Learn Behav (2013) 41:238–255
DOI 10.3758/s13420-012-0100-8

Serial discrimination reversal learning in pigeons as a function of signal properties during the delay of reinforcement

Bertram O. Ploog · Ben A. Williams

Published online: 6 January 2013
© Psychonomic Society, Inc. 2013

Abstract Pigeons learned a series of reversals of a simultaneous red–green discrimination with a 6-s delay of reinforcement. The signal properties during the 6-s reinforcement delay were varied across blocks of reversals, such that the delay was either unsignaled (intertrial interval conditions during the delay) or signaled by illumination of the center key. Four different signal conditions were presented: (1) signals only after S+ responses, (2) signals only after S– responses, (3) differential signals after S+ versus S– responding, and (4) the same nondifferential signals after S+ and S– responses. (A zero-delay control condition was also included.) Learning was at a high level in the S+-only and differential-signal conditions, and learning was at a low level during the unsignaled, nondifferentially signaled, and S– signal conditions. Thus, a differential stimulus contingent on correct choices was necessary for proficient learning-to-learn, even though within-reversal learning occurred in all conditions. During the S+ and differential-signal conditions, improvement in learning continued to occur even after more than 240 reversals (more than 38,000 trials).

Keywords Serial-discrimination reversal learning · Un/signaled reinforcement delay · Learning-to-learn effect · Key peck · Pigeon

B. O. Ploog
College of Staten Island, Center for Developmental Neuroscience, and Graduate School & University Center of the City University of New York, Staten Island, NY, USA

B. A. Williams
University of California, La Jolla, San Diego, CA, USA

B. O. Ploog (✉)
Department of Psychology, 4S-105, College of Staten Island, CUNY, 2800 Victory Boulevard, Staten Island, NY 10314, USA
e-mail: bertram.ploog@csi.cuny.edu

The concept of conditioned reinforcement has a venerable but disputed status. The core of the concept is that an initially neutral stimulus gains conditioned value via its association with a primary reinforcer such as food. Thus, conditioned reinforcers share the properties of primary reinforcers, although presumably those properties are weaker. Those challenging the concept have not questioned whether putative conditioned reinforcers have powerful behavioral effects, but whether conditioned value is the proper characterization of the basis of those effects. Consequently, they have proposed the notion of a "sign post" by which the discriminative properties of the nominal conditioned reinforcer guide behavior, rather than actually strengthen it (see Shahan, 2010, for a discussion).

Many different experimental procedures have been used to study conditioned reinforcement effects, so it is not surprising that different interpretations have ensued. Although the classic early studies used discrete-trial procedures (e.g., Grice, 1948, with rats), the majority of later studies have employed free-operant procedures that primarily have used response rate, or in some cases choice, rather than the rate of learning (see Williams, 1994, for a review). A notable exception to this pattern was Williams and Dunn (1991a, with rats and pigeons), who studied the rate of learning of a two-choice simultaneous discrimination in which the differential consequences of correct versus incorrect choices were a 1-s tone plus food versus no feedback. In order to study within-subjects learning effects, repeated reversals of the values of the S+ and S– stimuli were used, with different conditions used for different reversals.

In Experiment 1 of Williams and Dunn (1991a), the effects of three conditions of a simultaneous conditional discrimination task were compared: (1) 100 % of the correct choices were followed by a tone plus food; (2) only 50 % of the correct choices were followed by the tone plus food, with the unreinforced correct choices receiving no feedback (i.e., being treated like incorrect choices); (3) only 50 % of the correct choices were reinforced with tone plus food—

just as under Condition 2—but the remaining 50 % of correct choices were followed by the tone without food (this manipulation being the most critical one for the present study). The major findings were that adding the tone on the remaining 50 % of correct but unreinforced trials under Condition 3 greatly reduced the number of errors and food reinforcers required to learn the discrimination reversals, relative to 50 % reinforcement without the additional tones (Condition 2). Adding the tones (Condition 3) also reduced the number of food reinforcers, as compared to Condition 1, which had 100 % reinforcement. Additional experiments demonstrated that the tone had no consistent effect when it was not paired with food, and that the substitutability of the tone for food was diminished when the tone's pairing with food was reduced from 50 % to 33 %. These results are consistent with a conditioned-value interpretation of the stimulus effects, but are not obviously explainable by the idea of a sign post (see the Discussion section).

The present study extends the investigation of conditioned reinforcement effects in simultaneous-discrimination learning. Here, however, a delayed-reinforcement contingency was used, and the issue was how various stimulus conditions during the delay-of-reinforcement intervals would affect the rate of learning. In the basic condition to which the others were compared, a differential signal was presented in the delay-of-reinforcement interval leading to food, with no feedback on trials with incorrect responses other than a return to the conditions of the intertrial interval (ITI). Other conditions included (1) adding a second differential signal after incorrect responding, (2) having a differential stimulus appear only after an incorrect response, (3) having no signal in the delay interval, and (4) having the same stimulus occur after correct and incorrect responses.

Because the serial-discrimination reversal learning (SDRL) procedure has commonly been used to study "learning to learn" (LTL), its use here also provides information about whether LTL occurs with delayed reinforcement and how LTL is affected by different stimulus conditions during the delay-of-reinforcement interval. SDRL typically produces increasingly rapid learning as a function of the number of repeated reversals. Such improvement in the rate of learning has been interpreted as an index of comparative intelligence (see Sutherland & Mackintosh, 1971, for a review), although it varies as a function of several procedural variables, including the response requirements (Williams, 1971a) and the ITI (Ploog & Williams, 2010; Williams, 1971b, 1976). However, the range of procedural variables studied in conjunction with serial reversal learning has been limited, and LTL may not occur under all conditions. Ploog and Williams reported greatly attenuated learning of repeated reversals when an unsignaled 2-s delay separated correct responses from food, relative to when immediate reinforcement was used. Also, unpublished work indicated that unsignaled delays as short as 6 s prevented any improvement in learning over successive reversals, although within-reversal learning did occur.

In the present study, we again used a delay-of-reinforcement contingency during training on serial discrimination reversals. In most conditions, a 6-s delay intervened between a choice response and its outcome. The principal manipulations were the stimuli during the delay intervals.

Our prior unpublished work failed to show improvement across reversals when a 6-s unsignaled delay was interposed between the choice and outcome, despite the fact that the individual reversals themselves were learned to a reasonable level of accuracy. Therefore, the first set of conditions (Phases 1 through 4) were designed to determine the reliability of the prior unpublished findings and to compare the unsignaled delay to other conditions in which either a differential or a nondifferential signal was presented during the delay. A zero-delay condition was also included, to determine the extent to which the signals during the delay (putative conditioned reinforcers) substituted for an immediate primary reinforcement. Sample sizes of $n = 4$ or 8 were sufficient to accomplish this first part of our investigation. The second set of conditions (Phases 5–9) extended the investigation to the effects of added differential stimuli following incorrect responding. Because the effect size of this manipulation was unknown, an increase in statistical power seemed advisable, by increasing the sample size to $n = 15$.

Method

Subjects

A group of 16 White Carneaux pigeons served in this experiment. (One bird died during the course of the study.) The birds were housed in individual cages under a 10-h:14-h dark:light cycle, with water and grit always available. They were maintained at 80 % of their free-feeding weights by the food (Purina, Gold Pellets) obtained during experimental sessions and by supplements in the home cages. All had served in a previous experiment (Ploog & Williams, 2010) with a similar design, but in which the ITI and reinforcement delay, instead of the signal properties during the reinforcement delay, were manipulated.

Apparatus

The experiment was carried out in four identical three-key standard pigeon conditioning chambers (Scientific Prototype, Model B 200) 42 cm in height, with a 38.0 × 30.5 cm wire mesh floor (elevated by 7.0 cm). The keys, with diameters of 3.2 cm, were located 24.0 cm above the wire mesh floor. These keys were separated by 9.2 cm, center to center. The keys could be illuminated by 12-stimulus in-line projectors

with 2.8-W incandescent light bulbs (Type 1820X) and operated with a minimum force of 0.18 N. Red and green stimulus colors were used for the left and right keys, and yellow and blue stimulus colors for the center key. The grain magazine (BRS/LVE, Model GFM-001) was accessible through a 6.0-cm-wide and 5.0-cm-high opening, which was located 5.5 cm above the wire mesh floor below the center key. When operated, the magazine access opening was illuminated by a 1.1-W incandescent light bulb (Type 28PSB). The front wall (hinged on the bottom), the ceiling, and the back wall of each chamber were made of clear acrylic. The left and right walls were made of bronze-colored 2-mm-thick sheet metal. The right wall served as the intelligence panel on which the keys, the stimulus projectors, and the hopper were mounted. The chambers were placed in sound-attenuated enclosures (Scientific Prototype, Model SPEC 2) that opened at the front. The enclosures had inside dimensions of 45.5 × 62.5 × 33.0 cm (H × W × D). Each was equipped with a 0.3-W loudspeaker (7.5-cm diameter) mounted in the far upper left corner, providing white noise to mask extraneous sounds, and a fan to provide air circulation and additional masking noise. A 5-W incandescent light bulb (Type FG 616) was mounted on the ceiling, centered above the conditioning chamber. It served as a houselight, which was lit throughout each experimental session, including when the hopper was operated. The enclosures containing the conditioning chambers were located in a sound-attenuated room, which was dark during the experimental sessions. The experimental room was adjacent to the room that contained the computer (Apple, Macintosh G3 "Blue & White Minitower"), which controlled the experimental events and performed data collection. The interface consisted of a multipurpose I/O card (National Instruments, PCI-DIO-96) and opto-relays to read keypecks and to control stimulus presentations. The software was custom-written in C/C++ (Metrowerks, CodeWarrior 9.0).

Procedure

Given that the birds had prior experience, all were started with the simultaneous discrimination procedure using the delay-of-reinforcement contingencies. A session started with an 8-s ITI. After the ITI, the left and right keys were illuminated with a red and a green light. Pecks to one stimulus (S+) resulted in termination of both keylights and led to a 6-s reinforcement delay (except in Phase 4—see below—when the delay was 0 s), before 2 s of access to grain was provided. After reinforcement, the next ITI was initiated. Pecks to the alternative stimulus (S–) also resulted in termination of both keylights and led to the same delay, but no reinforcement was provided. Instead, the next ITI started immediately after the delay. A total of five pecks (FR 5) to either the S+ or the S– (counted separately) was required to initiate the delay of reinforcement. The positions of the S+ and S– were random, with the following restrictions: (1) A correction procedure was used after S– choices, such that the positions of the colors remained the same on the next trial. (2) The positions never remained the same for more than three consecutive trials, unless the correction procedure was in effect. (3) Not considering the correction trials, each stimulus was presented on both sides equally often. Each session comprised 80 trials (correction trials included). Depending on the pigeons' performance (i.e., time to complete the FR 5 requirement), the session duration varied but was about 25 to 30 min long on average, with each trial lasting about 20 s. With very rare exceptions, all of the birds completed each session.

The stimulus conditions during the delays were the variables of interest, yielding six conditions. The signal or signals were presented on the center key for the entire duration of the delay. Under Condition Unsig, the stimulus situation during the delay following S+ and S– responses was identical to the ITI situation (i.e., houselight but no keylight on). Under Condition SigS+, a yellow keylight was presented after responses to the S+ but not to the S–. Under Condition NDiff, the same yellow keylight was presented after S+ and S– responses. Under Condition Diff, a blue or yellow keylight was presented after S+ responses, and the alternative keylight was presented after S– responses. Under Condition SigS–, a blue or yellow keylight (the S– stimulus of Condition Diff) was presented after responses to the S– but not after responses to the S+. Finally, under Condition Zero, no signal was presented, because there was no delay.

Each condition began with two sessions with one color/trial-outcome correlation (e.g., red/S+, green/S–). After two sessions with the original correlation, the correlation was reversed. From then on, a reversal occurred after every two sessions, thus resulting in alternation between the original and the reversed correlation every other session. (For presentation of the results, the first and second sessions of a given reversal will be referred to as *Sessions 1 and 2*.) Each condition remained in effect until 20 reversals (40 sessions) had been completed.

The specific assignments of birds to the stimuli and the orders of conditions are presented in Table 1. Our initial interest was in determining whether our unpublished research showing minimal LTL with the unsignaled-delay condition was reliable, and also how it compared with different signaled-delay conditions using the same delay values. Consequently, the order of the conditions was counterbalanced, resulting in different orders of conditions for different subjects for the first four conditions. Thereafter, all birds received all conditions in the same order.

Data analysis

On the basis of previous experience (Ploog & Williams, 2010) and because of concerns about the range limitations (from 0 to 1.0) of proportions correct (p), we present all of the results as logit p—that is, $\log_{10}[p / (1 - p)]$—which is less susceptible to ceiling or floor effects (cf. Williams, 1998). To evaluate performance under the Results below,

Table 1 Assignments of birds to stimuli, conditions, and phases

	Phase (S+/S–)[1]								
Bird	1	2	3	4	5	6	7	8	9
Set 1	SigS+	NDiff	Unsig	Zero[2]	Diff	SigS+	SigS–	NDiff	Unsig
GR83	(Y/X)	(Y/Y)	(X/X)	—	(B/Y)	(Y/X)	(X/Y)	(Y/Y)	(X/X)
GR84	(Y/X)	(Y/Y)	(X/X)	—	(B/Y)	(Y/X)	(X/Y)	(Y/Y)	(X/X)
GR85	(Y/X)	(Y/Y)	(X/X)	—	(Y/B)	(Y/X)	(X/B)	(Y/Y)	(X/X)
GR95	(Y/X)	(Y/Y)	(X/X)	—	(Y/B)	(Y/X)	(X/B)	(Y/Y)	(X/X)
Set 2	Unsig	NDiff	SigS+	Zero[2]	Diff	SigS+	SigS–	NDiff	Unsig
GR81	(X/X)	(Y/Y)	(Y/X)	—	(B/Y)	(Y/X)	(X/Y)	(Y/Y)	(X/X)
GR92	(X/X)	(Y/Y)	(Y/X)	—	(B/Y)	(Y/X)	(X/Y)	(Y/Y)	(X/X)
GR94	(X/X)	(Y/Y)	(Y/X)	—	(Y/B)	(Y/X)	(X/B)	(Y/Y)	(X/X)
GR97	(X/X)	(Y/Y)	(Y/X)	—	(Y/B)	(Y/X)	(X/B)	(Y/Y)	(X/X)
Set 3	SigS+	NDiff	Unsig	SigS+	Diff	SigS+	SigS–	NDiff	Unsig
GR82	(Y/X)	(Y/Y)	(X/X)	(Y/X)	(B/Y)	(Y/X)	(X/Y)	(Y/Y)	(X/X)
GR87	(Y/X)	(Y/Y)	(X/X)	(Y/X)	(B/Y)	(Y/X)	(X/Y)	(Y/Y)	(X/X)
GR91	(Y/X)	(Y/Y)	(X/X)	(Y/X)	(Y/B)	(Y/X)	(X/B)	(Y/Y)	(X/X)
GR93	(Y/X)	(Y/Y)	(X/X)	(Y/X)	(Y/B)	(Y/X)	(X/B)	(Y/Y)	(X/X)
Set 4	Unsig	NDiff	SigS+	SigS+	Diff	SigS+	SigS–	NDiff	Unsig
BL90	(X/X)	(Y/Y)	(Y/X)	(Y/X)	Died				
BL91	(X/X)	(Y/Y)	(Y/X)	(Y/X)	(B/Y)	(Y/X)	(X/Y)	(Y/Y)	(X/X)
BL92	(X/X)	(Y/Y)	(Y/X)	(Y/X)	(Y/B)	(Y/X)	(X/B)	(Y/Y)	(X/X)
BL93	(X/X)	(Y/Y)	(Y/X)	(Y/X)	(Y/B)	(Y/X)	(X/B)	(Y/Y)	(X/X)

[1] (S+/S–)=Signals after S+ and S– during delay. B=blue, Y=yellow, and X=dark key. [2] Condition Zero did not allow for presentation of any signal

note that logit p = 1.0 is equivalent to 90 % correct. The $p < .05$ level was used for all significance tests.

Results

All phases

Figure 1 provides an overview of performance in all nine phases as a function of condition (in blocks of eight consecutive sessions, collapsed over reversals). Each graph panel represents one set of birds (all sets with $n = 4$ except for Set 4, Phases 5–9, with $n = 3$, because one of the birds died in the course of the experiment). The different conditions were presented as within-subjects factors in nine phases consisting of 20 two-session reversals, also allowing for some between-group comparisons.

Overall, Conditions SigS+, Zero, and Diff produced high levels of performance, whereas Conditions NDiff, Unsig, and SigS– produced low levels. For the high-performance conditions, Condition Zero, not surprisingly, produced the highest level of discrimination, indicating its superiority over all delay-of-reinforcement conditions, even when the delay was signaled. That is, an immediate signal (a putative conditioned reinforcer) is less effective than immediate primary reinforcement. For the cluster of low-performance conditions, the differences between conditions were small and dependent on the order of the conditions, and required the more detailed analyses to be presented below.

All conditions that generated high performance generally resulted in improved discrimination across the 20 reversals within a phase (i.e., the slope of the functions within a panel is positive). In contrast, Conditions NDiff (Phases 2 and 8, all sets) and Unsig (Phase 1 in Sets 2 and 4; Phase 3 in Sets 1 and 3; and Phase 9, all sets) generated either no improvement or even a decline in performance within a phase. Condition SigS– (Phase 7, all sets), which generally produced low performance, appeared to produce some improvement across the 20 reversals of Phase 7, but an analysis of variance (ANOVA) comparing performance for the first and the last ten reversals failed to reach significance, $F(1, 14) = 3.01$, $p = .105$.

In what follows, we will present more fine-grained analyses related to the effects of principal interest: the effects of the six different signal properties during the delay-of-reinforcement interval (SigS+, SigS–, Diff, NDiff, Zero, and Unsig) and LTL. Because of the complexity and extent of the statistical analyses, a summary of all significant effects will be provided in Tables 2, 3, 4, 5, 6, 7, and 8, each of which corresponds to Figs. 2, 3, 4, 5, 6, 7, and 8, respectively, presenting *df*, *MS*, *F*, and *p*, with partial η^2 as a measure of effect size. For the sake of brevity and

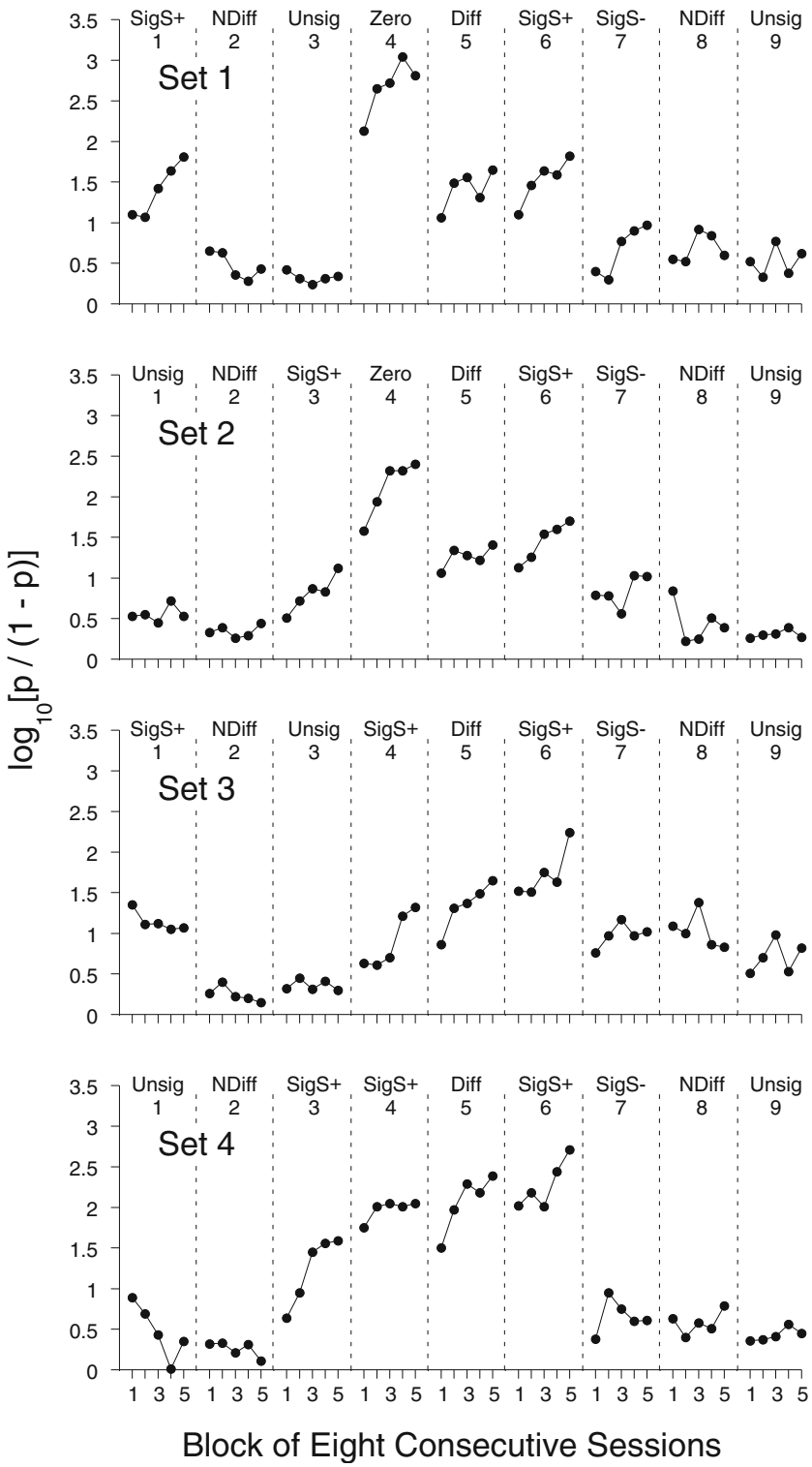

Fig. 1 Performance (logit p) for four sets of birds across nine phases as a function of block of eight sessions (collapsed over reversals) and signal properties (conditions). All birds within one set received the same treatment in a given phase ($n = 4$ for all sets, except $n = 3$ for Set 4 in Phases 5–9)

readability, the values of only the significant effects will be reported in the tables and not repeated in the text.

Learning to learn

Our experience with the SDRL preparation has been that performance continues to improve (albeit slowly) with continued training, despite hundreds of prior reversals (cf. Ploog & Williams, 2010). Thus, we were interested in the underlying basis of this improvement, as well as in the pattern of change with continued training. From Table 1, it is evident that several conditions were presented at least twice, in earlier and later phases of training. By comparing performance in the earlier versus later phases of training

Table 2 Summary of the ANOVA corresponding to Fig. 2

Source	*df*	*MS*	*F*	*p*	Partial η^2
Condition SigS+: Comparison of Phases 1 or 3 vs. 6 (Fig. 2, top):					
Phase	1, 14	40.37	25.73	< .001	0.65
Half-Phase	1, 14	11.18	25.15	< .001	0.64
Session	1, 14	56.91	168.88	< .001	0.92
Quarter-Session	3, 42	91.47	171.08	< .001	0.92
Phase × Session	1, 14	3.65	10.49	< .01	0.43
Half-Phase × Session	1, 14	1.75	6.00	< .05	0.30
Phase × Quarter-Session	3, 42	2.97	13.33	< .001	0.49
Half-Phase × Quarter-Session	3, 42	0.45	4.36	< .01	0.24
Session × Quarter-Session	3, 42	1.95	8.74	< .001	0.38
Condition SigS+: Comparison of Phases 4 vs. 6 (Fig. 2, bottom):					
Phase	1, 6	20.63	16.12	< .01	0.73
Half-Phase	1, 6	7.77	15.54	< .01	0.72
Session	1, 6	10.13	40.30	< .005	0.87
Quarter-Session	3, 18	55.29	58.05	< .001	0.91
Phase × Session	1, 6	2.14	7.67	< .05	0.56
Session × Quarter-Session	3, 18	0.38	3.50	< .05	0.37
Phase × Session × Quarter-Session	3, 18	0.41	4.01	< .05	0.40

under identical stimulus conditions, an assessment could be made of the molar changes over continued training (*LTL across phases*), and also of the specific aspects that improved when LTL did occur. A second measure of the LTL effect was the degree of improvement, under one condition, within a phase (i.e., comparing the first with the second half of a phase; *LTL within phases*). An open question is how these two measures of LTL are related, as it is possible that a molar improvement across phases could occur with or without an improvement within phases.

SigS+: Earlier versus later conditions Figure 2 illustrates the LTL effect across phases under Condition SigS+, which produced very high levels of performance. The data are plotted as values of logit p as a function of blocks of 20 trials (*quarter-sessions*). To illustrate LTL across phases, the top panel shows the comparison for Phase 1 (Sets 1 and 3) or 3 (Sets 2 and 4) versus Phase 6 (Sets 1–4; $n = 15$); the bottom panel shows the comparison for Phase 4 (Sets 3 and 4) versus Phase 6 (Sets 3 and 4; $n = 7$). Filled circles show early phases, whereas open circles show later phases. The functions are shown separately for the first and second halves of a phase (*half-phase*: Sessions 1–20 on the left and Sessions 21–40 on the right, to assess LTL within phases) and for Sessions 1 and 2 of each reversal (left and right pairs of functions in each panel, respectively). ANOVAs were conducted (summarized in Table 2) with Phase, Half-Phase, Session, and Quarter-Session as within-subjects factors (all under Condition SigS+).

For both comparisons—that is, the comparisons of Phases 1 or 3 versus 6 (Fig. 2, top) and of Phases 4 versus 6 (Fig. 2,

Table 3 Summary of the ANOVA corresponding to Fig. 3

Source	*df*	*MS*	*F*	*p*	Partial η^2
Condition Unsig: Comparison of Phases 1 or 3 vs. 9 (Fig. 3, top):					
Session	1, 14	13.34	39.29	< .001	0.74
Quarter-Session	3, 42	12.82	51.92	< .001	0.79
Phase × Session	1, 14	5.49	15.58	< .005	0.53
Phase × Session × Quarter-Session	3, 42	0.29	3.98	< .05	0.22
Condition NDiff: Comparison of Phases 2 vs. 8 (Fig. 3, bottom):					
Phase	1, 14	16.69	14.34	< .005	0.51
Session	1, 14	9.89	41.12	< .001	0.75
Quarter-Session	3, 42	12.11	49.63	< .001	0.78
Phase × Quarter-Session	3, 42	0.96	6.80	< .005	0.33

Table 4 Summary of the ANOVA corresponding to Fig. 4

Source	*df*	*MS*	*F*	*p*	Partial η^2
Comparison of SigS+ (Phase 3) vs. Diff (Phase 5)—Fig. 4, top:					
Condition	1, 6	23.38	35.57	< .005	0.86
Half-Phase	1, 6	5.57	8.65	< .05	0.59
Session	1, 6	41.89	279.90	< .001	0.98
Quarter-Session	3, 18	37.60	85.81	< .001	0.94
Condition × Quarter-Session	3, 18	1.84	14.69	< .001	0.71
Session × Quarter-Session	3, 18	0.96	5.70	< .01	0.49
Comparison of SigS+ (Phase 4) vs. Diff (Phase 5)—Fig. 4, middle:					
Condition	1, 6	4.77	8.78	< .05	0.59
Half-Phase	1, 6	10.43	19.39	< .01	0.76
Session	1, 6	30.54	158.58	< .001	0.96
Quarter-Session	3, 18	46.83	41.39	< .001	0.87
Half-Phase × Quarter-Session	3, 18	0.77	4.88	< .05	0.45
Session × Quarter-Session	3, 18	0.46	3.64	< .05	0.38
Comparison of SigS + (Phase 6) vs. Diff (Phase 5)—Fig. 4, bottom:					
Condition	1, 14	6.12	11.70	< .005	0.46
Half-Phase	1, 14	12.06	35.09	< .001	0.72
Session	1, 14	59.22	145.07	< .001	0.91
Quarter-Session	3, 42	109.66	148.34	< .001	0.91
Condition × Session	1, 14	4.25	23.94	< .001	0.63
Half-Phase × Session	1, 14	3.61	10.61	< .01	0.43
Half-Phase × Quarter-Session	3, 42	0.48	3.32	< .05	0.19
Session × Quarter-Session	3, 42	2.00	13.19	< .001	0.49
Half-Phase × Session × Quarter-Session	3, 42	0.59	4.85	< .01	0.26

bottom)—performance in the later Phase 6 was notably higher than that in the earlier phases (i.e., open circles are above filled circles, showing LTL across phases). The effect of Phase was significant for both comparisons (Table 2), confirming that performance improved with continued training (i.e., LTL across phases), even when additional phases with different delay-of-reinforcement stimulus conditions separated the early versus late phases over which the comparisons were made. This effect occurred despite the aforementioned extensive prior training (cf. Ploog & Williams, 2010).

The main effect of Half-Phase represents changes across successive reversals within a given condition (LTL within phases). Half-Phase was significant for both comparisons (Table 2). The remaining two main effects, Session and Quarter-Session, reflected learning within individual reversals and were significant for both comparisons (Table 2). However, because Session and Quarter-Session were of secondary theoretical interest, and were always significant, they will not be considered in further detail. (The Session×Quarter-Session interaction was also significant for both comparisons; see Table 2.)

Table 5 Summary of the ANOVA corresponding to Fig. 5

Source	*df*	*MS*	*F*	*p*	Partial η^2
Comparison of SigS + (mean of Phases 4 and 6) vs. Diff (Phase 5):					
Half-Phase	1, 6	9.43	33.88	< .005	0.85
Session	1, 6	22.96	175.17	< .001	0.97
Quarter-Session	3, 18	54.33	64.38	< .001	0.92
Condition × Session	1, 6	2.61	55.82	< .001	0.90
Half-Phase × Quarter-Session	3, 18	0.49	3.61	< .05	0.38
Session × Quarter-Session	3, 18	0.46	4.60	< .05	0.43
Half-Phase × Session × Quarter-Session	3, 18	0.25	4.16	< .05	0.41

Table 6 Summary of the ANOVA corresponding to Fig. 6

Source	*df*	*MS*	*F*	*p*	Partial η^2
Comparison of SigS+ (Phase 4) vs. Zero (Phase 4)—Fig. 6 (top):					
Condition (*Note*: *Between-subjects factor*)	1, 13	63.62	9.32	< .01	0.42
Half-Phase	1, 13	11.45	25.60	< .001	0.66
Session	1, 13	17.09	45.49	< .001	0.78
Quarter-Session	3, 39	71.02	110.55	< .001	0.90
Quarter-Session × Condition	3, 39	4.79	7.45	< .001	0.36
Half-Phase × Session × Condition	1, 13	2.36	7.55	< .05	0.37
Half-Phase × Quarter-Session	3, 39	0.65	3.94	< .05	0.23
Comparison of SigS+ (Phase 6) vs. Zero (Phase 4)—Fig. 6 (bottom):					
Condition (*Note*: *Within-subjects factor*)	1, 7	52.40	42.20	< .001	0.86
Half-Phase	1, 7	10.61	73.35	< .001	0.91
Session	1, 7	23.51	64.69	< .001	0.90
Quarter-Session	3, 21	80.36	163.43	< .001	0.96
Condition × Session	1, 7	1.58	7.34	< .05	0.51
Half-Phase × Session	1, 7	4.64	10.81	< .05	0.61
Condition × Quarter-Session	3, 21	4.67	15.64	< .001	0.69
Half-Phase × Quarter-Session	3, 21	0.60	6.13	< .005	0.47
Session × Quarter-Session	3, 21	0.97	6.11	< .005	0.47

Any interaction between Phase and Session and/or Quarter-Session implies that within-reversal learning is changing systematically as a function of continued training—that is, LTL across phases. For the comparisons of Phases 1 or 3 versus 6 and of Phases 4 versus 6, the Phase × Session interaction was significant (Table 2). From inspecting Fig. 2, this appears to be due to a greater difference between the early and late phases for Session 1 of new reversals than for Session 2, presumably due to the high level of learning on Session 1 during the later phases of training. The Phase × Quarter-Session interaction was significant for the comparison of Phases 1 or 3 versus 6 (Table 2), which reflects that the rate of within-reversal learning was greater in the later phase of training. The Phase × Quarter-Session interaction marginally failed to reach significance for the Phase 4 versus 6 comparison, $F(3, 18) = 2.88$, $p = .065$. The Phase × Session × Quarter-Session triple interaction was significant only for the Phase 4 versus 6 comparison (Table 2). Note that the initial level of performance during a session (first quarter) was slightly higher during Session 2 than during Session 1, but this was not differential for the early versus late phases. Note also that the superior performance in the later phase was characterized by faster within-reversal learning (across quarter-sessions), with this difference being greater on Session 1 than on Session 2.

The Half-Phase × Session interaction was significant for the comparison of Phases 1 or 3 versus 6 (Table 2), but not for the comparison of Phases 4 versus 6. Inspection of the marginal means revealed that improvement in performance from the first to the second half of a phase was more pronounced for Session 1 than for Session 2, possibly because of a ceiling effect caused by high performance at the end of Session 2. This does not seem to have any direct implications for LTL within phases.

Finally, an interaction between Half-Phase and Quarter-Session reflects changes in the pattern of within-reversal learning as a function of improvement in the rate of within-phase learning (LTL within phases). This interaction was significant only for the comparison of Phases 1 or 3 versus 6 (Table 2), but not for the Phase 4 versus 6 comparison.

Unsig and NDiff: Earlier and later conditions Figure 3 presents the results for two conditions that maintained low

Table 7 Summary of the ANOVA corresponding to Fig. 7

Source	*df*	*MS*	*F*	*p*	Partial η^2
Comparison of NDiff (mean of Phases 2+8) vs. Unsig (mean of Phases 1 or 3+9):					
Session	1, 14	11.59	51.07	< .001	0.79
Quarter-Session	3, 42	12.40	68.19	< .001	0.83
Half-Phase × Session	1, 14	0.57	9.61	< .01	0.41

Table 8 Summary of the ANOVA corresponding to Fig. 8

Source	*df*	*MS*	*F*	*p*	Partial η^2
Comparison of SigS– (Phase 7) vs. NDiff (Phase 8)—Fig. 8 (top):					
Session	1, 14	15.86	38.34	< .001	0.73
Quarter-Session	3, 42	26.52	72.45	< .001	0.84
Condition × Half-Phase	1, 14	1.15	5.78	< .05	0.29
Condition × Session	1, 14	1.36	5.48	< .05	0.28
Comparison of SigS– (Phase 7) vs. Unsig (Phase 9)—Fig. 8 (bottom):					
Condition	1, 14	10.64	13.98	< .005	0.50
Session	1, 14	10.43	20.98	< .001	0.60
Quarter-Session	3, 42	20.68	59.61	< .001	0.81
Condition × Session	1, 14	3.69	18.80	< .005	0.57
Condition × Quarter-Session	3, 42	1.75	11.37	< .001	0.45
Condition × Session × Quarter-Session	3, 42	0.16	3.13	< .05	0.18

discrimination performance levels, due to the absence of differential stimuli during their delay-of-reinforcement intervals. The top panel shows Phases 1 or 3 (earlier) versus 9 (later) under Condition Unsig; the bottom panel shows Phases 2 (earlier) versus 8 (later) under Condition NDiff.

For Condition Unsig (Fig. 3, top panel), performance was generally poor, with a small advantage for the later phase of training, even though the ANOVA failed to reach significance (p = .298). Also, the effect of Half-Phase was not significant (p = .91). Thus, we found no clear evidence for

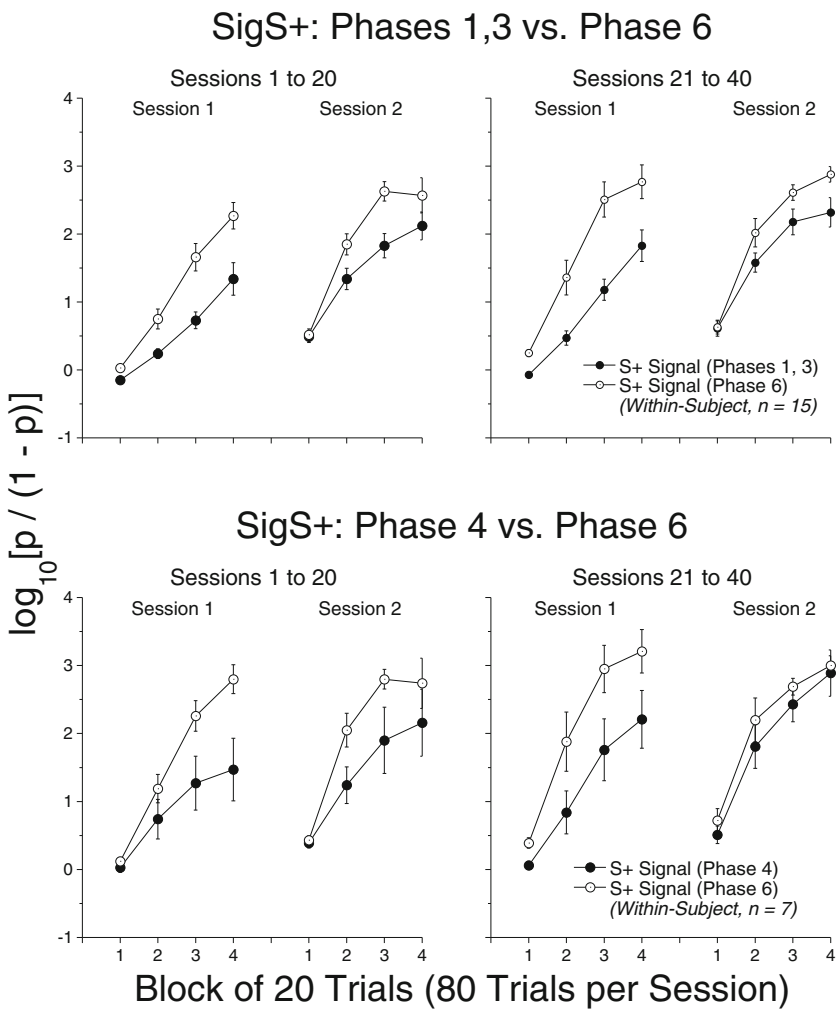

Fig. 2 Learning-to-learn effect across phases under Condition SigS+: Performance (logit p) as a function of block of 20 trials, Session 1 or 2 per reversal, and first or second half of the phase (Sessions 1–20 or 21–40). The top graphs show the within-subjects comparison (n = 15) of Condition SigS+ under Phase 1 or 3 (filled circles) versus Phase 6 (*open circles*). The bottom graphs show the within-subjects comparison (n = 7) of Condition SigS+ under Phase 4 (*filled circles*) and Phase 6 (*open circles*). Bars represent one standard error above and below the means

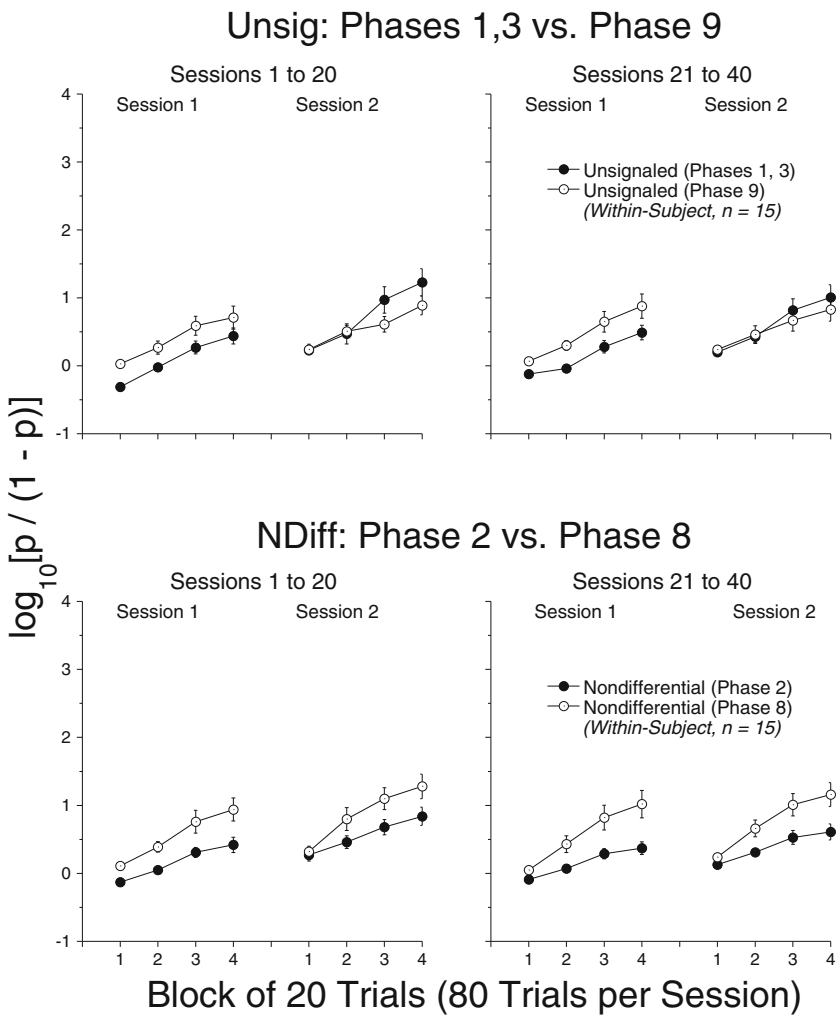

Fig. 3 Learning-to-learn effect across phases under Conditions Unsig (*top*) and NDiff (*bottom*): Performance (logit p) as a function of block of 20 trials, Session 1 or 2 per reversal, and first or second half of the phase (Sessions 1–20 or 21–40)

LTL across or within phases, which replicates our previous unpublished findings. The Phase × Session interaction was significant (Table 3), which reflects a smaller difference between sessions for the later phase of training, the reverse of an LTL effect. Phase × Quarter-Session was not significant (p = .462), nor was Phase × Half-Phase (p = .569). Thus, neither learning within reversals nor learning across reversals was increased by the additional training across phases. The triple Phase × Session × Quarter-Session interaction was significant (Table 3), apparently due to the respective performance for the two phases reversing between the first and second sessions. None of the other sources of variance yielded significant results, except the usual main effects of Session and Quarter-Session.

The effect of Phase was somewhat larger in the NDiff condition (Fig. 3, bottom panel) and was statistically significant (Table 3), indicating LTL across phases. However, Half-Phase was not significant (p = .128), thus providing no evidence for LTL within NDiff conditions.

The Phase × Quarter-Session interaction was significant (Table 3), indicating faster within-reversal learning in the later phase of training (LTL across phases). The Phase × Half-Phase interaction was not significant. Thus, even though within-reversal learning was faster later in training, learning did not improve across successive reversals. (Session and Quarter-Session were, again, significant.)

Summary of LTL effects The LTL effects can be summarized as follows: LTL across phases occurred for Conditions SigS+ and NDiff, but not for Unsig. (These three were the only conditions for which LTL across phases could be assessed, as these conditions were the only ones that were replicated in our design.) LTL within phases occurred for Conditions SigS+, Diff, and Zero, but not for Unsig, NDiff, and SigS–.

Effects of different stimulus conditions during the delay-of-reinforcement intervals

Condition SigS+ versus Condition Diff For Condition SigS +, signals occurred only following S+ choices, while for Condition Diff, one signal occurred after choice of the S+,

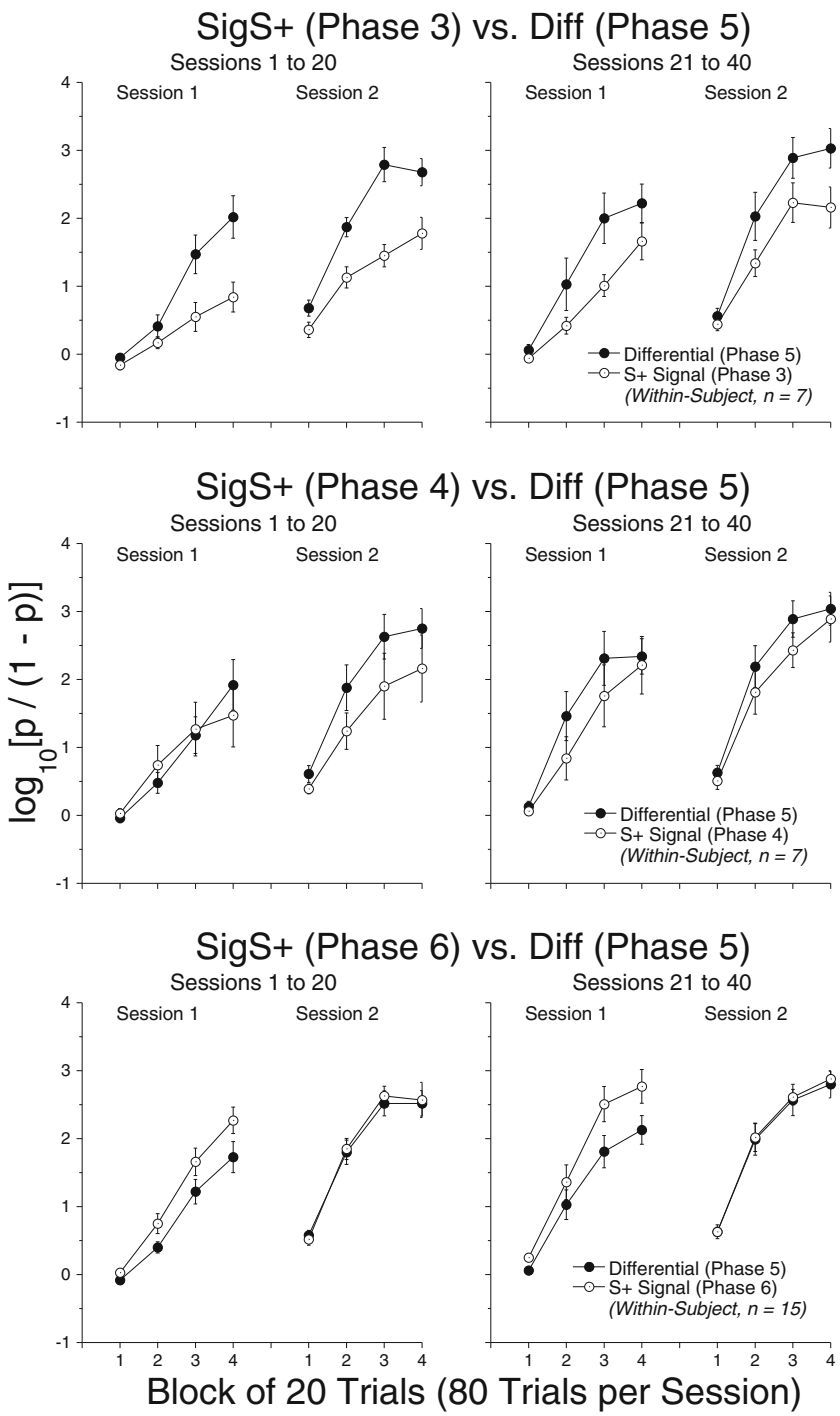

Fig. 4 Comparison of performance under Conditions SigS+ and Diff: Performance (logit p) as a function of block of 20 trials, Session 1 or 2 per reversal, and first or second half of the phase (Sessions 1–20 or 21–40). The top, middle, and bottom pairs of graphs show Phases 3, 4, and 6 (SigS+) versus Phase 5 (Diff), respectively

and a different stimulus (never paired with food) followed the choice of the S–. Although Condition Diff was presented only once (Phase 5 for all subjects), Condition SigS+ was presented multiple times (in Phases 3, 4, and 6), each of which could be compared with the one presentation of Condition Diff.

Figure 4 depicts these comparisons. The top panel shows the performance of Sets 2 and 4 (n = 7) under SigS+ in Phase 3, the middle panel shows the performance of Sets 3 and 4 (n=7) under SigS+ in Phase 4, and the bottom panel shows the performance of Sets 1 through 4 (n = 15) under SigS+ in Phase 6. Performance in Condition SigS+ (open circles) varied substantially over its three different presentations, but this variation was systematic, as presentations during later phases produced notably higher performance levels. Thus, any assessment of the effect of the additional S– signal in Condition Diff is confounded by the order of presentation of Condition SigS+. Because of this confound, the superior learning for Condition Diff shown in the top and middle panels of Fig. 4 was statistically significant, as

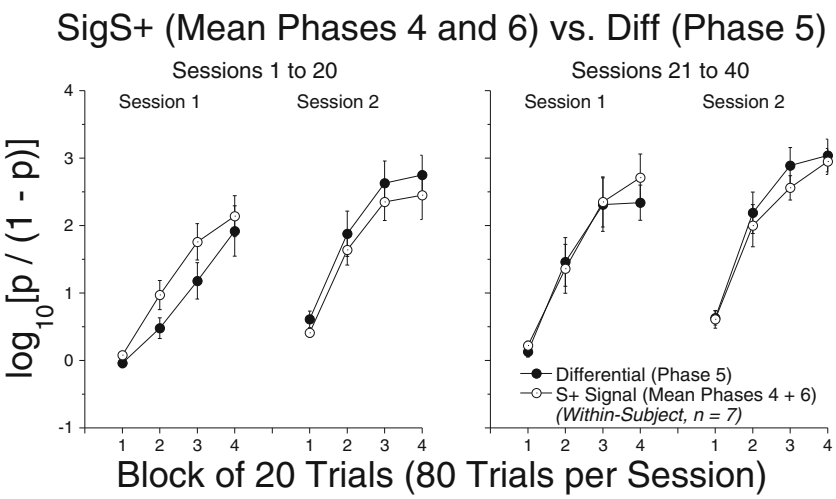

Fig. 5 Comparison of performance under Conditions SigS+ (mean of Phases 4 and 6) and Diff (Phase 5) for Sets 3 and 4—an ABA reversal design: Performance (logit p) as a function of block of 20 trials, Session 1 or 2 per reversal, and first or second half of the phase (Sessions 1–20 or 21–40)

was the superior learning for Condition SigS+ shown in the bottom panel (main effect of Condition in Table 4). As can be seen in Table 4, a number of effects were statistically significant, but these analyses are not discussed any further due to the order confound rendering them uninterpretable.

From Table 1, it is evident that Bird Sets 3 and 4 both had Condition SigS+ in Phases 4 and 6, while these same birds (n = 7) had Condition Diff during Phase 5. This allows for a comparison for these birds (Sets 3 and 4) according to an ABA reversal design. These data are shown in Fig. 5. The data

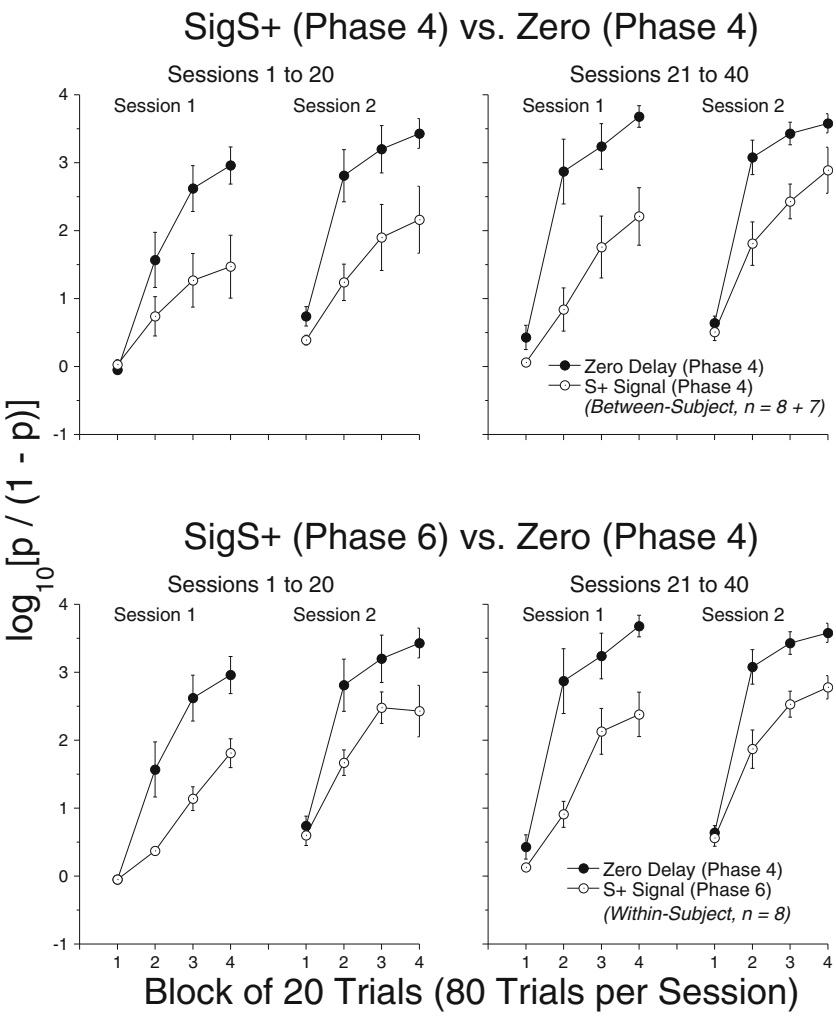

Fig. 6 Comparison of performance under Conditions SigS+ and Zero: Performance (logit p) as a function of block of 20 trials, Session 1 or 2 per reversal, and first or second half of the phase (Sessions 1–20 or 21–40). The top and bottom pairs of graphs show Phases 4 and 6, respectively (SigS+), versus Phase 4 (Zero). Note that the first comparison is a between-subjects comparison. The filled circles in the top and bottom graphs (Condition Zero) are identical and are only provided to facilitate comparisons

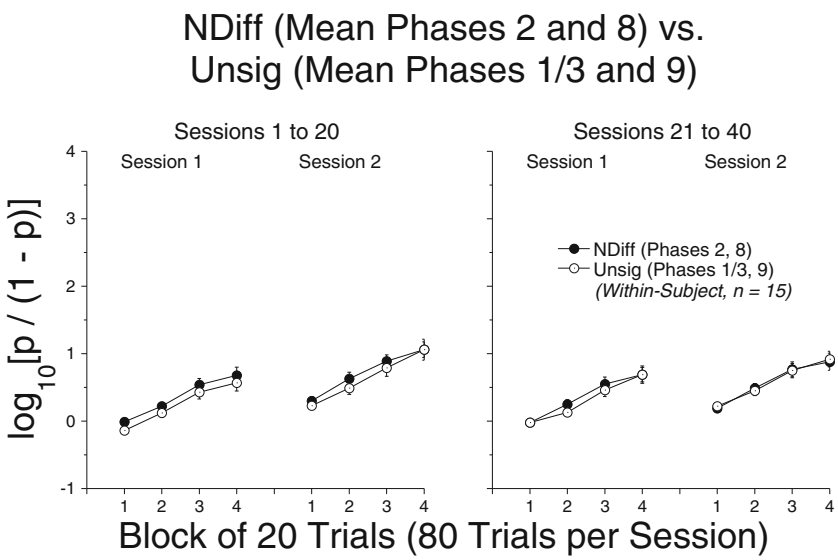

Fig. 7 Comparison of performance under Conditions NDiff (mean of Phases 2 and 8) and Unsig (mean of Phases 1 or 3 and 9): Performance (logit p) as a function of block of 20 trials, Session 1 or 2 per reversal, and first or second half of the phase (Sessions 1–20 or 21–40)

points are the means of Phases 4 and 6 (SigS+), to cancel out order effects and possible LTL across phases, with this average being compared with the average for Phase 5 (Diff). In an ANOVA we found that, although the effect of Half-Phase was significant (Table 5), the effect of Condition was not. The Condition × Half-Phase interaction was also not significant, nor was Condition × Quarter-Session. Condition × Session was significant (Table 5), which from Fig. 5 seems to be

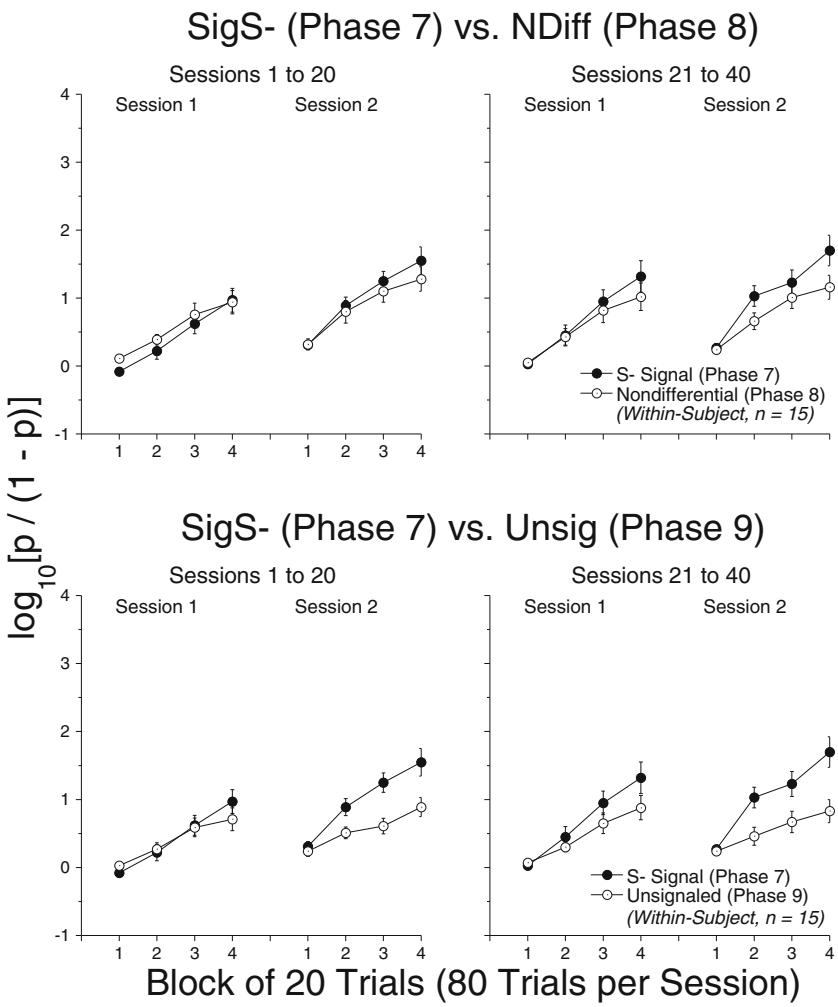

Fig. 8 Comparison of performance under Condition SigS– (Phase 7) versus Conditions NDiff (Phase 8) and Unsig (Phase 9): Performance (logit p) as a function of block of 20 trials, Session 1 or 2 per reversal, and first or second half of the phase (Sessions 1–20 or 21–40). The top pair of graphs shows Conditions SigS– (Phase 7) and NDiff (Phase 8); the bottom pair shows SigS– (Phase 7, same as in the top graphs, included for easy comparison) and Unsig (Phase 9)

due to Condition SigS+ producing better performance on Session 1, while Condition Diff produced better performance on Session 2, although in both cases the differences were very small. Half-Phase × Quarter-Session was significant, as was Session × Quarter-Session and the triple Half-Phase × Session × Quarter-Session interaction (Table 5), in addition to the usual main effects of Session and Quarter-Session (Table 5). In general, therefore, when the order confound was removed, there appeared to be no consistent difference between Condition SigS+ and Condition Diff. Note, however, that LTL within phases occurred regardless of the stimulus condition during the reinforcement delay, as the Condition × Half-Phase interaction was not significant.

Condition SigS+ versus Condition Zero Figure 6, top panel, depicts differences in performance under Conditions SigS+ (Phase 4, Sets 3 and 4) and Zero (Phase 4, Sets 1 and 2)—a between-subjects comparison—and under Conditions SigS+ (Phase 6, Sets 1 and 2) and Zero (Phase 4, Sets 1 and 2)—a within-subjects comparison. Performance was very consistent, in that Condition Zero (filled circles) generated higher performance than did Condition SigS+ (open circles), except in the first quarter of each session. The ANOVA, with Condition as a between-subjects factor (top pair of graphs), confirmed the significant differences between the conditions (Table 6). Two interactions involving condition—Condition × Quarter-Session and Condition × Half Phase × Session—were significant (Table 6). Furthermore, the effects involving Half-Phase (LTL within phases)—the Half-Phase main effect (LTL within phases) and Half-Phase × Quarter-Session interaction (performance across quarter-sessions improved more for the second than for the first half-phase)—were significant (Table 6), in addition to the usual Session and Quarter-Session effects, which were of secondary interest. The ANOVA, with Condition as a within-subjects factor (Fig. 6, bottom panel), produced significant differences for the Condition and Half-Phase effects and for the Condition × Session, Condition × Quarter-Session, Half-Phase × Session, and Half-Phase × Quarter-Session interactions, in addition to the usual Session, Quarter-Session, and Session × Quarter-Session effects. In summarizing these numerous findings, performance was higher under Condition Zero than under Condition SigS+, with improved performance over half-phases (LTL within phases), sessions, and quarter-sessions.

Condition NDiff versus Condition Unsig Figure 7 depicts the birds' performances under Conditions NDiff and Unsig. Because each bird was exposed to two NDiff and two Unsig phases, Phases 2 and 8 (NDiff, filled circles) were averaged, as well as Phases 1 or 3 and 9 (Unsig, open circles). Note that Sets 1 and 3 were exposed to Unsig in Phase 3, whereas Sets 2 and 4 were exposed to Unsig in Phase 1. Note also that Pigeon BL 90 died after Phase 4, and thus the analysis is based on $n=15$ only. We found no difference in performance under Conditions NDiff and Unsig (the filled and open circles overlap). Other than the usual significant effects of Session and Quarter-Session, the ANOVA revealed that only the Half-Phase × Session interaction was significant (Table 7), which upon closer inspection of the marginal means indicated that the improvement from Session 1 to Session 2 was greater for the first than for the second half-phase.

Conditions NDiff and Unsig versus Condition SigS– Figure 8 depicts the birds' performance under Conditions SigS– versus NDiff and under SigS– versus Unsig. For the top panels, comparing Conditions SigS– (Phase 7) and NDiff (Phase 8), the main effect of Condition was not significant, nor was Half-Phase. However, the Condition × Half-Phase interaction was significant, due to a bigger advantage for the SigS– condition in the second half-phase than in the first (Table 8). This was confirmed by a post-hoc ANOVA for each separate level of Half-Phase. For the first half-phase, no source of variance was significant, except the usual effects of session and Quarter-Session. Specifically, Condition was not significant. For the second half-phase, Condition also failed to reach significance, but the Condition × Quarter-Session interaction was significant, $F(3, 42) = 3.62$, $p = .021$ (sufficient for Bonferroni correction), reflecting faster within-reversal learning for Condition SigS– during the second half-phase.

Because the omnibus ANOVA (with Half-Phase included) indicated that the Condition × Session interaction was also significant (Table 8), a second post-hoc ANOVA was conducted for each separate session. Condition was not significant for Session 1 or 2, but the Condition × Quarter-Session interaction for Session 2 was significant, $F(3, 42) = 2.93$, $p = .044$ (even though it failed the test with Bonferroni correction), indicating faster within-reversal learning for Condition SigS–.

For the bottom panels, comparing Conditions SigS– (Phase 7) versus Unsig (Phase 9), the effect of Condition was significant (Table 8), showing faster learning for SigS–, whereas Half-Phase was not significant. Condition × Session was significant, with this interaction again being due to a bigger difference in favor of Condition SigS– on the second session of the reversal. Condition × Quarter-Session was also significant (Table 8), again with greater within-reversal learning in Condition SigS– than in Condition Unsig. Furthermore, Condition × Session × Quarter-Session was significant, indicating that the difference in within-reversal learning in favor of Condition SigS– was greater on Session 2 than on Session 1. (Session and Quarter-Session were significant as usual.) In summary, we found evidence that Condition SigS– produced higher performance than did Conditions NDiff and Unsig, although the differences were primarily on the second session of training in a given reversal.

Discussion

Learning to learn

Reversal learning continued to occur even after hundreds of prior reversals, as was also reported by Ploog and Williams (2010). Such improvement across repeated exposures to the same condition—that is, LTL across phases—was evident here for the SigS+ and NDiff conditions, although the effect was much larger in the former condition. However, no significant improvement across training phases was evident for the Unsig condition, although a nonsignificant trend in that direction did appear (see Fig. 3). Conditions SigS+, Diff, and Zero also produced a significant improvement across reversals within a phase—that is, LTL within phases, the more conventional measure of LTL. However, no such effect was apparent for the NDiff, Unsig, or SigS– conditions (Fig. 1). This dissociation suggests that two different types of LTL exist: one that transfers across phases (Conditions SigS+ and NDiff), and one that is specific to a given set of discrimination contingencies and signal properties (the within-phase effect and its interactions with Quarter-Session and Session). Previous interpretations of similar improvements in learning (e.g., Harlow, 1959) have attributed the improvement to the reduction of general "error factors," such as position habits, stimulus preferences, response perseveration, and response alternation. We did not attempt an analysis of the types of errors in the present results, because errors necessarily decrease whenever discrimination accuracy improves, making any causal analysis problematic, especially given that multiple potential error factors that are statistically interdependent are involved.

The improvement within phases, seen in the SigS+, Diff, and Zero conditions, seemingly requires a different or additional explanation. One possibility is that the birds learned to use the outcome of the preceding trial as a discriminative cue, as proposed by Williams (1976) and supported by the effects of ITI on reversal-learning performance (Williams, 1971b, 1976; Ploog & Williams, 2010).

One reason for the continued interest in LTL is that it is at least partially independent of simple discrimination learning. In all conditions, reversals were learned to conventionally high levels (note that a logit p value of 1.0 is equivalent to 90 % correct). Despite that degree of learning, neither the NDiff nor the Unsig condition produced any improvement in learning rates across the 20 reversals in each training phase. The issue posed is why an even higher level of discrimination learning was necessary for LTL to emerge across the 20 reversals presented within a phase. Given the delay-of-reinforcement contingency used here, the conditioned reinforcer during the delay interval was necessary for this second type of LTL, although presumably any of a variety of manipulations might similarly allow LTL to occur if they also caused discrimination learning within a reversal to increase to very high levels.

Effects of stimuli during the delay of reinforcement

The effects of the stimulus conditions during the delay-of-reinforcement intervals can be viewed at two levels. At a gross level, we observed an order-of-magnitude difference in the rates of learning when a differential stimulus was contingent on the correct choice, relative to conditions Unsig, NDiff, and SigS–, thus indicating that the conditioned-reinforcement properties of the delay-interval stimulus was the dominant variable. However, at a more fine-grained level, additional effects of the different stimulus conditions merited attention. Comparing the SigS+ and Diff conditions, the additional signal (a potential "sign post") contingent on S– choices failed to produce a significant main effect, although a significant Condition × Session interaction did emerge. This interaction is difficult to interpret because of the lack of an obvious rationale: Condition SigS+ produced better performance on Session 1, while Condition Diff produced better performance on Session 2—a reversal in direction—that cannot be aligned with the simple effect of an added sign post (S–). In terms of a conditioned-reinforcement account, adding S– was not expected to produce a consistent effect, because S– is not a conditioned reinforcer.

More provocative were the differences among the SigS–, Unsig, and NDiff conditions. While a direct comparison of Conditions Unsig versus NDiff yielded no significant differences, the separate comparisons of each with Condition SigS– did produce some possibly important differences. For the comparison of Conditions SigS– versus NDiff, the main effect of Condition was not significant, but the interaction between Condition and Half-Phase was significant, indicating relatively better learning for Condition SigS– during the second half of the training phase. Also significant was the interaction between Condition and Quarter-Session (indicating a higher rate of within-reversal learning), but again only during the second half of the training phase, primarily during the second session of training. A similar advantage of SigS– over NDiff was apparent for Session 2, but not for Session 1, of the reversals. For the comparison of SigS– versus Unsig, the main effect of Condition was significant, showing faster learning with the SigS– condition, which also occurred for the rate of learning within reversals. These differences were greatest in the second half of the training phase, and also during the second session of reversal learning, suggesting that the superiority of the SigS– condition gradually developed over the series of repeated reversals. Overall, this implied that a sign post (S–) may have enhanced discrimination learning, but only when contrasted with Condition Unsig, and not with Condition NDiff. Why a sign post should have this differential effect is not clear. A conditioned-reinforcement account would not predict any consistent effect under these conditions. Because of the multiple comparisons involved and the small effect sizes

that were evident, the possibility that these differences were Type I errors cannot be dismissed.

It should also be noted that the just-described differences were confounded by order of presentation, as Condition SigS– occurred before both the NDiff and Unsig conditions to which it was compared. However, this confound is biased against Condition SigS–, assuming that the rate of learning was continuously improving with continued training. Although the superiority of Condition SigS– over Condition Unsig may be theoretically important, its small size in comparison to the other stimulus effects suggests that it is less important in terms of determining the rate of discrimination learning.

With one exception, the traditional concept of conditioned reinforcement seems to encompass all of the results, as the essential feature of highly proficient reversal learning was the presentation of a stimulus associated with food immediately contingent on the correct choice response. Adding a second, different stimulus that was contingent on an incorrect choice did not improve performance, while adding the S+ signal following an S– response (the NDiff condition) greatly retarded learning.

Sign posts versus conditioned reinforcement

In the introduction, we noted that the traditional concept of conditioned reinforcement has been challenged by the idea that the stimulus serves as a "sign post" mediating the delay, which implies that the concept of "conditioned value" need not be invoked. Because the empirical ramifications of the sign-post idea have not been explicated, it is difficult to ascertain how it would be applied to the present data set. One possible interpretation is that sign posts serve as sources of information much like that invoked by the "information hypothesis," a viewpoint that has been heavily investigated in previous research, and generally rejected (cf. Fantino, 1977; see, e.g., Case, Ploog & Fantino, 1990). However, that rejection was based on whether S– stimuli were reinforcers, not whether they increased the rate of discrimination learning. The distinction is important because a purely informational role would, in principle, not require the stimulus to support behavior similarly to conventional reinforcers.

At the risk of distorting the views of others, we suggest that the essence of the sign-post notion is forward-looking, informing the animal of where it is on the path to the eventual reinforcement. In contrast, the notion of conditioned reinforcement is backward-acting, strengthening the behavior that preceded it. Given this distinction, one means to dissociate the two ideas would be to vary the response contingency for producing the stimulus. According to the sign-post idea, the response contingency would seem to be irrelevant, as the information about the future is the same for a conditioned reinforcer separated from the response by a 2-s delay as for a conditioned reinforcer that is immediately contingent on the response. Thus, the delay contingency should not interfere with the rate of learning, a prediction contradicted by previous results (Ploog & Williams, 2010; Williams & Dunn, 1994).

An alternative conception of a sign post is that it does have reinforcing properties, due to its providing forward-looking information about the eventual reinforcer, so that its ability to facilitate learning depends on its being immediately contingent on the response, essentially like the traditional concept of conditioned reinforcement. If so, it is not clear how the concept of a sign post can be empirically dissociated from the traditional concept, in which case the issue of interpretation is moot.

Other evidence pertaining to the sign-post concept

The primary evidence adduced to support the "sign-post hypothesis" has been demonstrations that conditioned reinforcers have discriminative properties that are similar to those of discriminative stimuli that are unpaired with the primary reinforcer. Davison and Baum (2006) provided one example in a study in which pigeons' responses produced the usual food plus magazine presentations, along with magazine presentations without food. The schedules of the two types of magazine presentations were manipulated independently, from a perfect positive correlation, to a zero correlation, to a perfect negative correlation. The dependent variables were the size and duration of the "preference pulse," defined as a local increase in preference for the alternative that has just produced a given outcome. When a positive correlation between the two types of magazine presentations was in effect, the preference pulses after food presentations and after magazine-only presentations, while they were substantially greater for food presentations, were similar in form: an immediate elevation, followed by a gradual decline over successive responses. However, when a negative correlation was in effect, there was a small elevation in preference after a magazine-only presentation, followed by a small shift in preference toward the alternative response—that is, preference opposed to the supposed reinforcing effects of the magazine-only presentations. Results with a zero correlation were intermediate between the extremes. Davison and Baum interpreted these results as showing that the magazine-only presentations were serving as discriminative cues for which of the response alternatives had a higher rate of reinforcement.

In a second experiment, Davison and Baum (2006) added a neutral keylight that never was paired with food, and presented it with schedules similar to the magazine-only presentations described above. The preference pulses for the magazine-alone and keylight presentations were generally similar, although the preference pulses for magazine-only presentations were slightly higher throughout, an effect possibly due to the magazine-only presentations having been paired with food.

While the results of Davison and Baum (2006) strongly support the idea that putative conditioned reinforcers can have substantial discriminative effects, they do not exclude the interpretation that they have reinforcing effects as well. There is no question that stimuli can serve multiple functions and that, depending on the contingencies, one function or another may dominate.

Davison and Baum (2006) also cited the literature on second-order conditioning as further evidence that pairing of the putative conditioned reinforcer with food is unnecessary for producing its behavioral effects. Such effects have been ascribed to a purely discriminative function (Marr, 1979). However, Rose and Fantino (1978) challenged the generality of that interpretation, as they demonstrated substantial differences in the effects of brief stimuli that were paired versus unpaired with food.

Although such domination by discriminative effects has occurred in some cases, this seems unlikely in simultaneous discrimination procedures like the one that was used here. The occurrence of the conditioned reinforcer during the delay-of-reinforcement interval was perfectly correlated with food at the end of the interval, so that any information provided by the conditioned reinforcer was redundant with that of the food itself, and the food was closer to the onset of the next trial. Moreover, the discrete trial procedures used for many simultaneous discriminations typically are separated by intertrial intervals, so that any discriminative effect would need to transcend the temporal interval between trials, in much the same way that delayed matching-to-sample requires memory over the retention interval, a task at which pigeons are notably deficient.

Perhaps the most serious problem with the sign-post concept is exemplified by further discussion of the results of Williams and Dunn (1991a), already described in the introduction. They provided strong evidence that any benefit of putative conditioned reinforcers required pairing of the stimulus with food. In their first experiment, a 50 % food reinforcement schedule was used for correct responses. In the baseline condition, correct but unreinforced choices had the same feedback as incorrect responses—that is, a return to the ITI. In the conditioned-reinforcement condition, a tone was presented after the correct but nonreinforced choices. When the tone accompanied the food on correct reinforced trials, learning was substantially enhanced relative to the baseline. However, in a second condition, the tone occurred only on correct trials that were nonreinforced, and hence was never paired with food. Here, no facilitation of learning was found, relative to the baseline with no tone. It is important to appreciate that the discriminative properties of the tone were the same in both cases. Hence, the critical difference in the results—whether or not tone-enhanced learning occurred—was dependent on the tone being paired with food.

Likewise, in the present study, the condition in which the differential signal appeared only after an incorrect response produced learning far inferior to that when the differential signal appeared only after a correct response, even though, logically, the two types of signals were equivalent in providing a cue about which choice would be correct on the next trial.

The results of Williams and Dunn (1991a) are consistent with many other data showing that discriminative effects are insufficient to explain putative conditioned-reinforcement effects, as such effects have occurred only when pairing with the primary reinforcer has occurred (Cronin, 1980; McDevitt & Williams, 2010; Williams & Dunn, 1991b).

As noted, one feature of the results, the superiority of Condition SigS– over Condition Unsig (and, to a lesser extent, over Condition NDiff) is problematic for the traditional concept of conditioned reinforcement. The facilitation produced by the S– signal could be interpreted as the S– signaling that the incorrect stimulus should be avoided. But, an alternative explanation would be that the S– signal acquired conditioned inhibitory properties, and thus punished incorrect responses (given that stimuli that are inhibitory with respect to food acquire aversive properties). Such an account would be consistent with the finding that the advantage of the S– signal primarily emerged during the last set of ten reversals, and was most apparent during the second session of learning for individual reversals, assuming that conditioned inhibition takes more time to develop than does conditioned excitation.

The benefit of the S– signal observed in the present experiment is conceptually consistent with the original "information hypothesis" of conditioned reinforcement, where the critical empirical issue is whether a negative discriminative stimulus can maintain responding. The present results suggest that such an effect is possible. However, it should be appreciated that the magnitude of this effect is trivial in comparison to the conditioned-reinforcement effects of a response-contingent S+ signal. The information idea is also not compelled by the results, given the alternative conditioned-inhibitor interpretation proposed above.

Author note This research was supported, in part, by a PSC-CUNY grant to the first author and by funds for animal maintenance provided by the College of Staten Island Divisional Dean and by the Department of Psychology.

References

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.

Cronin, P. B. (1980). Reinstatement of postresponse stimuli prior to reward in delayed-reward discrimination learning by pigeons. *Animal Learning & Behavior, 8,* 352–358. doi:10.3758/BF03199616

Davison, M., & Baum, W. M. (2006). Do conditional reinforcers count? *Journal of the Experimental Analysis of Behavior, 86,* 269–283.

Fantino, E. (1977). Conditioned reinforcement: Choice and information. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 313–339). Englewood Cliffs, NJ: Prentice Hall.

Grice, G. R. (1948). The relation of secondary reinforcement to delayed reward in visual discrimination learning. *Journal of Experimental Psychology, 38,* 1–16. doi:10.1037/h0061016

Harlow, H. F. (1959). Learning set and factor theory. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 2, pp. 492–537). New York, NY: McGraw-Hill.

Marr, M. J. (1979). Second-order schedules and the generation of unitary response sequences. In M. J. Zeiler & P. Harzem (Eds.), *Advances in the analysis of behavior: Vol. 1. Reinforcement and the Organization of Behavior* (pp. 223–260). Chichester, U.K.: Wiley.

McDevitt, M. A., & Williams, B. A. (2010). Dual effects on choice of conditioned reinforcement frequency and conditioned reinforcement value. *Journal of the Experimental Analysis of Behavior, 93,* 147–155.

Ploog, B. O., & Williams, B. A. (2010). Serial discrimination reversal learning in pigeons as a function of intertrial interval and delay of reinforcement. *Learning & Behavior, 38,* 96–102. doi:10.3758/LB.38.1.96

Rose, J. E., & Fantino, E. (1978). Conditioned reinforcement and discrimination in second-order schedules. *Journal of the Experimental Analysis of Behavior, 29,* 393–418.

Shahan, T. A. (2010). Conditioned reinforcement and response strength. *Journal of the Experimental Analysis of Behavior, 93,* 269–289.

Sutherland, N. S., & Mackintosh, N. J. (1971). *Mechanisms of animal discrimination learning*. New York, NY: Academic Press.

Williams, B. A. (1971a). Color alternation learning in the pigeon under fixed-ratio schedules of reinforcement. *Journal of the Experimental Anlaysis of Behavior, 15,* 129–140.

Williams, B. A. (1971b). The effects of intertrial interval on discrimination reversal learning in the pigeon. *Psychonomic Science, 23,* 241–243.

Williams, B. A. (1976). Short-term retention of response outcome as a determinant of serial reversal learning. *Learning and Motivation, 7,* 418–430. doi:10.1016/0023-9690(76)90047-3

Williams, B. A. (1994). Conditioned reinforcement: Neglected or outmoded explanatory construct? *Psychonomic Bulletin & Review, 1,* 457–475. doi:10.3758/BF03210950

Williams, B. A. (1998). Relative time and delay of reinforcement. *Learning and Motivation, 29,* 236–248. doi:10.1006/lmot.1997.0999

Williams, B. A., & Dunn, R. (1991a). Preference for conditioned reinforcement. *Journal of the Experimental Analysis of Behavior, 55,* 37–46.

Williams, B. A., & Dunn, R. (1991b). Substitutability between conditioned and primary reinforcers in discrimination acquisition. *Journal of the Experimental Analysis of Behavior, 55,* 21–35.

Williams, B. A., & Dunn, R. (1994). Context specificity of conditioned reinforcement effects on discrimination acquisition. *Journal of the Experimental Analysis of Behavior, 62,* 157–167.