RESPONSE ACQUISITION BY HUMANS WITH DELAYED REINFORCEMENT

HIROTO OKOUCHI

OSAKA KYOIKU UNIVERSITY

The present experiment examined whether a response class was acquired by humans with delayed reinforcement. Eight white circles were presented on a computer touch screen. If the undergraduates touched two of the eight circles in a specified sequence (i.e., touching first the upper-left circle then the bottom-left circle), then the touches initiated an unsignaled resetting delay culminating in point delivery. Participants experienced one of three different delays (0 s, 10 s, or 30 s). Rates of the target two-response sequence were higher with delayed reinforcement than with no reinforcement. Terminal rates of the target sequence decreased and postreinforcement pauses increased as a function of delay duration. Other undergraduates exposed to yoked schedules of response-independent point deliveries failed to acquire the sequence. The results demonstrate that a response class was acquired with delayed reinforcement, extending the generality of this phenomenon found with nonhuman animals to humans.

Key words: acquisition, reinforcement delay, points exchangeable for money, response sequence, screen touch, humans

———————————

Lattal and Gleeson (1990) demonstrated that, in the absence of shaping, bar pressing or key pecking could be acquired in free-operant sessions by food-deprived rats and pigeons when the response produced food after unsignaled resetting or nonresetting delays of up to 30 s. Since this seminal work appeared, the finding of response acquisition with delayed reinforcement has been replicated by several investigators (e.g., Anderson & Elcoro, 2007; Dickinson, Watt, & Griffiths, 1992; Galuska & Woods, 2005; Schlinger & Blakely, 1994; van Haaren, 1992; Wilkenfield, Nickel, Blakely, & Poling, 1992), with different response topographies (Critchfield & Lattal, 1993), different reinforcers (Galuska & Woods, 2005; Lattal & Metzger, 1994; LeSage, Byrne, & Poling, 1996; Snycerski, Laraway, & Poling, 2005), and different species (Galuska & Woods, 2005; Lattal & Metzger, 1994).

It is surprising that there is no published work demonstrating response acquisition with delayed reinforcement by humans because the generality of this effect extends from Siamese fighting fish (Lattal & Metzger, 1994) to rhesus

monkeys (Galuska & Woods, 2005). Based on casual observation of daily behavior, one may suppose that humans obviously can acquire a new response with delayed reinforcement. It seems self-evident, and, therefore, may have been ignored by investigators. Some findings from the experimental analysis of human behavior, however, suggest that this issue needs careful consideration.

First, human behavior is often said not to show the same sensitivity to schedules of reinforcement such as fixed-interval (FI), multiple fixed-ratio (FR) differential-reinforcement-of-low-rate (DRL), or concurrent variable-interval (VI) VI, as the behavior of nonhuman animals (e.g., Baron, Kaufman, & Stauber, 1969; Hayes, Brownstein, Zettle, Rosenfarb, & Korn, 1986; Horne & Lowe, 1993). Such human–nonhuman differences are still a matter of controversy. Some investigators have argued that important procedural variations between human and nonhuman research may account for many of the behavioral differences (e.g., Perone, Galizio, & Baron, 1988; Weiner, 1983), while others have suggested that there are fundamental differences in the principles that govern human and nonhuman behavior (e.g., Lowe, 1979; Wearden, 1988). Contingencies arranged in experiments of response acquisition with delayed reinforcement, such as tandem FR 1 differential-reinforcement-of-other-behavior (DRO) schedules (e.g., Lattal & Gleeson, 1990), are, of course, types of schedules of reinforcement.

Thus, it is unknown whether schedules of delayed reinforcement have the same effects on human and nonhuman behavior.

Second, verbal stimuli may influence the cases of response acquisition by humans with delayed reinforcement observed casually in daily life. Considerable research has shown formidable effects of verbal stimuli on human operant behavior (see reviews by Baron & Galizio, 1983; Kerr & Keenan, 1997; Vaughan, 1989). Instructing which response produces reinforcers can establish the response without shaping (e.g., Kaufman, Baron, & Kopp, 1966; Weiner, 1962) and describing a relation between responses and reinforcers can maintain the responses even though those responses never produce a reinforcer (e.g., Kaufman et al., 1966). Thus, it is plausible that humans can emit responses with delayed reinforcement when (a) the responses produce reinforcers or (b) the response–reinforcer relation was instructed. Previous findings provide no prediction whether humans can respond with delayed reinforcement without such instructions.

Finally, the generally employed class of response in the experimental analysis of human behavior may have made the demonstration of response acquisition difficult. Usually, one response to an operandum, such as a press to a key, has been defined as an operant (e.g., Weiner, 1962). The operant response has been established by instructions specifying that the response produces reinforcers in many studies (e.g., Catania, Matthews, & Shimoff, 1982; Okouchi, 2003; Torgrud & Holborn, 1990; Weiner, 1962) or by shaping in a few studies (Lowe, Beasty, & Bentall, 1983; Matthews, Shimoff, Catania, & Sagvolden, 1977). Because previous experiments have demonstrated that nonhuman animals acquired responses in the absence of shaping (e.g., Lattal & Gleeson, 1990), specific instructions or shaping should be omitted for a systematic replication using human subjects. Even though neither instructions nor shaping are given, some humans still may respond to the operandum. In Experiment 3 of Kaufman et al. (1966), for example, four college students pressed a key 7 to 161 times per minute with no reinforcer deliveries during the first 30 min of the experiment, although neither shaping nor instructions about the operandum were given. On the other hand,

the operant level of pressing a single bar for one college student was zero without instructions or shaping (DeGrandpre, Buskist, & Cush, 1990; also see Ader & Tatum, 1961; Ayllon & Azrin, 1964). Response acquisition is the transition from operant-level responding to steady-state responding (Sidman, 1960, pp. 117–119; Snycerski, Laraway, Huitema, & Poling, 2004). Thus, the operant level of the commonly-used class of response (a response to an operandum) often may be too high to detect reliable evidence of acquisition. This would be the case especially when the resetting delay is as long as 30 s, because this contingency restricts the rates of responding to around two responses per min to produce the maximum rates of reinforcement. By contrast, any response with an operant level at zero can not be acquired without shaping. Thus, an examination of the generality of response acquisition with delayed reinforcement using human participants requires a class of response with low but nonzero baseline rates (Critchfield & Lattal, 1993; Sidman, 1960, pp. 117–119), and it is unlikely that single responses to an operandum would meet this requirement.

The present experiment examined response acquisition by humans with delayed reinforcement. For this purpose, a two-response sequence was selected as the class of response because its operant level should be low but above zero. Eight filled white circles were presented on a black touch screen of a monitor. Touches on two of the eight circles in a specified sequence initiated an unsignaled 0-s, 10-s, or 30-s resetting delay, depending on the condition, culminating in point delivery exchangeable for money. Rates of the two-response sequences for participants exposed to these delayed reinforcement conditions were compared with (a) those when no reinforcer was delivered and (b) those for yoked-control participants to whom points were delivered independently from their performance.

METHOD

Participants

Nine male and 21 female undergraduates recruited from an educational psychology class at Osaka Kyoiku University served as participants. They were 19 to 29 years old, and none
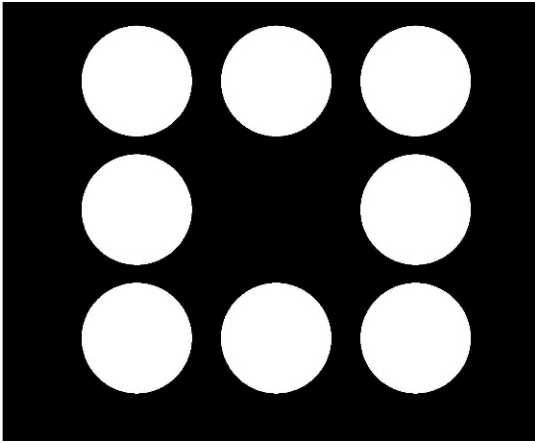
Fig. 1. A depiction of the stimuli displayed to the subject on the computer screen. A two-response sequence—a touch to the bottom-left circle following a touch to the upper-left circle—was the response class (target sequence) in the present experiment.

had experience with operant conditioning experiments.

Apparatus

The experimental room was 1.70 m wide, 2.20 m deep, and 2.17 m high. A Nihon Electric Company PC-9821AP microcomputer, located in an adjacent room, controlled the experiment. The participant sat at a desk facing a color display monitor (250 mm wide by 180 mm high) equipped with a Micro Touch Systems touch screen. During the experimental session, three, two, and three filled white circles (30-mm diameter for each) were presented on the top (45 mm below the top of the screen), middle (90 mm below the top of the screen), and bottom (135 mm below the top of the screen) of the black screen, respectively. Figure 1 shows the configuration of these eight circles. At the top and bottom of the screen, three circles were presented at the left, center, and right, respectively, 45 mm apart from center to center, whereas two circles were presented at the left and right in the middle of the screen, respectively, 90 mm apart from center to center. When the schedule requirement was met, these eight circles disappeared and one filled white circle (30 mm diameter) for the ''consummatory response'' was presented in the middle center of the screen (not shown in Figure 1). All interevent times were recorded in real time with a 50-ms resolution. Each

touch to the eight circles was accompanied by a low tone (880 Hz), and the appearance of the middle-center circle (delivery of a reinforcer) and each touch to that circle (consuming the reinforcer) were accompanied by a high tone (1760 Hz). A digital counter was located at the top right of the screen.

Class of Response

The operant level of the response class in the present experiment needed to be low. To determine that class, operant levels of touching the eight circles were measured for 5 undergraduates other than the ones participating in the present experiment. The procedure in this pilot experiment was identical to that of the nonreinforcement condition described below. Thus, these pilot participants were exposed to the screen showing the eight circles (Figure 1), were told that they might earn points by touching the circles, but actually no points were delivered. Under this condition, every touch to each circle was recorded for 10 min. One of the most infrequent two-response sequences, touching the upper-left circle then the bottom-left one, was selected as the present response class. During the 10 min, this sequence of responses did not occur for 4 of the 5 pilot participants and occurred only once for the other participant.

Procedure

Participants signed an informed consent agreement that specified the frequency and duration of their participation and the average earnings for such participation. They agreed to remain in the experiment for a maximum of eight 90-min experimental periods. At the beginning and at the end of the experiment, each participant was asked not to speak to anyone other than the experimenter about the study in an attempt to prevent discussion about the contingencies among participants (cf. Horne & Lowe, 1993). At the end of the experiment, each participant was asked whether he or she had any information to offer about the study. All reported that they did not.

A 90-min experimental period was conducted once per day, two times per week. During this period, a maximum of seven variable-duration sessions occurred. Sessions were separated by 2- to 3-min breaks. After every experimental period, participants were paid for their perfor-

mance (2 yen per 100 points, approximately .018 U.S. dollars). Upon completion of the entire experiment, participants were paid for their participation (200 yen per 90 min) and were debriefed. The overall earnings for each participant who completed the entire experiment ranged from 2,320 to 3,370 yen (approximately 20.88 to 30.33 U.S. dollars).

On the first day of the experiment, each participant was asked to leave wristwatches, cellular phones, and books outside the experimental room. Once in the room, the participant was asked to read silently the following instructions (translated here from Japanese to English):

> Your task is to earn as many points as you can. A hundred points are worth two yen. Payment for the points will be made at the end of each visit. In addition, you will be paid 200 yen for every day you spend in the experiment. Payment for participation will be made at the end of the last visit.
> Eight circles will be shown on the screen of the display monitor. Sometimes these circles may disappear, then a circle will appear in the center of the screen. By touching the center circle, you can earn points. Accumulated points will be shown in the top right of the screen.
> It is up to you whether you touch the eight circles or not. If you touch these circles, the center circle may appear. Or, the appearance of the center circle may be unrelated to touching the eight circles.
> The words ''READY'' and ''GO'' will appear in sequence on the screen. When the word ''GO'' disappears, do the task until the words ''GAME OVER'' appear on the screen.

The typed set of instructions remained on the desk throughout the experiment. Questions regarding the experimental procedure were answered by telling the participant to reread the appropriate sections of the instructions. Then the words ''READY'' and ''GO'' were presented in sequence at the top left of the display monitor. After the word ''GO'' disappeared, eight circles were presented on the screen as shown in Figure 1.

When the schedule requirement was met, the eight circles were darkened and the circle for the consummatory response was presented in the middle center of the screen. A touch darkened the circle and accumulated 100 points on the top right counter, followed by the reappearance of the eight circles. After the session terminated, the words ''GAME OVER'' appeared at the top left of the screen.

Participants were assigned to one of six conditions: one of three conditions of sequence-dependent delayed reinforcement, or one of three yoked control conditions of sequence-independent reinforcement. All participants were exposed first to a nonreinforcement condition, then to a reinforcement condition. The procedure in the nonreinforcement condition was the same for all participants and identical to that of the pilot experiment. In this condition, a 10-min single session was conducted to measure the operant level of the target response sequence (upper-left-then-bottom-left sequence).

The procedure of the reinforcement condition differed across participants depending on the condition to which they were assigned. Participants in the conditions of sequence-dependent delayed reinforcement were exposed to a tandem FR1 DRO schedule. Under this schedule, a touch to the bottom left circle after a touch to the upper left circle initiated a delay of 0 s (the 0-s delay condition), 10 s (the 10-s delay condition), or 30 s (the 30-s delay condition) to reinforcement, then the center circle appeared for the consummatory response (reinforced). Any two-response sequence of upper-left-then-bottom-left that occurred during a delay restarted the delay interval. Responses other than this target sequence were without consequences. Each session lasted until 45 reinforcers occurred or 60 min (including the time taken for consummatory responses) elapsed, whichever came first. Each participant was exposed to this schedule for 20 sessions or until eight 90-min experimental periods were completed, whichever came first.

Participants in the yoked control conditions of sequence-independent reinforcement were exposed to identical contingencies to those of the sequence-dependent delayed reinforcement conditions with the following exceptions. The frequency and distribution of sequence-independent reinforcers, the session duration, and the number of sessions for each participant were yoked to those of one of the participants exposed to the tandem FR1 DRO schedules.

Participant Screening

Although the pilot data suggested that operant levels of the present response class would be low, it was still uncertain whether the level for each participant would be suitable for

answering the present research question. Thus, participants were initially screened according to their operant levels.

First, because a relatively high level of responding makes valid evaluation of acquisition difficult, participants emitting the target sequence more than four times during the 10-min nonreinforcement session were eliminated from further study. Second, because responses must occur to be reinforced (Wilkenfield et al., 1992), participants never emitting the target sequence during the nonreinforcement and the first reinforcement sessions (70 min in total) were screened out. Finally, participants emitting the sequence less than a total of five times from the nonreinforcement session through the second reinforcement session (130 min in total) also were screened out because a sequence with such a low operant level might not be strengthened by reinforcement. For example, one pigeon in Experiment 1 of Lattal and Gleeson (1990) responded nine times during the first two sessions of a tandem FR1 fixed-time (FT) 30-s schedule but ceased responding for the next five sessions. The first criterion was applied to all participants, whereas the other two were applied only to participants in the sequence-dependent delayed reinforcement conditions.

The first, second, and third criteria, respectively, resulted in the elimination of one (the 30-s delay condition), three (one in the 0-s delay and two in the 10-s delay conditions) and two participants (the 30-s delay condition). Furthermore, two participants in the 30-s delay condition discontinued the experiment with no sign of response acquisition. One of them dropped out after exposure to two sessions of the tandem FR1 DRO 30-s schedule because of personal reasons (unrelated to the ongoing experiment) and another was dismissed during the fourth session of the delayed contingency because of a violation of the contract (bringing a book into the experimental room and reading it during the session). Thus, four, four, and three pairs of response-dependent delayed reinforcement participants and their yoked partners completed the 0-s, 10-s, and 30-s delay conditions, respectively.

RESULTS

Figures 2, 3, and 4, respectively, show the number of target sequences per minute for each participant exposed to the 0-s, 10-s, and 30-s delays to reinforcement in each session. Data for the yoked participants are shown on the same panels as their partners in the delayed reinforcement condition. The rates of the sequence during the nonreinforcement session were zero or close to zero for all participants, suggesting that the operant levels of the target sequence were extremely low.

When each target sequence produced a reinforcer immediately (the 0-s delay condition), rates of the target sequence increased for all 4 participants (Figure 2). Although there was some variability in target sequence rate across sessions and participants, the rates for at least the last 19 sessions for all participants were higher than those during the nonreinforcement session and those of the yoked participants.

Figure 3 shows the data from the 10-s delay condition. For all sessions for Participants E5 and E6, and for the last 19 sessions for Participant E8, rates of the target sequences under the tandem FR1 DRO 10-s schedule were higher than those for the no reinforcement session and those for the yoked participants under sequence-independent reinforcement. Participant E7 experienced only 12 sessions of the 10-s delay of reinforcement, because his erratic performance during the initial 10-s delay sessions consumed considerable time. During the first, second, third, fourth, and fifth sessions for this participant, numbers of single responses to any of the eight circles, target sequences, and reinforcers, respectively, were 387, 5, and 5, 1077, 3, and 3, 72, 0, and 0, 5154, 15, and 15, and 5526, 28, and 22. Each of these sessions lasted 60 min. For the last seven sessions, however, rates of the target sequences for Participant E7 were consistently higher than that for the nonreinforcement session and those of his yoked partner, C7.

Participants and their yoked partners received fewer than 20 reinforcement sessions under the tandem FR1 DRO 30-s schedule (Figure 4). Because a tandem FR1 DRO 30-s schedule required relatively longer times to obtain each reinforcer, none of the delayed-reinforcement participants completed 20 sessions within the contracted eight, 90-min, experimental periods. For all 15 and 14 sessions for Participants E13 and E15, respectively, and for the last 11 sessions for Partici-
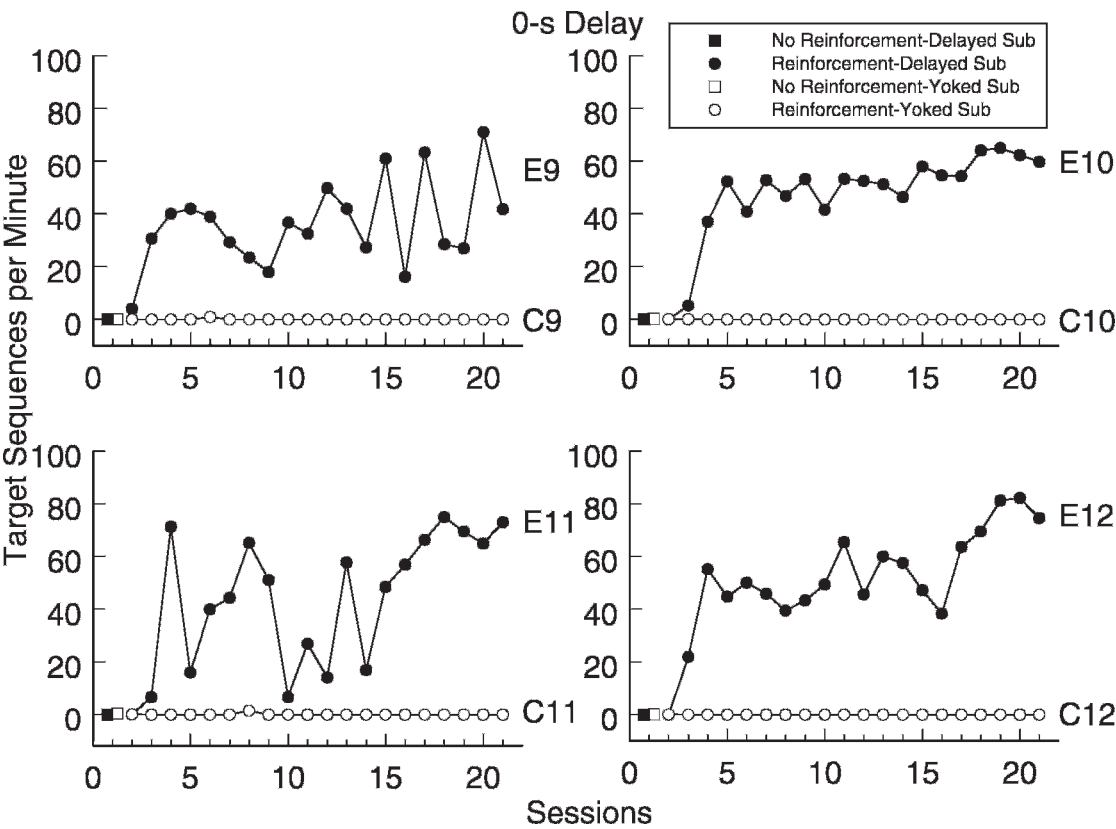
Fig. 2. Rates of the target sequence in each session for each participant exposed to the 0-s delay to reinforcement and for the yoked participant. Filled and open squares represent rates of the sequence in the no-reinforcement condition for participants subsequently exposed to the delayed reinforcement and for their yoked partners, respectively. Filled circles represent rates of the sequence in the delayed-reinforcement condition, whereas open circles represent rates of the sequence in the sequence-independent reinforcement condition.

pant E14, rates of the target sequences under the tandem FR1 DRO 30-s schedule were higher than those under no reinforcement and those for the yoked participants under the sequence-independent reinforcement.

Another measure of response acquisition is the ratio of the number of responses that produced reinforcers to the total number of responses. Some investigators have used two levers (with rats as subjects) and compared responses on a lever that produced food with those on a lever on which responses had no programmed consequences (e.g., Keely, Feola, & Lattal, 2007; Sutphin, Byrne, & Poling, 1998). In the present experiment, the frequency of the two-response target sequence (touching first the upper-left circle, then the bottom-left circle) can be compared to the frequency of all other two-response sequences. These two-response sequences were counted

only from the immediately preceding reinforcer to the next reinforcer, but were not counted across a reinforcer. For example, events may occur in the following order:

UL, BL, UL, SR+, BL, UL,

BL, MR, UC, SR+,

where UL, BL, MR, and UC are touches to the upper-left, bottom-left, middle-right, and upper-center circles, respectively, and SR+ is a reinforcer. In this example, the two-response sequences are UL–BL, BL–UL, BL–UL, UL–BL, BL–MR, and MR–UC. The numbers of target sequences, nontarget sequences, and total sequences are two, four, and six, respectively, whereas the number of total responses is eight.

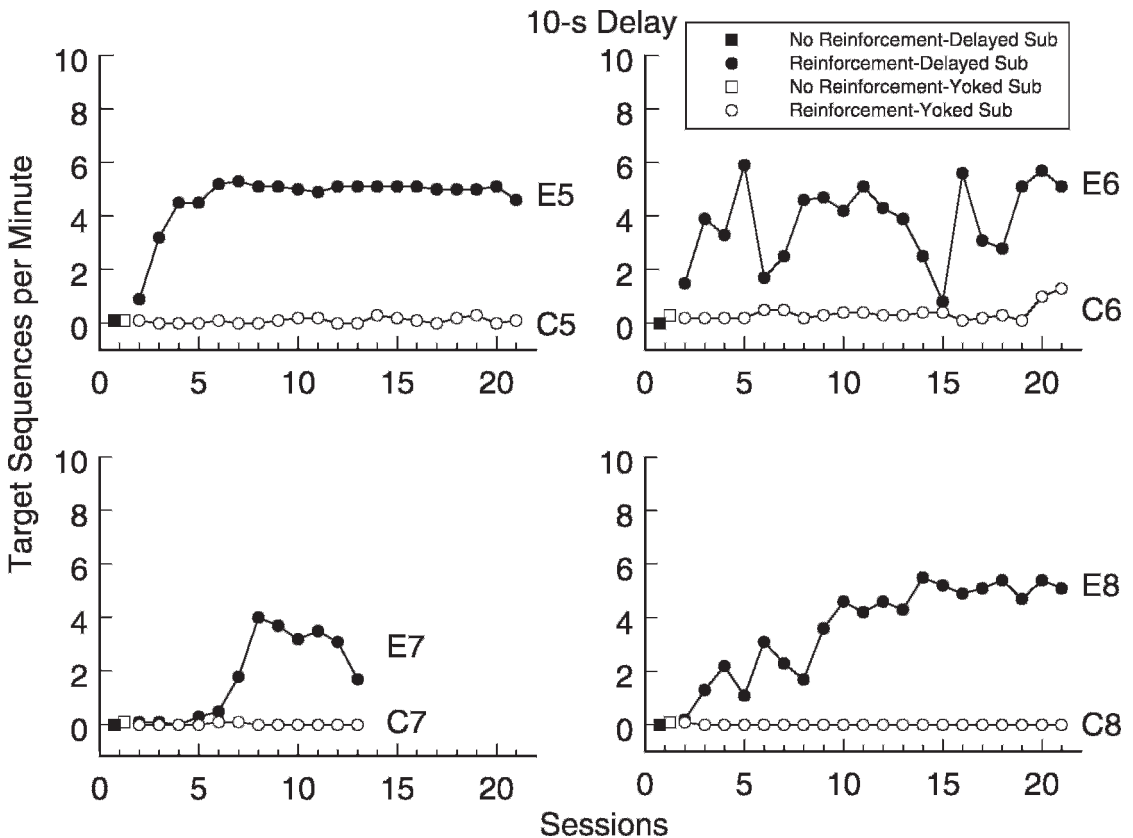Figures 5, 6, and 7 show target sequences as a percentage of total sequences for each

Fig. 3. Rates of the target sequence in each session for each participant exposed to the 10-s delay to reinforcement and for the yoked participant. Details are as in Figure 2.

participant exposed to the 0-s, 10-s, and 30-s delays to reinforcement in each session and that for the yoked participant. With this measure, a value of 100 indicates that the target sequence occurred exclusively throughout the session, whereas a value of 0 indicates that nontarget sequences occurred exclusively. Missing data points indicate that there were no two-response sequences during those sessions. Table 1 shows the median of single responses (not sequences) to any of the eight circles per minute in the last five sessions of the delayed reinforcement or the sequence-independent yoked reinforcement conditions.

Consistent with the previous analysis of target sequence rates (Figure 2), the percentages of target sequences for at least the last 19 sessions of the 0-s delay to reinforcement were higher for all participants than those during the nonreinforcement session and those for the last 19 sessions for the yoked participants with sequence-independent reinforcement

(Figure 5). The percentages during at least the final nine sessions with the 10-s delayed reinforcement were higher for all participants than those during the nonreinforcement session and those during respective sessions for the yoked participants (Figure 6). The percentages during the last 7 and 13 sessions with the 30-s delayed reinforcement for Participants E13 and E15, respectively, were higher than those during the nonreinforcement session and those during respective sessions for the yoked participants (Figure 7). Although the percentages for Participant E14 were lower than those for the other two participants in the 30-s delayed reinforcement condition, close scrutiny revealed that these percentages showed an increasing trend with exposure to this contingency and those during the last nine sessions were higher than that during the nonreinforcement session. The missing data for the yoked participant's (C14) respective sessions indicate that she
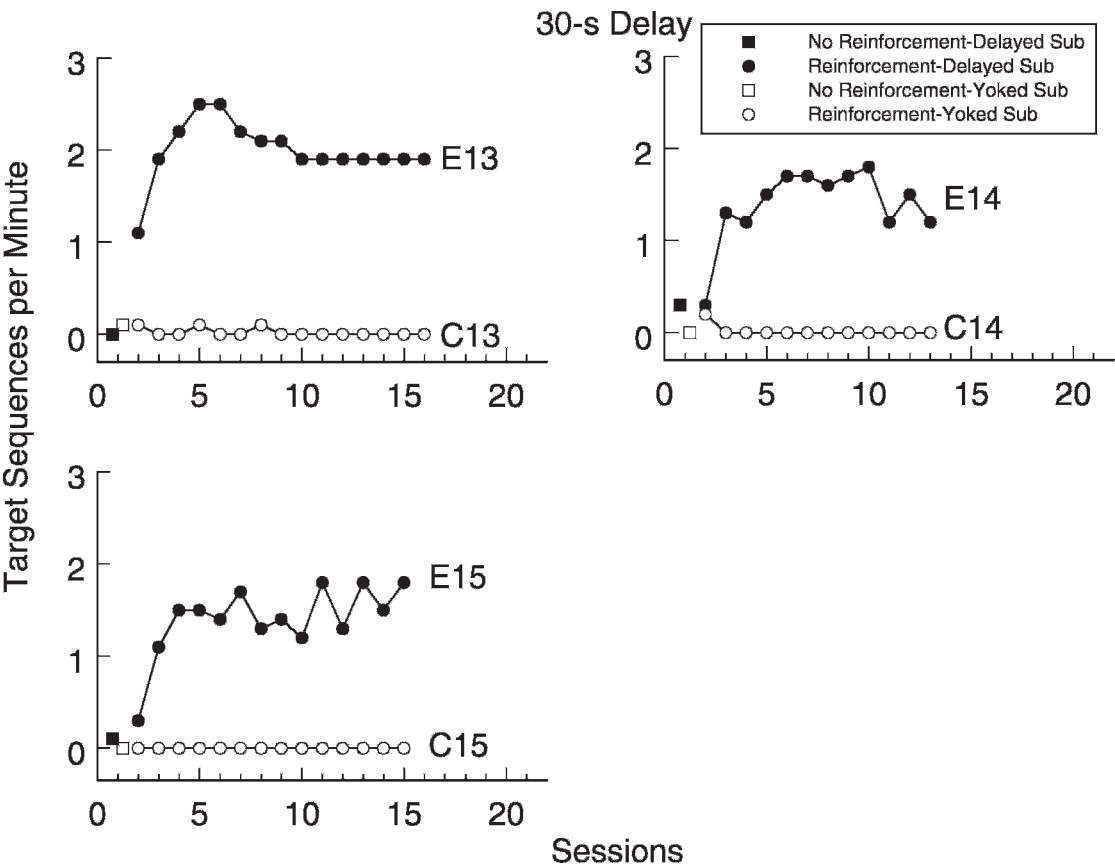
Fig. 4.   Rates of the target sequence in each session for each participant exposed to the 30-s delay to reinforcement and for the yoked participant. Details are as in Figure 2.

ceased all responding (see also Table 1), and this suggests that the target sequence developed differentially for her partner E14 who was exposed to delayed reinforcement.

Percentages of the target sequences were generally low for some participants exposed to delayed reinforcement (i.e., Participants E6, E7, and E14), although they were higher than those for the yoked partners (Figures 6 and 7). For these participants, rates of single responses to any of the eight circles were relatively high (Table 1). These results raise two questions. First, did the target sequence occur more frequently than any of the nontarget two-response sequences? Second, was the target sequence acquired as a single unit or was it embedded in a larger sequence of responses? Table 2 shows the two-response sequence that occurred most frequently and the rank of the frequency of the target sequence during the last five sessions of the 10-s and 30-s delays to

reinforcement. The target sequence occurred most frequently for all participants except for Participants E7 and E14 for whom the frequency of the target sequence ranked quite low. It should be noted, however, that relative frequencies of the most frequent sequences for these participants (MR–BC for E7 and BL–UR for E14) were low also, and comprised only 4.5% (E7) and 6.2% (E14) of all two-response sequences.

Table 2 shows the sizes of total sequences that occurred from the immediately preceding reinforcer to the next reinforcer. Sequences which were repeated most frequently (dominant sequences) also are shown. Participants E5, E8, and E15 emitted exclusively the target sequence 222, 164, and 163 times out of 225 opportunities, respectively, during the last five sessions, suggesting that the target sequence was acquired as a single unit. By contrast, certain long sequences that included the
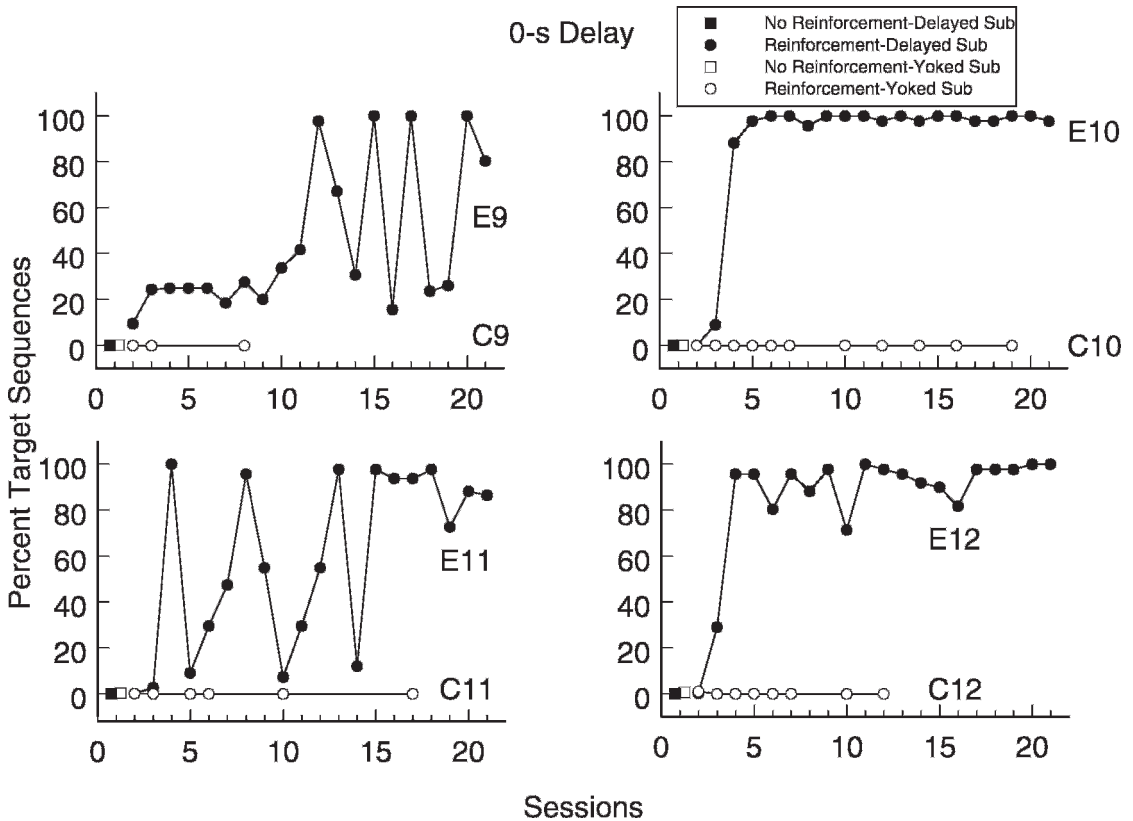
Fig. 5.   Percentages of the number of the target sequence to the number of all two-response sequences in each session for each participant exposed to the 0-s delay to reinforcement and for the yoked participant. Filled and open squares represent percentages of the target sequence in the no-reinforcement condition for participants subsequently exposed to the delayed reinforcement and for their yoked partners, respectively. Filled circles represent percentages of the target sequence in the delayed reinforcement condition, whereas open circles represent percentages of the target sequence in the sequence-independent reinforcement condition. Missing data indicate that percentages could not be calculated because no two-response sequences occurred.

target sequence tended to be repeated for Participants E13 and E14. Participant E13 emitted the sequence UL–BL–UR–BR 224 times out of 225 opportunities, whereas Participant E14 emitted a 64-response sequence, in which the target sequence was embedded in the 18th and 19th responses, 150 times out of 225 opportunities. Long sequences occurred for Participants E6 and E7 also, but none were repeated more than 12 times.

It is also possible to compare the results of each response-dependent condition as a function of the delay duration. The left panel of Figure 8 plots the medians of the target sequence rates across the last five sessions for participants in each condition of delayed reinforcement. The rates decreased as the delay interval increased. The right panel of

Figure 8 plots the medians of the postreinforcement pauses (PRPs; time between the termination of the consummatory response period and the first occurrence of the target sequence) across the last five sessions for participants in each condition of delayed reinforcement. The PRPs increased as the delay interval increased, again suggesting a delay-of-reinforcement-gradient consistent with the results of the sequence rates.

DISCUSSION

In general, response acquisition under conditions of delayed reinforcement has been assessed by comparing response rates with (a) those under no reinforcement (e.g., Lattal & Gleeson, 1990; Wilkenfield et al., 1992), (b)
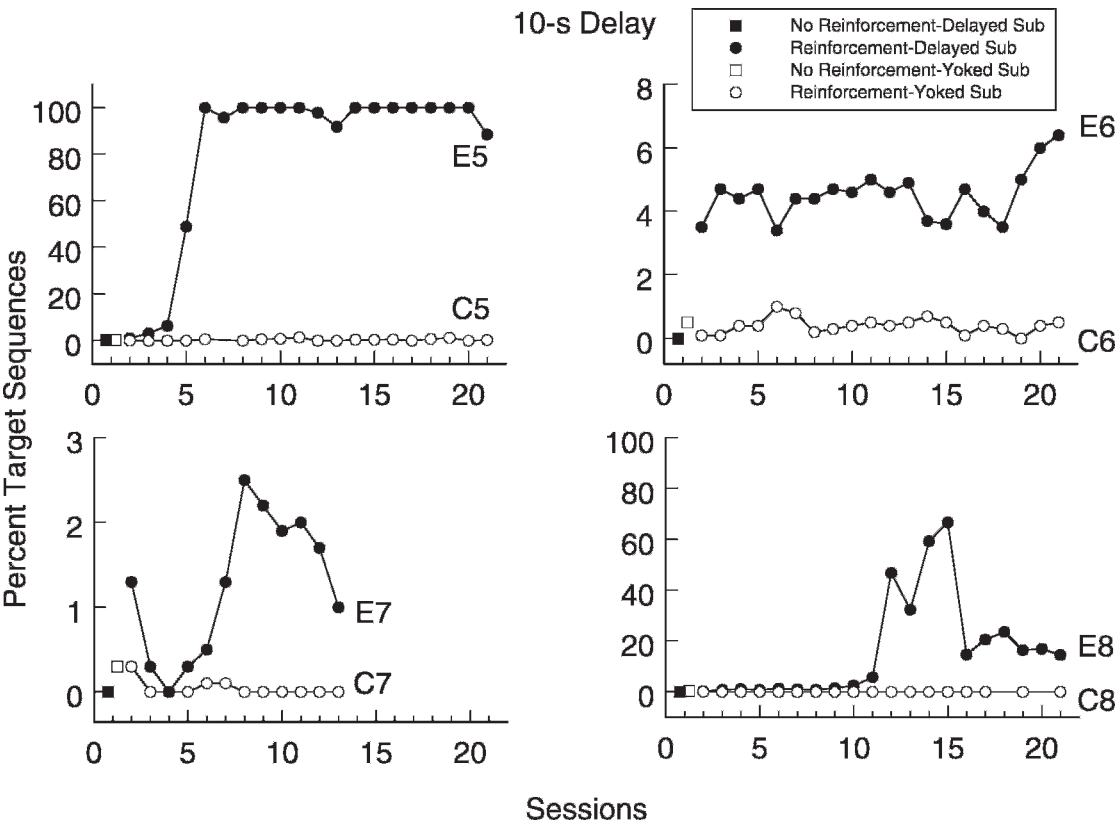
Fig. 6. Percentages of the number of the target sequence to the number of all two-response sequences in each session for each participant exposed to the 10-s delay to reinforcement and for the yoked participant. Details are as in Figure 5.

those under response-independent-yoked reinforcement (e.g., Dickinson et al., 1992; Galuska & Woods, 2005; Lattal & Gleeson, 1990), and (c) rates of different responses that produce no reinforcer (e.g., Critchfield & Lattal, 1993; Galuska & Woods, 2005; Keely et al., 2007; Wilkenfield et al., 1992). In terms of the present experiment, a response may be considered acquired by delayed reinforcement when the number of the target sequences per min and the number of target sequences as percentage of all two-response sequences are higher (a) when the target sequences were reinforced than when they were not reinforced, and (b) for participants delivered reinforcers depending on their performance than for participants delivered reinforcers independently. By these criteria, all 7 participants who were exposed to the sequence-dependent delayed reinforcement conditions acquired the two-response sequence.

Previous experiments have shown response acquisition with a delay of 30 s in pigeons (Lattal & Gleeson, 1990) and rats (Critchfield & Lattal, 1993; Lattal & Gleeson, 1990; Lattal & Williams, 1997; Snycerski et al., 2005; but see also van Haaren, 1992). The present experiment replicates these results with 3 human participants.

A number of experiments using nonhuman subjects have found that response rates show a delay-of-reinforcement gradient (e.g., Dickinson et al., 1992; Lattal & Metzger, 1994; Reilly & Lattal, 2004), whereas, to our knowledge, only Reilly and Lattal have found gradients for PRPs. The delay-of-reinforcement gradients found for both response rates and PRPs in the present experiment provide additional evidence of the generality of this phenomenon across species, and they demonstrate the consistency of the present results with those of previous experiments. However, the delay-
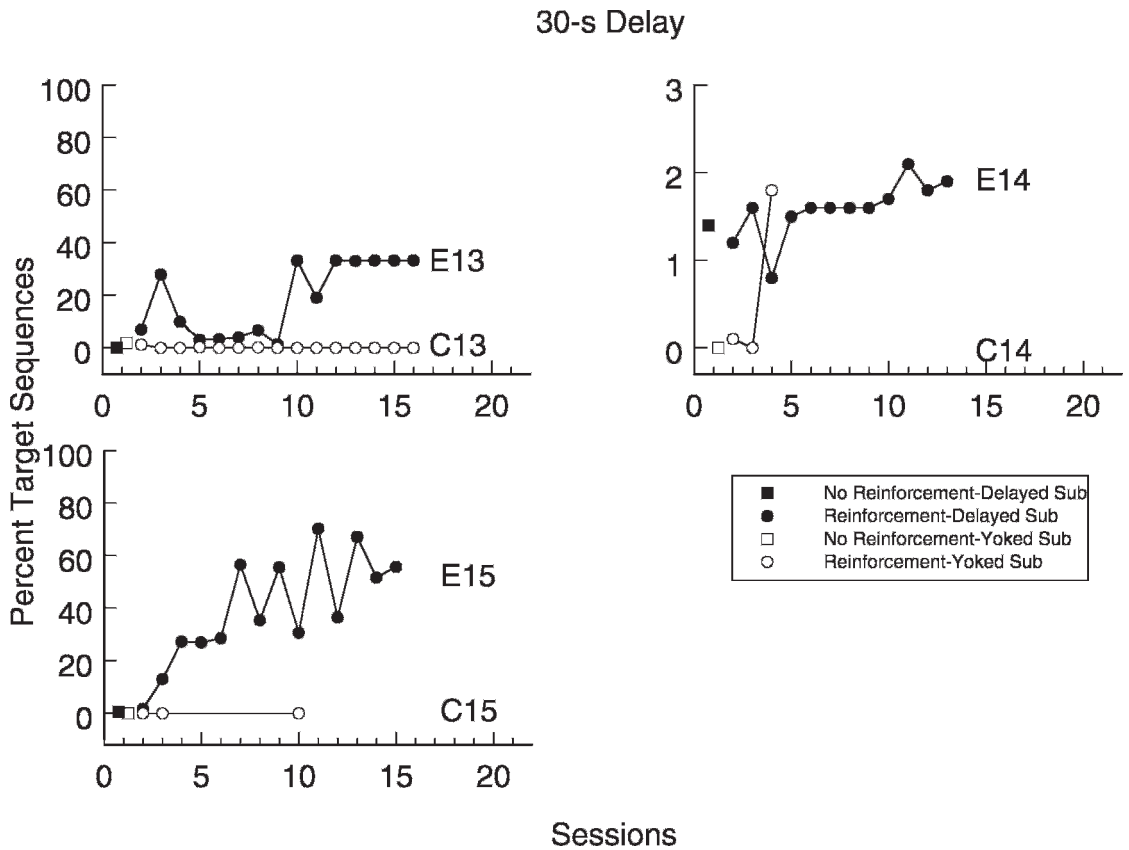
30-s Delay

Fig. 7. Percentages of the number of the target sequence to the number of all two-response sequences in each session for each participant exposed to the 30-s delay to reinforcement and for the yoked participant. Details are as in Figure 5.

Table 1

The median of single touches to any of the eight circles per minute (ranges in parentheses) for each participant exposed to delayed reinforcement (left) and for the yoked participant (right) in the last five sessions of each condition.

Participant	Delayed (E)	Yoked (C)
0-s Delay		
9	130.1 (93.7–149.2)	0.0 (0.0–0.0)
10	124.6 (109.8–130.1)	0.0 (0.0–10.1)
11	151.6 (138.3–165.2)	1.4 (0.0–5.9)
12	149.3 (128.6–164.6)	0.0 (0.0–0.0)
10-s Delay		
5	10.0 (9.7–10.1)	34.1 (22.8–47.2)
6	82.6 (77.4–106.5)	241.3 (49.2–280.6)
7	173.4 (169.9–184.7)	17.0 (3.9–21.7)
8	33.3 (28.3–40.2)	5.4 (5.4–7.0)
30-s Delay		
13	7.7 (7.7–7.7)	35.4 (31.9–41.7)
14	80.7 (59.5–108.6)	0.0 (0.0–0.0)
15	4.5 (4.3–5.0)	0.0 (0.0–0.0)

Table 2

Two-response sequences occurring most frequently, ranks of frequency of the target sequence, median sizes of sequence per reinforcer (ranges in parentheses), and dominant sequences. Data are pooled across the last five sessions of training for each participant exposed to delayed reinforcement.

Participant	Two-response sequence		Total sequence	
	Most frequent sequence	Rank of target sequence	Size of sequence	Dominant sequence[a]
10-s delay				
E5	UL-BL	1	2 (2–4)	UL-BL
E6	UL-BL	1	23 (8–256)	No dominant sequence
E7	MR-BC	24	90 (36–564)	No dominant sequence
E8	UL-BL	1	2 (2–109)	UL-BL
30-s delay				
E13	UL-BL	1	4 (4–8)	UL-BL-UR-BR
E14	BL-UR	19	64 (35–129)	64-response sequence
E15	UL-BL	1	2 (2–16)	UL-BL

Note. U, M, B, L, C, and R, denote upper, middle, bottom, left, center, and right, respectively, and indicate the location of circles touched. UL–BL, for example, is the target sequence of touching first the upper-left then the bottom-left circles. a. See text for details.

of-reinforcement gradients obtained from the present experiment have a limitation: the resetting delay contingency can restrict rates of responding (Critchfield & Lattal, 1993; Sutphin et al., 1998). Under such a contingency, the most efficient performance is to respond once immediately after the prior reinforcement, then to pause until a reinforcer occurs. Thus, even response rates produced by efficient performance can decrease as the delay interval is increased. It is very possible, therefore, that the resetting delay contingency contributed to the delay-of-reinforcement gradients found with response rates. Although the resetting delay contingency should not affect PRPs directly, it may have had an indirect
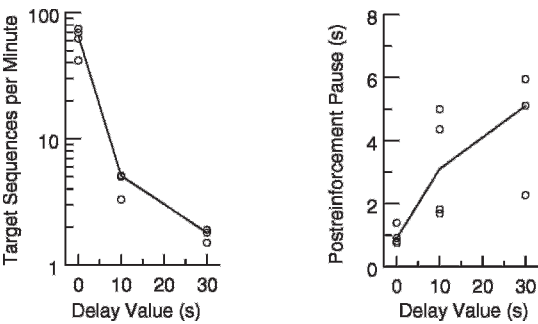

Fig. 8. Medians of the rates of the target sequence (left) and the postreinforcement pauses (right) for the last five sessions in each condition of delayed reinforcement. Unfilled circles show data from individual participants and solid lines connect group medians.

influence; that is, a general reduction in response rate produced by the resetting delay contingency could contribute to the PRP's results.

On time-based or interval schedules, organisms can emit behavior that is not required by the contingency (Pierce & Epling, 1995, p. 139). When lever pressing produced food according to a VI schedule, for example, rats also drank an extreme amount of water even though this was unrelated to contingencies of reinforcement (Falk, 1961). In the context of response acquisition with delayed reinforcement, rats pressed not only a lever which produced food with resetting delay but also another lever on which responses had no programmed consequences (e.g., Sutphin et al., 1998; Wilkenfield et al., 1992). Consistent with these previous findings, 4 of 7 participants exposed to the resetting delay contingency frequently emitted responses other than the target sequences. These results suggest that for some participants the target two-response sequence was acquired not as a single unit but as a member of a longer sequence. These results, however, do not alter the conclusion that the target sequence was acquired. Whether the target sequence was a single unit or it was embedded in a longer sequence of responses, the target two-response sequences occurred at higher overall and relative rates with delayed reinforcement than with no reinforcement or with response-independent reinforcement.

The present results may suggest that response acquisition with delayed reinforcement is general across classes of responses. Pecking a key (Lattal & Gleeson, 1990), pressing a lever (e.g., Lattal & Gleeson, 1990; Wilkenfield et al., 1992), and breaking a photocell beam (Critchfield & Lattal, 1993; Lattal & Metzger, 1994) have been acquired in previous experiments. The present experiment found that not only single responses but more complex classes of response (i.e., touching two circles in a particular sequence) also can be acquired with delayed reinforcement. However, a feature of the present results, by contrast, limits the generality across response classes with humans as subjects. As described in the Introduction, the present experiment employed a two-response sequence as a response class because humans often respond to a single operandum without reinforcement (Kaufman et al., 1966). Performance during the nonreinforcement session of the present experiment highlights this problem; there were often high operant levels of single responses. The 22 participants touched any of the eight circles 10 to 1055 (median = 191) times during the 10-min nonreinforcement session. Thus, although the present experiment showed acquisition of a two-response sequence, these high rates of single responses suggest it would be more difficult to demonstrate acquisition of a simple response to a single operandum in humans.

REFERENCES

Ader, R., & Tatum, R. (1961). Free-operant avoidance conditioning in human subjects. *Journal of the Experimental Analysis of Behavior, 4,* 275–276.

Anderson, K. G., & Elcoro, M. (2007). Response acquisition with delayed reinforcement in Lewis and Fischer 344 rats. *Behavioural Processes, 74,* 311–318.

Ayllon, T., & Azrin, N. H. (1964). Reinforcement and instructions with mental patients. *Journal of the Experimental Analysis of Behavior, 7,* 327–331.

Baron, A., & Galizio, M. (1983). Instructional control of human operant behavior. *The Psychological Record, 33,* 495–520.

Baron, A., Kaufman, A., & Stauber, K. A. (1969). Effects of instructions and reinforcement feedback on human operant behavior maintained by fixed-interval reinforcement. *Journal of the Experimental Analysis of Behavior, 12,* 701–712.

Catania, A. C., Matthews, B. A., & Shimoff, E. (1982). Instructed versus shaped human verbal behavior: Interactions with nonverbal responding. *Journal of the Experimental Analysis of Behavior, 38,* 233–248.

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.

DeGrandpre, R. J., Buskist, W., & Cush, D. (1990). Effects of orienting instructions on sensitivity to scheduled contingencies. *Bulletin of the Psychonomic Society, 28,* 331–334.

Dickinson, A., Watt, A., & Griffiths, W. J. H. (1992). Free-operant acquisition with delayed reinforcement. *The Quarterly Journal of Experimental Psychology, 45B,* 241–258.

Falk, J. L. (1961, January 20). Production of polydipsia in normal rats by an intermittent food schedule. *Science, 133,* 195–196.

Galuska, C. M., & Woods, J. H. (2005). Acquisition of cocaine self-administration with unsignaled delayed reinforcement in rhesus monkeys. *Journal of the Experimental Analysis of Behavior, 84,* 269–280.

Hayes, S. C., Brownstein, A. J., Zettle, R. D., Rosenfarb, I., & Korn, Z. (1986). Rule-governed behavior and sensitivity to changing consequences of responding. *Journal of the Experimental Analysis of Behavior, 45,* 237–256.

Horne, P. J., & Lowe, C. F. (1993). Determinants of human performance on concurrent schedules. *Journal of the Experimental Analysis of Behavior, 59,* 29–60.

Kaufman, A., Baron, A., & Kopp, R. E. (1966). Some effects of instructions on human operant behavior. *Psychonomic Monograph Supplements, 1,* 243–250.

Keely, J., Feola, T., & Lattal, K. A. (2007). Contingency tracking during unsignaled delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 88,* 229–247.

Kerr, K. P. F., & Keenan, M. (1997). Rules and rule-governance: New directions in the theoretical and experimental analysis of human behaviour. In K. Dillenburger, M. F. O'Reilly, & M. Keenan (Eds.), *Advances in behaviour analysis* (pp. 205–226). Dublin: University College Dublin Press.

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

Lattal, K. A., & Metzger, B. (1994). Response acquisition by Siamese fighting fish (*Betta splendens*) with delayed visual reinforcement. *Journal of the Experimental Analysis of Behavior, 61,* 35–44.

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.

LeSage, M. G., Byrne, T., & Poling, A. (1996). Effects of *D*-amphetamine on response acquisition with immediate and delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 66,* 349–367.

Lowe, C. F. (1979). Determinants of human operant behaviour. In M. D. Zeiler, & P. Harzem (Eds.), *Advances in analysis of behavior: Vol. 1. Reinforcement and the organization of behaviour* (pp. 159–192). Chichester, England: Wiley.

Lowe, C. F., Beasty, A., & Bentall, R. P. (1983). The role of verbal behavior in human learning: Infant performance on fixed-interval schedules. *Journal of the Experimental Analysis of Behavior, 39,* 157–164.

Matthews, B. A., Shimoff, E., Catania, A. C., & Sagvolden, T. (1977). Uninstructed human responding: Sensitivity to ratio and interval contingencies. *Journal of the Experimental Analysis of Behavior, 27,* 453–467.

Okouchi, H. (2003). Stimulus generalization of behavioral history. *Journal of the Experimental Analysis of Behavior, 80*, 173–186.

Perone, M., Galizio, M., & Baron, A. (1988). The relevance of animal-based principles in the laboratory study of human operant conditioning. In G. Davey, & C. Cullen (Eds.), *Human operant conditioning and behavior modification* (pp. 59–85). New York: Wiley.

Pierce, W. D., & Epling, W. F. (1995). *Behavior analysis and learning*. Englewood Cliffs, NJ: Prentice Hall.

Reilly, M. P., & Lattal, K. A. (2004). Within-session delay-of-reinforcement gradients. *Journal of the Experimental Analysis of Behavior, 82*, 21–35.

Schlinger, H. D., & Blakely, E. (1994). The effects of delayed reinforcement and a response-produced auditory stimulus on the acquisition of operant behavior in rats. *The Psychological Record, 44*, 391–409.

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. New York: Basic Books.

Snycerski, S., Laraway, S., Huitema, B. E., & Poling, A. (2004). The effects of behavioral history on response acquisition with immediate and delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 81*, 51–64.

Snycerski, S., Laraway, S., & Poling, A. (2005). Response acquisition with immediate and delayed conditioned reinforcement. *Behavioural Processes, 68*, 1–11.

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.

Torgrud, L. J., & Holborn, S. W. (1990). The effects of verbal performance descriptions on nonverbal operant responding. *Journal of the Experimental Analysis of Behavior, 54*, 273–291.

van Haaren, F. (1992). Response acquisition with fixed and variable resetting delays of reinforcement in male and female Wistar rats. *Physiology and Behavior, 52*, 767–772.

Vaughan, M. (1989). Rule-governed behavior in behavior analysis. A theoretical and experimental history. In S. C. Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 97–118). New York: Plenum.

Wearden, J. H. (1988). Some neglected problems in the analysis of human operant behavior. In G. Davey, & C. Cullen (Eds.), *Human operant conditioning and behavior modification* (pp. 197–224). New York: Wiley.

Weiner, H. (1962). Some effects of response cost upon human operant behavior. *Journal of the Experimental Analysis of Behavior, 5*, 201–208.

Weiner, H. (1983). Some thoughts on discrepant human-animal performances under schedules of reinforcement. *The Psychological Record, 33*, 521–532.

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.