

The Surprising Influence of Delayed Primary Reinforcement on Choice

Margaret A. McDevitt

Abstract

It is well known that the duration of the delay between a response and consequence is inversely related to the impact of that consequence on future responding, and even short delays can greatly undermine the effectiveness of a consequence. However, several studies have shown that delayed primary reinforcement can have a substantial impact on responding in situations in which it was assumed to exert little or no influence. For example, delayed primary reinforcement has produced surprisingly strong effects on responding in procedures with simple concurrent schedules and concurrent chains schedules. This article will highlight two studies (McDevitt & Williams 2001; Ploog, 2001) that demonstrate that delayed primary reinforcement can have direct effects on choice.

Keywords: Delay of reinforcement, unsignaled delay, choice, key peck, pigeons

As behavior analysts, we are all familiar with the fundamental concept that delaying the delivery of reinforcement comprises its effectiveness. For a primary reinforcer to be directly effective (i.e., without resorting to the mediating influence of conditioned reinforcement), it should be presented immediately following the behavior. It is well established that increasing the delay to primary reinforcers systematically decreases the effectiveness of those reinforcers. We do well to provide immediate consequences for behavior when we want those consequences to influence future behavior, and just as importantly, we should increase the delay to consequences when we want to minimize the influence on future behavior. I sometimes use this latter approach when I find that I have inadvertently reinforced a behavior. For example, several years ago I was embarrassed to realize that I had unintentionally reinforced my cat's behavior of crying (painfully loudly) by feeding her shortly after the behavior. By simply increasing the delay from the behavior (crying) to the primary reinforcer (feeding), I easily eliminated the behavior.¹ My cat now sits quietly by her food dish staring at me when she wants to be fed.

Some notable research studies have highlighted the importance of the temporal relationship between responses and consequences. For example, Williams (1976, Experiment 1) interposed a delay between responding and reinforcement after training pigeons to respond to a variable-interval (VI) 2-min schedule of reinforcement. Across conditions, the delay period was varied from 3 s to 15 s, and was unsignaled (the peck that met the VI requirement started the delay timer but was not associated with any stimulus change, and food was delivered at the end of the delay period). Even with the shortest (3 s) delay period, response rates were reduced 70-80% compared to baseline responding. Thus, a small delay had a profound effect on responding. The results are even more striking when one considers that the actual delay between the last response and the reinforcer was likely shorter than 3 s since the pigeons could continue to respond during the delay period.

Similarly, other researchers have found considerable attenuation of response rates when short delays are interposed between responses and reinforcers (e.g., Black, Belluzzi, & Stein, 1985; Royalty, Williams, & Fantino, 1987; Sizemore and Lattal, 1977). The observation that short delays can have large

¹ In this case, I used a differential-reinforcement-of-other behavior contingency (DRO, also known as a resetting delay), such that each subsequent vocalization from my cat restarted the delay period. One characteristic of using the DRO contingency is that it ensures that the obtained delay is equivalent to the programmed delay (e.g., Critchfield & Lattal, 1993; Lattal & Gleeson, 1990; Lattal & Metzger, 1994).

negative effects is also evident in the widespread practice of using very short changeover delays (CODs) when assessing preference in a choice procedure, often as short as 1.5 s (e.g., Herrnstein, 1961). In other words, simply delaying the reinforcer by less than 2 s appears sufficient to disrupt adventitious reinforcement of switching behavior.

Although the delay to primary reinforcement clearly *is* important, can we assume that a delayed reinforcer is always ineffective? It has long been presumed that primary reinforcement delivered after a long delay does not directly affect behavior, but acts indirectly due to intervening conditioned reinforcers (e.g., Skinner, 1953, p. 76; Spence, 1947). I will argue that we should not assume, simply because a reinforcer is delayed, that it will have no impact on future behavior. Research is accumulating to suggest that delayed primary reinforcement may in fact have a much more powerful effect than previously assumed. To demonstrate this, I will highlight two studies that illustrate the influence of delayed primary reinforcement, separate from the effects of conditioned reinforcement.

Delayed Primary Reinforcement Effects in Concurrent Schedules

Ben Williams and I conducted a study (McDevitt & Williams, 2001) in which we presented pigeons with a choice between two delay-of-reinforcement alternatives. Because most research on delayed reinforcement had been conducted in single-response situations (e.g., Catania & Keller, 1981; Sizemore & Lattal, 1977; Williams, 1976), we were interested in how delayed reinforcement might affect choice responding. In our procedure, pigeons chose between two delay-of-reinforcement alternatives. In Experiment 1, the delay periods were 5 s and 15 s. The first peck to satisfy the choice schedule began a delay timer, and food was delivered at the end of the delay period. In the unsignaled condition, there were no stimulus changes during the delay period. The choice stimuli remained illuminated and further pecks were recorded but had no scheduled consequence. Thus, there was nothing to indicate to the pigeons that any given choice peck was effective in starting the delay timer.

One might expect that responding for delayed reinforcement in a choice situation might be even more adversely affected than in a single-response situation. First, as in the single-response situation, subjects might not discriminate the response-reinforcer contingency. Second, the fact that there is not just one, but two, delay periods operating complicates the procedure. Third, the choice situation also presents the opportunity for a response from one alternative to be associated with a reinforcer from the other alternative. For example, in the condition with no delay signals, subjects could continue responding to the choice stimuli during the delay, and could even switch back and forth between alternatives. In fact, subjects did both.

Surprisingly, not only were the subjects clearly able to discriminate the response-reinforcer contingencies despite the delays to reinforcement, they did so to a high degree in all conditions. The results from Experiment 1 are shown below in Figure 1, which compares the condition with no delay signals with conditions in which the delay interval was signaled. In the non-differential signals condition, a center horizontal line was illuminated during all delay intervals. In the differential signals condition, the center keylight was illuminated with the same color as the selected choice stimulus (so that the 5-s delay was signaled by a different color than the 15-s delay). Surprisingly, the condition in which the delay interval was completely unsignaled was not significantly different from the condition in which the delay interval was differentially signaled. Despite the complete lack of feedback in the unsignaled condition regarding which choice pecks were effective and which weren't, subjects showed an extraordinary degree of preference for the alternative providing food after a 5-s delay. With FR1 initial links, the mean choice proportion for the 5-s delay was .97. Even the condition with the lowest level of preference, with concurrent VI 60 VI 60 initial links, still produced a very high degree of discrimination, with a mean choice proportion of .77.

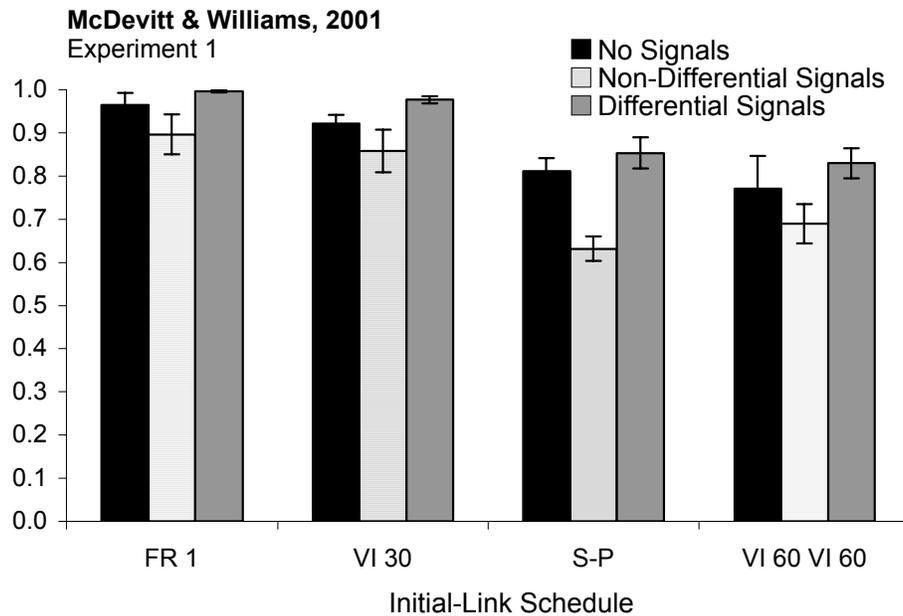


Figure 1. Preference for the 5-s delay alternative over the 15-s delay alternative in Experiment 1 of McDevitt and Williams (2001). The presence and type of signals presented during the delay period, and the initial-link schedule was varied across conditions.

Needless to say, the extreme levels of preference that developed in Experiment 1 with unsignaled delayed reinforcement surprised us. In Experiment 2 we replicated the unsignaled condition from Experiment 1 and included an additional unsignaled condition in which we increased the absolute delays four-fold to 20 s and 60 s. The results of Experiment 2 (shown in Figure 2) confirmed the surprising results of Experiment 1. The mean choice proportion was .87 for the 5-s delay (over the 15-s delay) and .88 for the 20-s delay (over the 60-s delay). Thus, not only was preference more extreme than the .75 choice proportion predicted by matching (Herrnstein, 1970), but increasing the absolute delays four-fold had absolutely no impact on the degree of preference. We expected that as the absolute delay increased, behavior would weaken as it does in the single-response situation, and preference would decrease towards indifference. Instead, there appears to be no effect of dramatically increasing the absolute delays. Of course, further research is needed to determine how far the delay periods can be extended before behavior breaks down. Of interest will be whether or not preference changes as the delays are further increased, or if the choice proportion remains the same, until responding simply ceases.

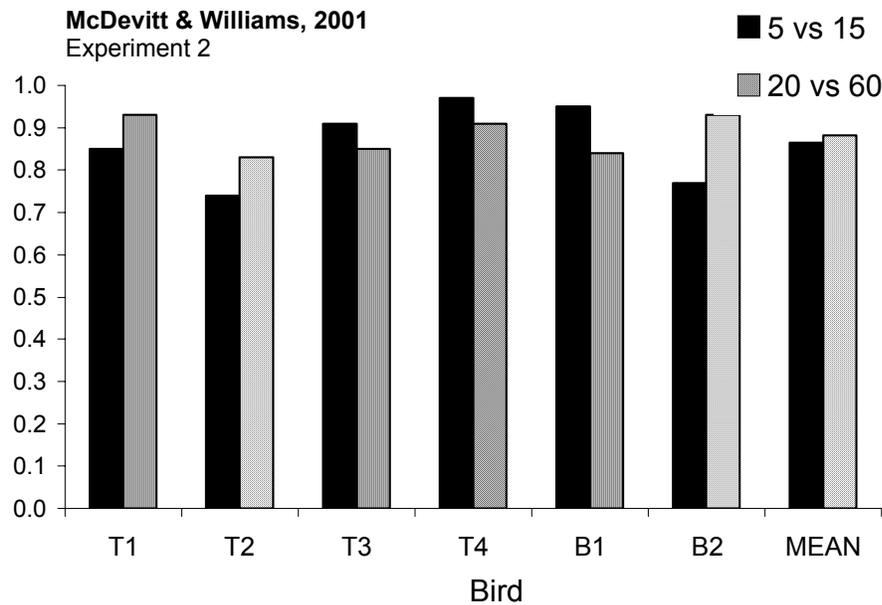


Figure 2. Preference for the shorter delay for each subject in Experiment 2 of McDevitt and Williams (2001). All birds chose between two VI 60-s schedules. Reinforcement was delayed for both alternatives, and the choice stimuli remained illuminated during the delay period. The reinforcers were delayed 5 s and 15 s in one condition, and 20 s and 60 s in the other condition.

Delayed Primary Reinforcement Effects in Concurrent Chain Schedules

Concurrent chains have long been used to study the effects of conditioned reinforcement, and in fact, it has been assumed that initial-link responding is governed solely by conditioned reinforcement. In some theoretical accounts, choice has been assumed to be based entirely on conditioned reinforcement, and the value of delayed primary reinforcement has been disregarded (e.g., Fantino, 1977; Mazur, 1997). Ploog (2001), however, conducted an ingenious experiment to assess the impact of delayed primary reinforcement, separate from conditioned reinforcement, in concurrent chain schedules. He presented pigeons with a choice between two equal initial-link schedules. The initial-link stimuli were signaled by different keylights, but each could appear on either the left or the right response keys. The terminal-link stimuli were always the same color and presented on the center key, regardless of which alternative was selected in the initial link. The two alternatives also employed the same terminal-link schedules. What differed between them was the amount of primary reinforcement. The point of this procedure was to isolate the differences in delayed primary reinforcement as the only potential influence on choice responding.

Ploog's (2001) main result is that he found preference for the alternative leading to the greater reinforcer amount for 57 of 60 cases. Figure 3 shows the most striking results from condition 5 of Ploog's study, in which birds chose between alternatives providing 1-s and 6-s hopper durations. Initial links were variable-interval (VI) 20-s schedules and terminal links were either 10-s (left bars, mean = .84) or 5-s (right bars, mean = .96) VI schedules. These results show that delayed primary reinforcement can produce reliable preference in a concurrent chains procedure.

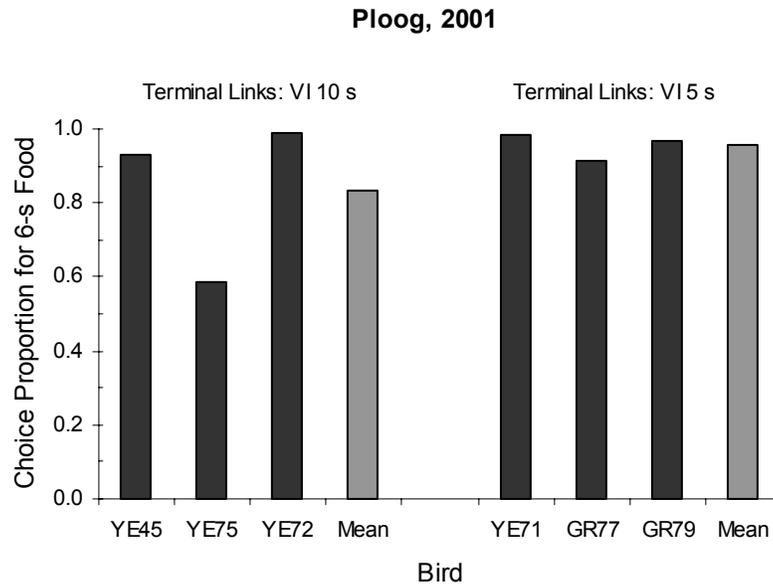


Figure 3. Preference for 6-s food over 1-s food in Ploog's (2001) study. The data are shown for subjects with VI 20-s initial-link schedules from condition 5.

For our purposes here, it is important to emphasize the degree to which Ploog's (2001) procedure can successfully eliminate explanations in terms of conditioned reinforcement. The same color and response key location was used for the terminal links of both chains in order to neutralize any differential effect of conditioned reinforcement on choice responding. Because it is possible that differential stimuli might inadvertently be generated if a given choice response was associated with a particular location (e.g., if a left choice peck was always associated with the larger reinforcer, perhaps subjects would generate their own differential terminal-link stimuli by standing on the left side of the chamber during the terminal-link delay), Ploog's procedure equally often assigned each choice stimulus to the left and the right response keys. Thus, the variable initial-link stimulus locations eliminated side position as a potential bridge to the delayed primary reinforcement at the end of the terminal links. However, one might argue that the same terminal-link could function differently depending on the choice peck that preceded it (e.g., if the center yellow stimulus preceded by a peck to the red stimulus was discriminated from the center yellow stimulus preceded by a peck to the green stimulus). If such a discrimination were to occur, one would expect to see different rates of responding in the terminal link depending on which alternative had been chosen. There was no evidence that such a discrimination took place, as responding during the terminal links did not differ for the two chains. In fact, it is probable that the influence of the delayed primary reinforcement was attenuated by the presence of the same terminal-link stimulus for the two chains, as the terminal-link stimulus should function as a conditioned reinforcer for initial-link responding. The addition of equal conditioned reinforcement to the two chains (comparable to the non-differential signals condition in McDevitt & Williams, 2001) should drive down preference, to some degree masking the impact of the differential primary reinforcement at the end of the chains.

Ploog's (2001) provocative study reveals that in a concurrent-chains schedule, primary reinforcement can affect initial-link responding directly and independently of conditioned reinforcement. This finding is both surprising and problematic since it is usually assumed, as noted earlier, that responding in the initial link of concurrent-chains schedules is a pure measure of the conditioned reinforcement present in the terminal links.

Conclusions

The results of McDevitt & Williams (2001) and Ploog (2001) show that delayed primary reinforcement can have important effects on behavior. What both studies have in common is that they assess the direct effects of delayed primary reinforcement in choice procedures and eliminate explanations in terms of conditioned reinforcement. In McDevitt & Williams' study, an analysis in terms of conditioned reinforcement is precluded by the absence of any stimulus change to indicate transition to a delay period. In Ploog's study, initial-link stimulus location was varied and the same terminal link stimulus was used for both chains, eliminating differential conditioned reinforcement as an explanation.

Why is it that these studies clearly show the effects of delayed primary reinforcement, when other studies have shown remarkable deterioration of responding with reinforcer delays of just a few seconds? First, the two studies highlighted here both used relative rate of responding (preference) as a dependent variable, not absolute rate of responding. It is possible delayed reinforcement may affect these measures differently. The absolute rate of responding might be more sensitive to reinforcement delays than relative rate of responding. Second, there are some studies in single-response situations that have shown surprisingly robust effects with delayed primary reinforcement. For example, Lattal and Gleeson (1990) established and maintained responding in rats and pigeons with un signaled delay intervals up to 30 s. It is likely that procedural considerations modulate the degree to which delayed primary reinforcement effects are evident.

Overall, the results of these studies call into question our current understanding of how and when delayed primary reinforcement will affect responding. The finding that high, reliable preference can be established in situations in which conditioned reinforcement has been eliminated or neutralized challenges the prevailing assumption that conditioned reinforcement is the primary mechanism responsible for choice.

References

- Black, J., Belluzzi, J. D., & Stein, L. (1985). Reinforcement delay of one second severely impairs acquisition of brain self-stimulation. *Brain Research*, 359, 113-119.
- Catania, A. C., & Keller, K. J. (1981). Contingency, contiguity, correlation, and the concept of causation. In P. Harzem & M. D. Zeiler (Eds.), *Advances in analysis of behaviour: Vol. 2. Predictability, correlation, and contiguity* (pp. 125-167). Chichester, England: Wiley.
- Critchfield, T. S., & Lattal, K. A. (1993). Acquisition of a spatially defined operant with delayed reinforcement. *Journal of the Experimental Analysis of Behavior*, 59, 373-387.
- Fantino, E. (1977). Conditioned reinforcement: Choice and information. In W. K. Honig & J.E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 313-339). Englewood Cliffs, NJ: Prentice Hall.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4, 267-272.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, 13, 243-266.
- Ploog, B. O. (2001). Effects of primary reinforcement on pigeons' initial-link responding under a concurrent-chains schedule with nondifferential terminal links. *Journal of the Experimental Analysis of Behavior*, 76, 75-94.

- Lattal, K. A., & Gleeson, S. (1990). Response acquisition with delayed reinforcement. *Journal of the Experimental Psychology: Animal Behavior Processes*, 16, 27-39.
- Lattal, K. A., & Metzger, B. (1994). Response acquisition by Siamese fighting fish (*Beta Splendens*) with delayed visual reinforcement. *Journal of the Experimental Analysis of Behavior*, 61, 35-44.
- Mazur, J. E. (1997). Choice, delay, probability, and conditioned reinforcement. *Animal Learning & Behavior*, 25, 131-147.
- McDevitt, M. A., & Williams, B. A. (2001). Effects of signaled versus unsignalled delay of reinforcement on choice. *Journal of the Experimental Analysis of Behavior*, 75, 165-182.
- Royalty, P., Williams, B. A., & Fantino, E. (1987). Effects of delayed conditioned reinforcement in chain schedules. *Journal of the Experimental Analysis of Behavior*, 47, 41-56.
- Sizemore, O. J., & Lattal, K. A. (1977). Dependency, temporal contiguity, and response-independent reinforcement. *Journal of the Experimental Analysis of Behavior*, 25, 119-125.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Free Press.
- Spence, K. W. (1947). The role of secondary reinforcement in delayed reward learning. *Psychological Review*, 54, 1-8.
- Williams, B. A. (1976). The effects of unsignalled delayed reinforcement. *Journal of the Experimental Analysis of Behavior*, 26, 441-449.

Author Contact Information:

Margaret A. McDevitt
Department of Psychology
McDaniel College
2 College Hill
Westminster, MD 21157
Phone: 410-857-2523
Email: mmcdevit@mcdaniel.edu