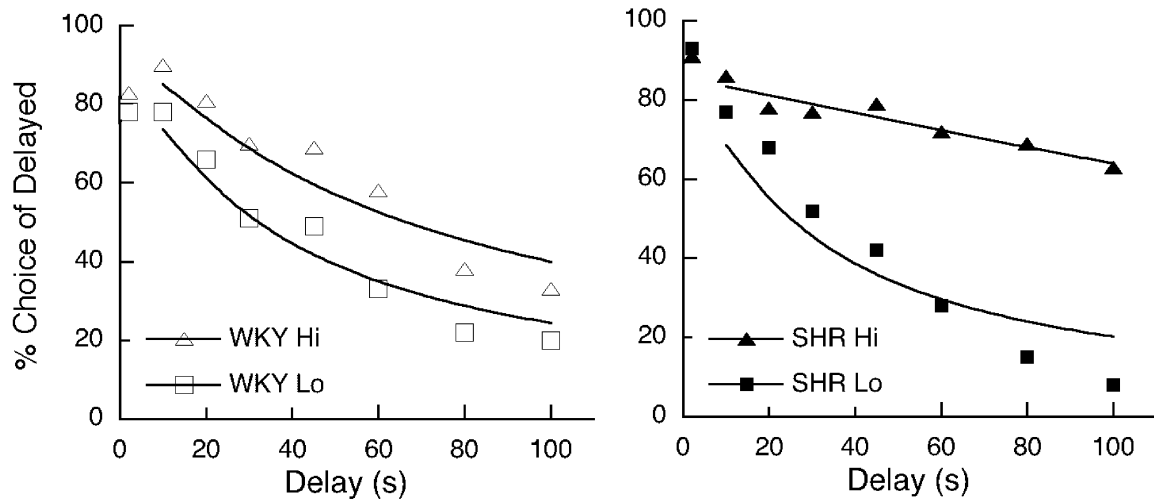


Figure 5 (Killeen). Preference for a large (five pellets) over a small (one pellet) reinforcer as a function of the delay to the larger. Median splits on preferences yielded different characteristics for the two strains. These are parsed by Equation 4 as differences in valence of the large reward for the two WKY groups, with both groups having the same rate of decay ($\lambda = 0.2$). For the SHR strains both valences and gradients ($\lambda = 0.04, 0.2$) differed. (Adriani et al. 2003)

tion 4 is applied to the interesting data of Adriani et al. (2003), shown in Figure 5. These authors found large intersubject variability in the performance of SHR (spontaneously hyperactive) rats given the choice between a small immediate reinforcer and a large delayed one. They therefore did a median split on the overall preference to yield the graph shown in the right panel. A similar median split on the control animals yielded very different profiles. Equation 4 drew the curves through the data, yielding estimates of the two key parameters. For the WKY (Wistar Kyoto Rat), all of the effect was due to variation in the subjective value of the reinforcers, the Hi group preferring the large reward twice as much as the Lo group, with λ remaining constant at $0.2 s - \lambda$. The SHR Lo group had about the same v_i and λ as found in the WKY Lo. The SHR Hi group preferred the large reward six times as much as the Lo group and had a much flatter delay gradient ($\lambda = 0.04$). These data give no support for steeper gradients for the SHR strain, nor for failure of impulse control, but rather, underscore the high variability of operating characteristics in these populations, and the need for care when drawing inferences from pooled data.

Conclusion. The *Gradus ad Parnassum* – Steps to Parnassus – was a guide to the elements of Greek and Latin for those who would write proper prose. The aforementioned considerations concerning delay gradients are also elements that only find their meaning in a larger theoretical text, such as that provided by the target article and precommentary, and by Figure 5. The elementary issue I discussed in this commentary is whether the steps lead up to a reinforcer, or down from a response. A case was made for the latter.

ACKNOWLEDGMENTS

The analysis and commentary was made possible by the generous support of the Norwegian Center for Advanced Study and by NSF IBN 0236821 and NIMH 1R01MH066860. I thank Jonathan Williams for comments on a draft.

ADHD, comorbidity, synaptic gates and re-entrant circuits

Florence Levy

School of Psychiatry, University of New South Wales, Prince of Wales Hospital, Sydney, New South Wales, 2031, Australia. f.levy@unsw.edu.au

Abstract: The “dynamic developmental” theory of attention-deficit/hyperactivity disorder (ADHD) has come full circle from Wender’s (1971) reinforcement hypothesis. By specifying the principle of time constraints on reinforcement and extinction, the present theory allows for empirical validation. However, the theory implies, but does not discuss, implications for the neurophysiology of comorbidity in ADHD. The authors’ attribution of comorbid oppositional behavior to parental and societal reinforcement leaves out biological factors.

Sagvolden et al. are to be congratulated on their comprehensive “dynamic developmental” theory of attention-deficit/hyperactivity disorder (ADHD) (predominantly hyperactive/impulsive and combined subtypes), which integrates behavioural analysis with neurobiological factors. The authors describe the consequences of a hypo-functioning mesolimbic dopamine circuit as altered reinforcement of behavior and deficient extinction of previously reinforced behavior. In some ways the theory has come full circle from Wender’s (1971) theory of minimal brain dysfunction, which postulated a reinforcement deficit. However, the strength of the dynamic developmental theory is that it is based on empirical animal studies in addition to clinical observations. Sagvolden et al. quote the three-factor Hebbian learning rule (Hebb 1949) that synaptic transmission is facilitated when presynaptic input, postsynaptic activation, and the dopamine signal occur simultaneously at the same neuron. Thus, the effect of a reinforcer is more potent when the delay between the response and the reinforcer is short rather than when the delay is long (delay of reinforcement gradient). In ADHD children, a steeper “delay of reinforcement gradient” allows a shorter time window for effective reinforcement contingencies, accounting for the necessity for immediate reinforcement to control the impulsive and hyperactive behavior of ADHD children.

Although the elaborated theory is comprehensive and heuristic in terms of parental and societal influences, and allows for empirical validation, it leaves out important aspects of the behavior of ADHD children. The authors attribute the frequent co-occurrence of oppositional and conduct disorders in ADHD children to