JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1995, **28,** 389–394 NUMBER 3 (FALL 1995)

# *THE PARADOX OF PREFERENCE FOR UNRELIABLE REINFORCEMENT: THE ROLE OF CONTEXT AND CONDITIONED REINFORCEMENT*

JOSEPH S. LALLI AND BENJAMIN C. MAURO

THE UNIVERSITY OF PENNSYLVANIA

We discuss Belke and Spetch's (1994) work on choice between reliable and unreliable reinforcement. The studies by Belke and Spetch extend a line of basic research demonstrating that under certain experimental conditions in a concurrent chains procedure, pigeons prefer an alternative that produces unreliable reinforcement. The authors describe the variables that influence preference for unreliable reinforcement, including the signaling and the duration of the reinforcement schedules, the context in which the signaling stimuli occur, and the effects of conditioned reinforcement. Hypothetical applied examples that address these variables are provided, and their influence on preference for unreliable reinforcement in humans is discussed. We conclude by suggesting a line of applied research to examine the relationship between these variables and a preference for unreliable reinforcement.

DESCRIPTORS: context, conditioned reinforcement, delay-reduction theory, unreliable reinforcement

Basic and applied researchers have recently focused considerable attention on the nature of reinforcement and its effects on behavior (see the *Journal of the Experimental Analysis of Behavior,* Vol. 60, No. 1, and the *Journal of Applied Behavior Analysis,* Vol. 27, No. 4). In applied research, matching theory, response deprivation, and reinforcer substitutability have been studied more frequently than other theoretical accounts of reinforcement. By contrast, one area that has received little attention from applied researchers is that of conditioned reinforcement.

In a recent paper on the nature and applications of reinforcement, Iwata and Michael (1994) discussed conditioned reinforcement in terms of the delay-reduction theory (Fantino, Preston, & Dunn, 1993). Briefly, the theory states that a stimulus is established as a conditioned reinforcer because its onset signals a reduction in delay to primary reinforcement relative to that signaled by another stimulus in an experimental situation (Fantino et al., 1993). That is, the context in which the stimuli occur helps to establish one of the stimuli as a more effective conditioned reinforcer. Iwata and Michael said that context, as emphasized by the delay-reduction theory, is important to applied researchers because of its contributions to the value of a stimulus as a conditioned reinforcer. Therefore, the present paper is an effort to further the discussion on context and conditioned reinforcement by examining the basic work of Belke and Spetch (1994) on choice between reliable and unreliable reinforcement and to relate these findings to matters of applied significance.

Reliable reinforcement refers to a response that always produces reinforcement (i.e., 100% reinforcement), and unreliable reinforcement refers to a response that produces either reinforcement or no reinforcement with an equal probability (i.e., 50% reinforcement). Belke and Spetch (1994) suggest that the effect of a stimulus as a conditioned reinforcer can, under certain experimental conditions, contribute to an organism's preference for an alternative that produces unreliable reinforcement in a choice situation. Their study contributed to a line of research showing that under certain conditions pigeons prefer unreliable reinforcement over re-

Reprints can be obtained from Joseph S. Lalli, Children's Seashore House, 3405 Civic Center Blvd., Philadelphia, Pennsylvania 19104, or Benjamin C. Mauro, Animal Learning Laboratory, The Children's Hospital of Philadelphia, 34th and Civic Center Blvd. Room 9044, Philadelphia, Pennsylvania 19104.

liable reinforcement (Dunn & Spetch, 1990; Kendall, 1974, 1985; Spetch, Belke, Barnet, Dunn, & Pierce, 1990). Pigeons' preference for unreliable reinforcement is paradoxical because they prefer an alternative that produces a relatively lower rate of primary reinforcement. This research has relevance for applied work on conditioned reinforcement because it may help to identify variables that could produce preference for unreliable reinforcement in humans. Specifically, why is there a preference for problem behavior in a choice situation when a concurrently available appropriate response is continuously reinforced and, therefore, results in more reliable reinforcement?

The standard procedure used to study the effects of percentage reinforcement on choice with pigeons is the concurrent chains schedule. In a concurrent chains procedure, a pigeon chooses between two concurrently available schedules of reinforcement. The time during which the concurrent schedules are available is called the initial link (i.e., a choice phase). The initial links are typically signaled by stimuli (e.g., response keys illuminated white) that are correlated with independent and equal concurrent schedules of reinforcement (e.g., a concurrent fixed-ratio [FR] 1 FR 1 schedule). Thus, a single response on either key during an initial link produces a period of access to one of two independent reinforcement schedules during a terminal link (the outcome phase). The terminal-link reinforcement schedules used in percentage reinforcement studies are usually fixed-time (FT) schedules. The terminal-link schedules can be correlated with distinct stimuli (e.g., red, green, or yellow keys). In studies with pigeons, the specific FT outcomes are usually either access to food or a blackout period (i.e., no food). A concurrent chains schedule differs from a concurrent schedule because completing the initial-link schedule of a concurrent chains schedule produces access to a terminal-link schedule of reinforcement rather than a specific reinforcer.

The pigeon's preference for either reliable or unreliable reinforcement depends on the particular schedule parameters used in the concurrent chains procedure. Two variables that influence preference are the signaling and the duration of the terminal links. Some researchers have studied percentage reinforcement under unsignaled conditions; that is, terminal-link schedules of reinforcement were not correlated with distinct stimuli (Dunn & Spetch, 1990; Kendall, 1974, 1985; Spetch et al., 1990; Spetch & Dunn, 1987). These studies showed that pigeons prefer the alternative that produced a higher total rate of primary reinforcement (i.e., reliable reinforcement). That is, responding in the initial link corresponded to the relative rate of primary reinforcement obtained in the terminal link.

An applied example of a concurrent chains schedule is a situation in which an adolescent male comes down from his bedroom prior to going out (compare to a pigeon in a chamber with the lights off). The adolescent may select one of the two concurrently available parents by going into the kitchen to see his father or into the family room to see his mother (illumination of the two white keys on the operant panel). This is analogous to the initial link of the concurrent chains schedule. Upon the son's appearance in one of the rooms, the parent says, "I thought that you were going out." In a loud and threatening voice or in a polite manner, the adolescent requests use of the family car; this produces access to a terminal-link schedule of reinforcement. For example, in an unsignaled condition the adolescent politely asks his father, who always provides access to the car according to an FT schedule (i.e., 100% reliable reinforcement). The car is provided based on the passage of time (i.e., the FT schedule) rather than any subsequent response by the adolescent. By contrast, the adolescent requests the car from his mother in a threatening manner, which results in access only 50% of the time (i.e., 50% unreliable reinforcement) according to an FT schedule. The terminal-link reinforcement schedules are not correlated with a distinct reaction from either the mother or father. Basic

research (e.g., Dunn & Spetch, 1990; Kendall, 1974, 1985; Spetch et al., 1990; Spetch & Dunn, 1987) suggests that the adolescent will prefer the reliable reinforcement and will be more likely to select father.

Direct support for the effects of reliable unsignaled reinforcement with humans is difficult to find because of a lack of applied research on concurrent chains schedules. However, indirect evidence may be found in applied research with concurrent schedules of reinforcement. For example, Mace, McCurdy, and Quigley (1990) evaluated the effects of unequal concurrent ratio (FR 1 variable-ratio [VR] 2) schedules of reinforcement (candy) for academic performance. The authors found that the students preferred the alternative with a richer schedule of reinforcement. Although these results are consistent with the basic research findings for unsignaled concurrent chains schedules (i.e., preference for the higher rate of primary reinforcement), comparisons between the findings are cautiously made because FT schedules were not used by Mace et al. (1990).

Belke and Spetch (1994) and other researchers (Dunn & Spetch, 1990; Kendall, 1974, 1985; Spetch et al., 1990) have also studied percentage reinforcement under signaled conditions. They used multiple correlated stimuli to signal the terminal-link outcomes. In the typical concurrent chains signaled procedure, a pigeon is presented with two identical stimuli (e.g., two white keys). Responding on the right key produces the terminal-link stimulus (i.e., a red key), which always results in food (100% reliable reinforcement) according to an FT schedule. Responding on the left key produces one of two terminal-link stimuli (i.e., a green or a yellow key), each with an equal probability ($p = .5$) of occurring. Each stimulus is correlated with a specific FT outcome. The green key is correlated with food reinforcement 100% of the time, and the yellow key is correlated with a blackout period. Thus, responding on the left key produces food only 50% of the time. Their findings showed that pigeons prefer a response alternative that produced a lower total rate of primary reinforcement (i.e., unreliable reinforcement). Dunn and Spetch (1990), Kendall (1974, 1985), and Spetch et al. (1990) also reported that preference for the signaled unreliable alternative was strongest when the delay (i.e., FT value) to reinforcement was longest.

Our hypothetical applied example may be modified to include correlated signals for the terminal-link outcomes. For example, dad (the right key in the above example) could say, "Of course. Let me go get the keys." Mom (the left key in the above example) could say "I'll give you the keys after I put my coat in the closet" (the green key). Or she could say, "You cannot have the car tonight because I have to use it" (the yellow key). Perhaps counterintuitively, basic research (Belke & Spetch, 1994; Dunn & Spetch, 1990; Kendall, 1974, 1985) suggests that the adolescent will prefer the unreliable reinforcement and will be more likely to request the car from his mother rather than his father (although the father always provides access to the car!). We may question this outcome for our applied example because choosing the mother (unreliable reinforcement) results in a lower overall rate of reinforcement compared with the alternative that produces reliable reinforcement. Next, we wish to speculate about some variables that may influence an adolescent to respond in this manner.

Several studies have shown that preference for unreliable reinforcement may be influenced by the duration of the terminal link, that is, the length of time the pigeon or the adolescent has to wait for a specific reinforcer following an initial-link response. Basic research has provided preliminary evidence that increasing FT terminal-link durations results in an increase in the existing preference for the unreliable reinforcement alternative (Dunn & Spetch, 1990; Kendall, 1974, 1985). However, these findings hold only with signaled terminal-link outcomes. It appears that the distinct terminal-link stimuli play a role in the preference for unreliable reinforcement because when outcomes are unsig-

naled, subjects prefer reliable reinforcement (Spetch et al., 1990; Spetch & Dunn, 1987). In the applied example, these findings suggest that the adolescent will show a greater preference for mom (i.e., unreliable reinforcement) than for dad (i.e., reliable reinforcement) when the reinforcement schedules (car or no car) are correlated, or signaled by, specific reactions (i.e., terminal-link stimuli). Also, the preference for mom will be strongest when the delay to the car is the longest. What is it about the terminal-link stimuli that influence selection of an alternative that produces unreliable reinforcement?

Belke and Spetch (1994) and others (Dunn & Spetch, 1990; Fantino, 1969; Spetch et al., 1990) suggest that preference for unreliable reinforcement in pigeons is due to the effects of conditioned reinforcement. The authors propose that the terminal-link stimulus correlated with food (S+) in the unreliable alternative (50% reinforcement, mom) is established as a conditioned reinforcer because it occurs in the context of a terminal-link stimulus correlated with no food (S−). The S+ terminal link signals a reduction in delay to reinforcement following a response on the initial-link stimulus correlated with unreliable reinforcement. By contrast, the S+ terminal link correlated with reliable reinforcement always occurs following a response for that alternative during the initial link. The stimulus correlated with reliable reinforcement provides little information regarding a reduction in the delay to reinforcement. Therefore, the terminal-link stimulus correlated with reinforcement (on the unreliable alternative) is established as a relatively stronger conditioned reinforcer due to the context in which it occurs. This may override the lesser percentage reinforcement of the unreliable alternative.

To further evaluate the effects of context on conditioned reinforcement, Belke and Spetch (1994) modified the concurrent chains procedure used in previous studies. Their procedural variation prevented a pigeon from switching to the reliable reinforcement alternative (100% reinforcement, dad) in the choice phase following a no-food outcome on the unreliable reinforcement alternative (50% reinforcement, mom). That is, after a no-food outcome on the 50% alternative, only the initial-link stimulus for the 50% alternative (i.e., mom) was available in the initial link (i.e., a forced choice). A choice between the two alternatives was available only after reinforcement on the unreliable alternative. In our applied example, this is analogous to dad being unavailable as a choice (in the choice phase) until after a night when mom provides the car to the adolescent. Results showed that the pigeons preferred unreliable reinforcement. The authors suggested that the forced-choice procedure enhanced the conditioned reinforcement effects of the S+ terminal-link stimulus (for unreliable reinforcement) and strengthened the pigeons' preference for unreliable reinforcement because the S+ terminal-link stimulus occurred in a greater context of no-food outcomes on the unreliable alternative (i.e., mom sometimes provides access to the car). Given the basic research findings on context and conditioned reinforcement, we present two questions regarding our applied example. First, will mom's distinct reaction (i.e., reaction correlated with reinforcement) function as a relatively stronger conditioned reinforcer due to the context in which it occurs? Second, if so, will its conditioned reinforcing properties outweigh the greater relative rate of reinforcement available with the reliable reinforcement option and result in preference for unreliable reinforcement?

The effects of context and conditioned reinforcement on preference have not been as extensively investigated in applied studies as they have in basic research (Iwata & Michael, 1994; Redmon & Farris, 1987). Thus, it is not surprising that we lack applied empirical support for a prediction that mom's reaction (when she provides access to the car) will function as a conditioned reinforcer. Or, if the reaction from mom does function as a conditioned reinforcer, what is the mechanism responsible for its effect? Iwata and Michael stated that the effects of con-

ditioned reinforcement in a concurrent chains procedure are due to a change in stimulus conditions (i.e., white key to red, green, or yellow key). Therefore, the effectiveness of conditioned reinforcement will depend on the pre- and postchange conditions. In our applied example, when the adolescent chooses mother there is a degree of uncertainty regarding the specific outcome (i.e., car or no car). The reaction from mom correlated with no car (S−) shows that reinforcement is not forthcoming. Thus, there is no improvement from pre- to postchange conditions. The reaction from mom correlated with access to the car suggests that reinforcement is forthcoming, which is an improvement from pre- to postchange conditions. Based on this interpretation, its appears to be reasonable that the reaction from mom correlated with access to the car can function as a stronger conditioned reinforcer. However, can its effects result in the adolescent's preference for unreliable (mom) over reliable (dad) reinforcement?

To our knowledge, applied researchers have not experimentally examined human choice responding on concurrent chains schedules. Therefore, we lack direct evidence to answer our question regarding the adolescent's preference for unreliable reinforcement. However, we may turn to research on observing responses to answer our question indirectly. In basic research, an observing response occurs in the presence of a stimulus that produces access to a stimulus correlated with reinforcement or nonreinforcement. For example, using a multiple-schedule procedure, a pigeon is trained to respond in the presence of two successive stimuli. Each stimulus is correlated with a simple reinforcement schedule. After differentiated responding is obtained, the stimuli (correlated with the reinforcement schedules) are removed and are produced contingent on an observing response. Thus, we may consider the observing response as analogous to the choice phase (i.e., initial link) in the concurrent chains procedure because it produces access to stimuli correlated with a specific reinforcement schedule (rather than a specific outcome). The observing response is established as a conditioned reinforcer because of its pairing with a stimulus predictive of reinforcement. The observing response occurs more frequently when followed by a stimulus predictive of reinforcement compared with a stimulus correlated with extinction (Case & Fantino, 1981).

Findings from research on observing responses in humans showed that it is maintained by conditioned reinforcement (Case, Ploog, & Fantino, 1990; Mulvaney, Hughes, Jwaideh, & Dinsmoor, 1981). Case et al. (1990) studied the effects of conditioned reinforcement on college students' observing responses during a computer game. Observing responses were commands that produced stimuli that were either correlated with reinforcement (i.e., S+, information regarding an impending attack by invaders) or extinction (i.e., S−, uncertain information regarding an impending attack by invaders). In each condition of the experiment, the subjects preferred the stimulus correlated with reinforcement. In a similar study, Mulvaney et al. (1981) trained subjects to key press in the presence of a red light (S+) and to withhold pressing in the presence of a blue light (S−). After training was completed, the red and blue lights were turned off and a press on a white key (i.e., an observing response) was required to activate the key color and reinforcement schedule in effect. Subjects produced the observing response more frequently when it was followed by the red light than by the blue light. These findings appear to be consistent with the basic findings reported previously on the effects of conditioned reinforcement and tend to support the adolescent in our applied example favoring mom over dad. In the above examples, the stimuli correlated with reinforcement may have been established as a conditioned reinforcer because they occurred in the context of a stimulus correlated with extinction in the multiple-schedule arrangements.

The extensions of Belke and Spetch's (1994) findings to our applied example suggests a line

of applied research on the effects of context and conditioned reinforcement on human choice. The goal of this paper is not to provide mere speculation but to encourage applied researchers to pursue experiments that test the generality of the basic findings on context and conditioned reinforcement. The hypothetical examples given and the implications drawn suggest several possible lines of study. First, are the basic findings regarding preference for unreliable reinforcement in a choice situation applicable to humans in applied situations? Second, if the basic findings hold in applied situations, do the applied terminal-link stimuli (e.g., the distinct parental responses) play a similar role as their basic counterparts? Third, what mechanism accounts for the influence of conditioned reinforcement (i.e., applied terminal-link stimuli) in a preference for unreliable reinforcement? We hope that this paper encourages productive investigations regarding the role of context and conditioned reinforcement and their paradoxical effects on preference for unreliable reinforcement.

## REFERENCES

Belke, T. W., & Spetch, M. L. (1994). Choice between reliable and unreliable reinforcement alternatives revisited: Preference for unreliable reinforcement. *Journal of the Experimental Analysis of Behavior, 62,* 353–366.

Case, D. A., & Fantino, E. (1981). The delay-reduction hypothesis of conditioned reinforcement and punishment: Observing behavior. *Journal of the Experimental Analysis of Behavior, 35,* 93–108.

Case, D. A., Ploog, B. O., & Fantino, E. (1990). Observing behavior in a computer game. *Journal of the Experimental Analysis of Behavior, 54,* 185–199.

Dunn, R., & Spetch, M. L. (1990). Choice with uncertain outcomes: Conditioned reinforcement effects. *Journal of the Experimental Analysis of Behavior, 53,* 201–218.

Fantino, E. (1969). Choice and rate of reinforcement. *Journal of the Experimental Analysis of Behavior, 12,* 723–730.

Fantino, E., Preston, R. A., & Dunn, R. (1993). Delay reduction: Current status. *Journal of the Experimental Analysis of Behavior, 60,* 159–169.

Iwata, B. A., & Michael, J. L. (1994). Applied implications of theory and research on the nature of reinforcement. *Journal of Applied Behavior Analysis, 27,* 183–193.

Kendall, S. B. (1974). Preference for intermittent reinforcement. *Journal of the Experimental Analysis of Behavior, 21,* 463–473.

Kendall, S. B. (1985). A further study of choice and percentage reinforcement. *Behavioral Processes, 10,* 399–413.

Mace, F. C., McCurdy, B., & Quigley, E. A. (1990). A collateral effect of reward predicted by matching theory. *Journal of Applied Behavior Analysis, 23,* 197–205.

Mulvaney, D. E., Hughes, I. H., Jwaideh, A. R., & Dinsmoor, J. A. (1981). Differential production of positive and negative discriminative stimuli by normal and retarded children. *Journal of Experimental Child Psychology, 32,* 389–400.

Redmon, W. K., & Farris, H. E. (1987). Applications of basic research to the treatment of children with autistic and severely handicapped repertoires. *Education and Treatment of Children, 10,* 326–337.

Spetch, M. L., Belke, T. W., Barnet, R. C., Dunn, R., & Pierce, W. D. (1990). Suboptimal choice in a percentage-reinforcement procedure: Effects of signal condition and terminal-link length. *Journal of the Experimental Analysis of Behavior, 53,* 219–234.

Spetch, M. L., & Dunn, R. (1987). Choice between reliable and unreliable outcomes: Mixed percentage-reinforcement in concurrent chains. *Journal of the Experimental Analysis of Behavior, 47,* 57–72.

*Received May 22, 1995*
*Initial editorial decision June 7, 1995*
*Revision received June 14, 1995*
*Final acceptance June 15, 1995*
*Action Editor, F. Charles Mace*