THE ROLE OF THE RESPONSE–REINFORCER RELATION IN DELAY-OF-REINFORCEMENT EFFECTS

APRIL MICHELE WILLIAMS AND KENNON A. LATTAL

WEST VIRGINIA UNIVERSITY

The role of the response–reinforcer relation in maintaining operant behavior under conditions of delayed reinforcement was investigated by using a two-operandum (i.e., two-key) procedure with pigeons. Responding on one key was reinforced under a tandem variable-interval differential-reinforcement-of-other-behavior (tandem VI DRO) schedule. The schedule defined a resetting unsignaled delay-of-reinforcement procedure in that a response was required when the interfood interval of the VI schedule lapsed, but further responding during the DRO component on either key reset the time interval. This ensured a fixed delay duration between any response and reinforcement. Responding on another key, physically identical to the first one except for spatial location, otherwise was without consequence. The location of the key correlated with the delay-of-reinforcement procedure varied between sessions according to a semirandom sequence. Differences in response rates between the two keys were greater, with proportionally higher rates on the key correlated with the delay-of-reinforcement procedure, the longer the delay-of-reinforcement procedure remained correlated with the same key. Differences in responding on the two keys also increased within individual sessions. These results suggest that the response–reinforcer relation is the primary determinant of responding when responding is acquired and maintained with delayed reinforcement.

Key words: delay of reinforcement, tandem VI DRO schedule, response–reinforcer relation, key peck, pigeons

---

The presentation of a stimulus as a reinforcer may control responding through its direct, response-strengthening, effects and also indirectly through other stimulus functions (cf. Zeiler, 1977). For example, the delivery of a reinforcer can serve as a discriminative stimulus to evoke a particular type or form of responding as a function of its prior correlation with different conditions of reinforcement (Franks & Lattal, 1976). Putative reinforcers such as food also can evoke responses in the absence of any history of reinforcement of those responses (Schoenfeld, Antonitis, & Bersh, 1949; Segal, 1959). Responses evoked in response to discriminative stimuli or as a result of establishing operations such as food deprivation often occur concurrently with operant responses and usually are indistinguishable from these operant responses. Such additional effects of reinforcer delivery

are rarely considered in most investigations in which reinforcement is immediate, because the number of responses that occur as a result of those effects is likely to be small relative to the number that occur as a direct result of the reinforcement contingency.

When response rates are persistent but low, it becomes more important to assess the relative contributions of direct and indirect effects of reinforcer delivery to such responding. One such instance occurs when responding is established and thereafter maintained when an unsignaled delay is imposed between the response and reinforcer (Lattal & Gleeson, 1990). Particularly when each response during the delay restarts the delay interval (a resetting delay), response rates may be on the order of only one or two per minute. These low response rates invite the question of whether it is the response–reinforcer dependency or some other variable that is responsible for the maintained responding.

The role of the response–reinforcer dependency in response maintenance under delay-of-reinforcement conditions has been assessed in different ways. Lattal and Gleeson (1990) compared responding by rats and pigeons under an unsignaled resetting delay-of-reinforcement procedure to responding by

other animals under a schedule on which response-independent food presentations occurred at the same rate and with the same distribution as those in the delay-of-reinforcement procedure. The response-independent food presentations maintained minimal responding relative to that sustained by the delay-of-reinforcement procedure. Wilkenfield, Nickel, Blakely, and Poling (1992) and Critchfield and Lattal (1993) added a second operandum on which responding of rats had no consequence (hereafter labeled the irrelevant operandum) to an unsignaled resetting delay-of-reinforcement procedure. In the Wilkenfield et al. experiment the operandum correlated with the delay-of-reinforcement procedure (hereafter labeled the relevant operandum) and the irrelevant operandum were identical levers located on either side of a work panel. Critchfield and Lattal employed a lever as the irrelevant operandum and the breaking of a photocell at the rear of the chamber as the operant response that initiated an unsignaled delay interval that terminated with food delivery. They found in two experiments that responding on an irrelevant operandum persisted at a low rate while responding on the relevant operandum systematically increased across the 20 sessions of the experiment. Only 1 animal responded consistently on the irrelevant operandum. With 16- and 32-s delays, Wilkenfield et al. observed substantial responding on the irrelevant operandum. Average response rates over the single 8-hr session were higher on the irrelevant operandum for 16 of 18 rats.

One interpretation of Wilkenfield et al.'s (1992) results is that irrelevant lever responding was evoked by the food presentations as noted above. This interpretation suggests that responding on the relevant lever may have been similarly evoked instead of occurring as a result of the response–reinforcer dependency. This interpretation is in conflict with the earlier conclusion of Lattal and Gleeson (1990) that the sustained responding observed in their experiment resulted primarily from the response–reinforcer relation.

The issues raised by Wilkenfield et al.'s (1992) data about the role of the response–reinforcer dependency in such delay-of-reinforcement effects, and the discrepancies between their findings and those of Critchfield and Lattal (1993) who used identical delay-of-reinforcement procedures and almost identical delay values (30 vs. 32 s), gave rise to the present experiment. Another way of assessing the control of responding by the response–reinforcer relation is to periodically reverse the location of the relevant and irrelevant operanda. If responding tracks the response–reinforcer relation over time, that is, if responding is more frequent on the relevant operandum over time and less frequent on the irrelevant operandum, it would suggest that such responding is controlled directly by the response–reinforcer relation and not by an indirect variable such as reinforcer evocation of responding.

METHOD

Subjects

Each of 4 adult male White Carneau pigeons was maintained at 70% ($\pm$10 g) of its ad libitum body weight (cf. Lattal & Williams, 1997). These 70% weights ranged between 390 g and 462 g for the 4 animals. Each had a history of responding on different reinforcement schedules.

Apparatus

An operant conditioning chamber with a work area measuring 33.5 cm by 30 cm by 31.5 cm was used. Three walls were made of wood. The fourth wall was an aluminum work panel that contained two response keys and a feeder aperture. Each response key required approximately 0.15 N to operate. The two keys were located 24 cm from the chamber floor and 5 cm from the left and right walls of the chamber. The keys could be transilluminated by white, green, or yellow bulbs located behind them. The keys were transilluminated at all times except during reinforcement. Reinforcers consisted of 4-s access to mixed grain from a food hopper accessible through an aperture (4.5 cm by 6 cm) located 0.5 cm from the chamber floor and centered on the midline of the work panel. The aperture was illuminated by a white light when grain was available. A ventilating fan operated continuously and also provided masking noise in the chamber. A Tandy 1000ex computer using Med-PC® interfacing and software controlled the experiment and recorded data.

Procedure

Even though each subject had a prior history of reinforced key pecking, magazine training and shaping still were implemented to help ensure consistent responding and thereby perhaps to reduce intersubject variability. During magazine training, access to the reinforcer for 4 s was provided approximately every 15 s until each subject ate promptly after the hopper was raised. This required an average of three sessions. In the session following reliable eating by each subject, key pecking was shaped by differentially reinforcing successive approximations. In each shaping session, a single key was illuminated yellow, and that key alternated daily from left to right. Each shaping session lasted for approximately 150 food presentations.

After key pecking occurred reliably for each subject, the key that was illuminated continued to alternate daily and was correlated with a tandem fixed-ratio (FR) 1 differential-reinforcement-of-other-behavior (DRO) 10-s reinforcement schedule. Thus, each response on the lighted key initiated the DRO schedule, which defined an unsignaled 10-s period that terminated with food delivery. Responses occurring during the 10-s period restarted the delay interval. This condition was in effect for approximately 12 sessions, after which the schedule was changed to a tandem variable-interval (VI) 15-s DRO 10-s schedule for six final training sessions. The VI schedule was constructed using the constant probability distribution described by Fleshler and Hoffman (1962).

In subsequent sessions, both keys were transilluminated yellow, and a tandem VI 15-s DRO 10-s schedule of reinforcement was correlated with only one of the two keys. This key was designated the relevant operandum. Responding on the other key had no programmed consequence and therefore was the irrelevant operandum. Responses on the relevant operandum initiated the DRO 10-s schedule an average of once every 15 s, and responses on either operandum during the DRO schedule reset the delay interval. Thus, food delivery never occurred closer than 10 s to a response on either operandum, even though initiation of the delay interval resulted only from responses on the relevant operandum. Selection of the left or right key as the relevant operandum occurred before each session. Selection was according to a semirandom sequence that limited assignment of the same key as the relevant operandum to no more than three consecutive sessions. The sequences of relevant-operandum assignment can be determined from the symbol type in Figure 1. Sessions occurred at the same time 7 days a week, and each ended after 90 food presentations. The experiment was terminated after 40 sessions.

RESULTS

The data in Figure 1 show that key pecking of each of the 4 pigeons was established and maintained at low (between <1 and 7 responses per minute) but consistent response rates (despite greater variability for Pigeon 2228) throughout the experiment. These data also show that response rates typically were higher on the relevant operandum as opposed to the irrelevant operandum.

The data in Figure 2 show response rates for each pigeon on both operanda in successive 10-min segments of each session. Each data point is an average calculated over successive segments for all sessions of the experiment. The 90-reinforcer session cut-off criterion resulted in uneven session durations. As a result, only responses occurring within the first t minutes of each session were included in the data analysis in Figure 2, where t was either 70 or 80 min for different pigeons. Responding beyond t was eliminated from the data analysis for this figure, but the trends shown in the data depicted continued when the truncated responses were divided by the time remaining after t to yield a response-rate measure for the "overflow" period (session duration $- t$). Response rate usually was higher on the relevant operandum than on the irrelevant operandum throughout the session. In addition, for Pigeons 2405, 2408, and 2485, response rates on the relevant operandum were lower at the beginning of the session and systematically increased as the session progressed. The data of Pigeon 2228, which had the highest response rates and the greatest separation of rates between the relevant and irrelevant operanda, did not show this effect.

The orderly nature of the data averaged across individual sessions does not account
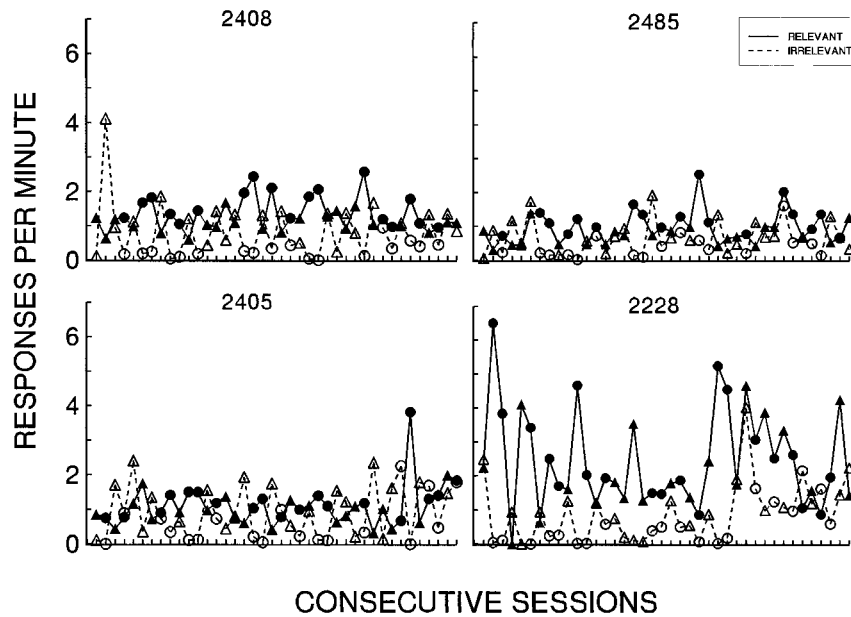
Fig. 1.   Response rates for each subject on the relevant and irrelevant operanda during each session. Triangles represent the first session with a new relevant operandum location. Subsequent identical locations of the relevant and irrelevant operanda are represented as circles. Response rates were calculated by dividing the total number of responses on each operandum by the total session time (in minutes), excluding the time during reinforcer presentations.
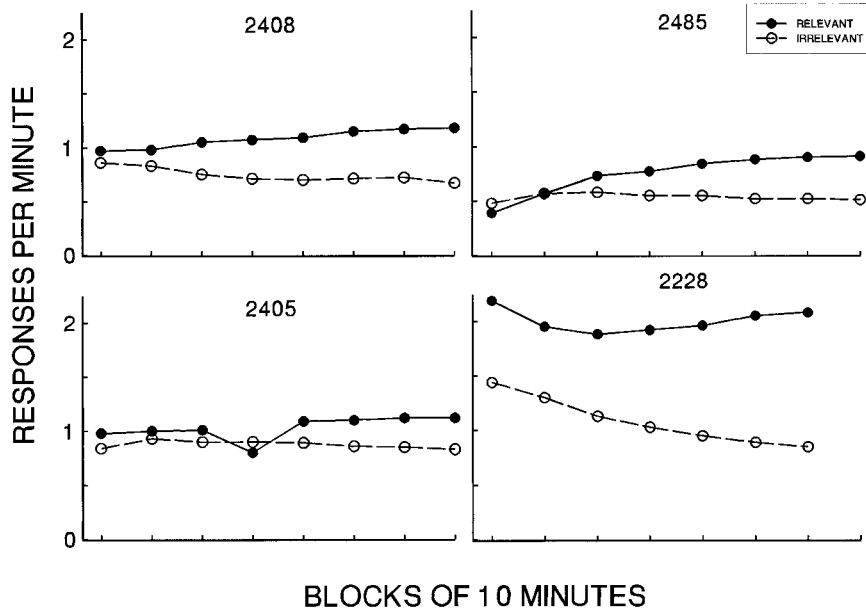

Fig. 2.   Response rates for each subject on the relevant and irrelevant operanda at consecutive 10-min intervals, averaged across all sessions. Response rates were calculated as in Figure 1.
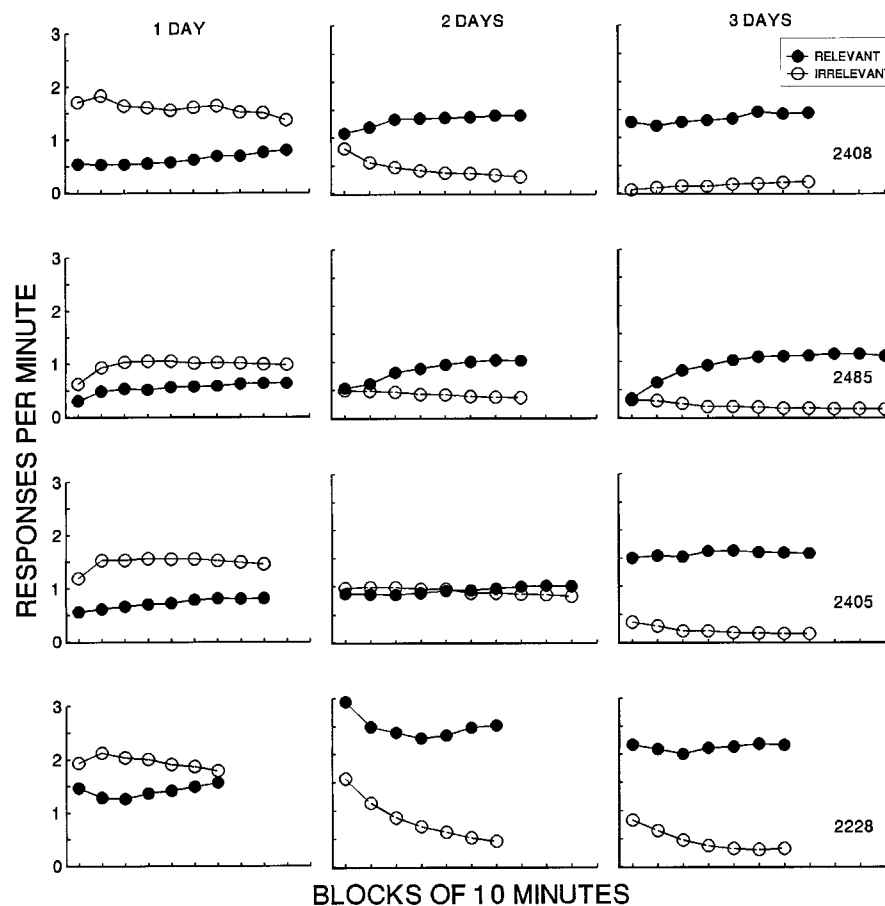
Fig. 3.   Response rates for each subject on the relevant and irrelevant operanda at consecutive 10-min intervals, averaged across sessions with 1 day (left column), 2 days (middle column), or 3 days (right column) at the same relevant-operandum location. Response rates were calculated as in Figure 1.

well for the variability in response rates between sessions shown in Figure 1. Figure 3 therefore shows response rates for each subject during successive 10-min segments for a given session as a function of whether the relevant operandum had been in its present location for only the single day shown (i.e., 1 day), for the day before (i.e., for 2 days), or for the 2 preceding days and the day shown (i.e., for 3 days). As in Figure 2, only responses in the first t minutes of the session were included in the analysis shown in Figure 3. The value of t sometimes differed across the graphs because of the different amounts of time required to accumulate 90 food presentations. Also as in Figure 3, the trends shown in the data depicted continued when the truncated responses were divided by the time

remaining after t to yield a response-rate measure for the overflow period (session duration $- t$). If the relevant operandum had been in the indicated location for a single session (graphs in the left column), response rates were higher on the irrelevant key. When the relevant operandum had been in the same location in the preceding session (graphs in the middle column), response rates were consistently higher on the relevant operandum for all pigeons except Pigeon 2405. For Pigeons 2408 and 2485, response rates on the relevant operandum increased across the session. When there were two previous sessions of the relevant operandum in the same location, response rates were higher on the relevant operandum. For Pigeons 2408 and 2485, response rates on the relevant

Table 1

Average session length (in minutes) for individual subjects in the first, second, or third consecutive sessions of responding with the same relevant operandum.

Subject	1 day (21 sessions)	2 days (13 sessions)	3 days (6 sessions)
2408	127	95	94
2485	192	118	105
2405	147	116	115
2228	126	93	94

operandum increased throughout the session, but for the other 2 pigeons response rates on the relevant operandum were more or less constant throughout the session. For all but Pigeon 2408, response rates on the irrelevant operandum decreased across the session.

The average session durations for each pigeon are shown in Table 1. Each pigeon obtained the maximum number of reinforcers more quickly during sessions that were the second or third with the same relevant operandum location.

DISCUSSION

The location of the dependency between responding and its delayed consequence determined response rates on, and the distribution of responses between, the two operanda. Further evidence of the control of behavior by the response–reinforcer relation comes from the fact that the rates of responding changed both within and across sessions as a function of the location of the relevant operandum. If responding were simply a matter of evocation by the presentation of food, equal responding on both operanda would have been expected.

Despite consistent tracking of the response–reinforcer relation, responding on the irrelevant operandum persisted throughout the experiment. This persistence is not surprising given that the location of the response–reinforcer relation was changing constantly. When that location remained unchanged for several sessions, response rate on the irrelevant operandum tended to decrease over those sessions. If the location had remained unchanged for a larger block of sessions, perhaps response rates would have decreased further. In support of this observation,

Critchfield and Lattal (1993) observed some responding on the irrelevant operandum during early sessions of exposure to the unsignaled delay procedure but, over the course of the 20-session experiment, responding diminished to low levels and remained there.

The present results with a 10-s unsignaled resetting delay were similar to Wilkenfield et al.'s (1992) effects under an 8-s delay in that in both experiments, in relative terms, substantial responding occurred on the irrelevant operandum. Wilkenfield et al. showed that response rates on the irrelevant operandum tended to be higher with even longer delays. Two features of the Wilkenfield et al. procedure suggest possible interpretations of the relatively high rate of irrelevant-operandum responding. First, only a single 8-hr session was conducted with each subject, and there was no report of the distribution of irrelevant responses over the course of the session. As a result, it is not known whether irrelevant-lever responding was stable or in transition over the 8-hr session. Even if such responding was stable over the 8-hr session, it might have decreased, as both the present results and those of Critchfield and Lattal (1993) show, had the number of sessions been extended as in these other experiments. Second, Wilkenfield et al. did not employ a changeover delay between responding on the irrelevant operandum and the delivery of food following a response and delay interval on the relevant operandum, but then neither did Critchfield and Lattal (1993). Wilkenfield et al. suggested that responding on the irrelevant operandum was related to the resetting contingency on the relevant operandum and that these irrelevant-operandum responses may have mediated the required period of nonresponding on the relevant operandum (cf. Laties, Weiss, Clark, & Reynolds, 1965; Schwartz & Williams, 1971). They noted that delays between an irrelevant-lever response and food generally were shorter than delays following relevant-lever responses. If such responding can be followed either immediately by a reinforcer (cf. Catania & Cutts, 1963) or a reinforcer after a shorter delay than that correlated with the relevant operandum, responding could be maintained on the irrelevant operandum by the direct effect of relatively closer temporal contiguity between such irrelevant-operandum responding and

food delivery. In the present experiment, responding on the irrelevant operandum was delayed from food delivery by 10 s, thereby minimizing the likelihood of adventitious correlations between irrelevant-operandum responding and food delivery. Thus, either the short duration of exposure to the delay-of-reinforcement procedure, the shorter delays to reinforcement on the irrelevant operandum, or both, may have contributed to the effects observed by Wilkenfield et al. In conjunction with the present results, these alternative accounts of irrelevant-operandum responding make remote the likelihood that the irrelevant-operandum responding is due to the reinforcement-induction effects described in the introduction.

The potential role of consequences in maintaining irrelevant-operandum responding is further suggested in an experiment reported by Sutphin, Byrne, and Poling (1998). They employed a procedure in which each response of rats on an irrelevant operandum during a delay period initiated by a response on a relevant operandum canceled the forthcoming reinforcer. This procedure reduced the number of irrelevant-lever responses relative to the number emitted by other animals in the absence of such a cancellation contingency. Nonetheless, the number of irrelevant-lever presses remained relatively high, though not as high as responses on the relevant lever, when the delay was 8 s. Similar effects also were obtained with longer delays. As in Wilkenfield et al. (1992), only a single 8-hr session was conducted, and it seems plausible that the number of responses on the irrelevant lever would diminish further with continued training.

The tracking of both response-dependent versus response-independent food delivery and immediate versus no reinforcement has been investigated previously. Lattal (1973) studied the responding maintained by a mixed schedule with response-dependent food presentations in one component arranged according to a VI schedule and response-independent food presentations occurring according to a variable-time (VT) schedule in the second component. Over successive conditions of the experiment, the duration of the mixed-schedule components was varied. In different conditions, components alternated semirandomly every 5, 10,

30, or 60 min. With shorter component durations, response rates in VI and VT schedules were identical, but as component duration was lengthened, higher response rates were controlled by the VI relative to those controlled by the VT schedule. Furthermore, responding over the course of, particularly, 30- and 60-min components in VI and VT tended to increase and decrease, respectively, systematically. Willson and Wilkie (1993) made four operanda, each on a different wall of the chamber and each associated with a different food hopper, concurrently available to pigeons. During any particular session, however, only one of the four was correlated with a VI schedule of positive reinforcement and the other three were correlated with extinction. The operandum that was correlated with reinforcement varied across sessions. At the beginning of a given session, before reinforcement was available in that session, the subjects responded at higher rates on the operandum that had been correlated with reinforcement in the preceding session than on the other three. The next highest rate of responding occurred on the operandum that had been correlated with reinforcement two sessions previously, and the third highest response rates occurred on the operandum that had been correlated with reinforcement three sessions before the current one. Once a response on one of the operanda was reinforced in the session, the response rate on that operandum increased above chance levels. As each session progressed, the response rate on the operandum correlated with reinforcement for that session continued to increase while response rates on each of the remaining operanda progressively decreased over the session.

The present results demonstrate that a response–reinforcer tracking effect, similar to that reported by Lattal (1973) and Willson and Wilkie (1993), is obtained when reinforcers occur after unsignaled resetting delays. The results of all three experiments also suggest that control by a particular response–reinforcer relation increases with longer histories of exposure to that relation. More important, the present results suggest that it is the response–reinforcer relation and not some other behavioral process that is primarily responsible for the sustained, low response rates that have been obtained consistently

when reinforcers follow unsignaled delays that restart with each response during the delay.

REFERENCES

Catania, A. C., & Cutts, D. (1963). Experimental control of superstitious responding in humans. *Journal of the Experimental Analysis of Behavior, 6,* 203–208.

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.

Fleshler, M., & Hoffman, H. S. (1962). A progression for generating VI schedules. *Journal of the Experimental Analysis of Behavior, 5,* 529–530.

Franks, G. J., & Lattal, K. A. (1976). Antecedent reinforcement schedule training and operant response reinstatement in rats. *Animal Learning & Behavior, 4,* 374–378.

Laties, V. G., Weiss, B., Clark, R. L., & Reynolds, M. D. (1965). Overt ''mediating'' behavior during temporally spaced responding. *Journal of the Experimental Analysis of Behavior, 8,* 107–116.

Lattal, K. A. (1973). Response-reinforcer dependence and independence in multiple schedules. *Journal of the Experimental Analysis of Behavior, 20,* 265–271.

Lattal, K. A., & Gleeson, S. (1990). Response acquisition with delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes, 16,* 27–39.

Lattal, K. A., & Williams, A. M. (1997). Body weight and response acquisition with delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 67,* 131–143.

Schoenfeld, W. N., Antonitis, J. J., & Bersh, P. J. (1949). Unconditioned response rate of the white rat in a bar pressing apparatus. *Journal of Comparative and Physiological Psychology, 43,* 41–48.

Schwartz, B., & Williams, D. R. (1971). Discrete-trials spaced responding in the pigeon: The dependence of efficient performance on the availability of a stimulus for collateral pecking. *Journal of the Experimental Analysis of Behavior, 16,* 155–160.

Segal, E. F. (1959). The stability of operant level and its relation to deprivation. *Journal of Comparative and Physiological Psychology, 52,* 713–716.

Sutphin, G., Byrne, T., & Poling, A. (1998). Response acquisition with delayed reinforcement: A comparison of two-lever procedures. *Journal of the Experimental Analysis of Behavior, 69,* 17–28.

Wilkenfield, J., Nickel, M., Blakely, E., & Poling, A. (1992). Acquisition of lever-press responding in rats with delayed reinforcement: A comparison of three procedures. *Journal of the Experimental Analysis of Behavior, 58,* 431–443.

Willson, R. F., & Wilkie, D. M. (1993). Pigeons remember briefly trained spatial location-food associations over extended time periods. *Journal of Experimental Psychology: Animal Behavior Processes, 19,* 373–379.

Zeiler, M. (1977). Schedules of reinforcement: The controlling variables. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 201–232). New York: Prentice Hall.