AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology:
Animal Learning and Cognition

ISSN: 2329-8456

2024, Vol. 50, No. 3, 210–224
https://doi.org/10.1037/xan0000382

Human Causality Detection and Judgment With Unsignaled and Signaled Delayed Outcomes

Phil Reed
Department of Psychology, Swansea University

Four experiments examined human ratings of causal effectiveness, and ability to detect causal relationships, in a nonverbal paradigm. Participants responded on a concurrent random interval, extinction schedule. In the presence of one stimulus, responses produced an outcome (triangle flash); in the presence of the other stimulus, they did not. Following making a judgment of causal effectiveness, two further stimuli were presented simultaneously with one another, and participants had to select one depending on which of the previous two stimuli were associated with effective responses. In all experiments, immediate outcomes were associated with higher causal ratings and better causal detection than outcomes delayed by 3 s. A signal inserted between response and outcome improved ratings and detection (Experiments 2 and 4), even when it was contiguous with the response but not the outcome (Experiments 2 and 3). Stimuli associated with both components (marking cues) did not impact judgments or detection (Experiment 3). Stimuli signaling the availability of an outcome if a response was made (signaled reinforcement) did not improve causal judgments, but did improve detection of stimuli associated with the outcome (Experiment 4). Responses during the delay interfered with detection of the actual relationship when delays were unsignaled (Experiments 1–4), but not with fully or briefly signaled delays (Experiments 2–4), or with signaled reinforcement (Experiment 4). The results suggest a delay stimulus serves to signal the response has been successful and demark the delay period by serving a discriminative function. These findings mirror those seen in nonhuman conditioning.

Keywords: human causal judgment, causal detection, concurrent schedule, matching-to-sample

Perceptions and ratings of causal relationships diminish as the temporal delay between a response and its outcome increases (Greville & Buehner, 2010; Michotte, 1946; Reed, 1992a, 1992b; van Elk et al., 2014; Young & Nguyen, 2009), and the presentation of a stimulus during the delay period can alleviate these effects (Greville & Buehner, 2010; Reed, 1999; Young & Falmier, 2008). Unfortunately, few theories have been fully developed to account for the effects of a stimulus signaling an outcome delay in these paradigms (Buehner & May, 2003; Reed, 1999). Similar effects are found in instrumental conditioning procedures when a reinforcer is presented following a delay; in that response rates decreases (Lattal, 1984; Renner, 1964; Tarpy & Sawabini, 1974), and a stimulus presented during the delay alleviates this attenuation (Lattal, 1984; Richards & Marcattilio, 1978; Schaal & Branch, 1988). In contrast to the case of causal judgment, a number of accounts of the function of the delay stimulus have been developed for instrumental learning, including conditioned reinforcement (Kelleher & Gollub, 1962; Spence, 1947), the marking hypothesis (Lieberman et al., 1979), perceptual bridging (Rescorla, 1982), and discriminative/informational views (Ferster, 1953; Keller & Schoenfeld, 1950; see also Buehner & May, 2003; Young, 1995).

Conditioned reinforcement accounts of the effects of a signal presented during a reinforcement delay in instrumental conditioning have been widely considered and long been accepted as having great explanatory power (Dinsmoor, 1983; Royalty et al., 1987; Williams, 1991). However, the applicability of the concept across all paradigms is questionable, and a straightforward application of this view to human causality judgment procedures, where the outcome is biologically neutral, and often personally meaningless to the participant, is far from obvious (Buehner & May, 2003; Reed, 1992a, 1992b). The notion that such delay signals serve as reinforcers is also not always accepted (Rachlin, 1976; Staddon, 1983/2016). As a consequence, alternative theories have been proposed to accommodate the action of a signal presented during a reinforcement delay (Kaplan & Hearst, 1982; Williams, 1994), and these views may have more face validity when considering the role of a delay stimulus in human judgment of causality paradigms (Buehner & May, 2003; Reed, 1999; Young, 1995).

One view that has received empirical support in conditioning procedures is the "marking hypothesis" (Lieberman et al., 1979, 1985). In studies of marking, a stimulus is presented contiguously with a response initiating the delay, but not with the outcome. Schaal and Branch (1988) noted that unsignaled reinforcement delays decreased

Andrew R. Delamater served as action editor.
Phil Reed https://orcid.org/0000-0002-8157-0747

Phil Reed served as lead for conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, visualization, writing–original draft, and writing–review and editing.
Correspondence concerning this article should be addressed to Phil Reed, Department of Psychology, Swansea University, Singleton Park, Swansea SA2 8PP, United Kingdom. Email: p.reed@swansea.ac.uk

responding relative to immediate reinforcement, but a brief response-contingent stimulus presented during the initial portion of the delay increased rates to near baseline levels (see also Lieberman et al., 1985; Reed, 1992a, 1992b). As the same effect is noted when brief stimuli are presented contingent on both correct and incorrect responses in a discrimination paradigm (Lieberman et al., 1985; Reed, 1992a, 1992b), it suggests conditioned-reinforcement effects are not the sole mechanism at play. Rather, the brief stimulus is taken to make the preceding response more salient in memory, allowing its representation to be retrieved more easily on the delivery of an outcome (Lieberman et al., 1979, 1985). Such an interpretation has been given to a variety of effects in instrumental conditioning (e.g., Reed, 1989; Reed & Hall, 1989), and similar effects have been noted in human judgments of effectiveness when there is a response–outcome delay (Reed, 1996, 1999).

Another similar view has been termed "perceptual catalysis" and suggests a delay stimulus facilitates target–outcome associations by perceptually bridging the temporal gap and making the two events "appear to go together" (Rescorla, 1982, p. 140). Several within-subject experiments have shown that when one stimulus is followed by a signaled delayed outcome, and the other by an unsignaled outcome and, separately, by the signal, conditioning is improved in the former relative to the latter condition (Rescorla, 1982). According to this view when the delay is entirely filled the target and outcome become linked together perceptually. In support of this, Rescorla (1982) noted that when a delay was not entirely filled, the facilitatory effect of the stimulus was reduced. This view is very similar to some accounts presented to explain the effect of a stimulus filling a response–outcome delay in human causal judgments (Young & Nguyen, 2009).

That facilitating effects of delay signals are not always reducible to conditioned reinforcement may be important in the context of human judgments of causality (Buehner & May, 2003; Young, 1995). Pavlov (1927) highlighted the potential importance of a "second signaling system" mediating learning with a delayed outcome, and Keller and Schoenfeld (1950) argued the development of conditioned reinforcement is dependent on the discriminative function of a stimulus (see also Pliskoff et al., 1964; Wyckoff, 1952). Although such views do not accommodate all effects seen in all instrumental conditioning paradigms with signaled delays (Dinsmoor, 1983; Fantino, 1977; Williams, 1991), they may be important in discrimination procedures, and in studies of human decision making, such as are involved in causal judgments, where informational cues are of prime significance to the judgment and detection of causal relationships (Buehner & May, 2003; Young, 1995).

One such discriminative/informational function of a delay stimulus was highlighted by Ferster (1953), who interpreted a reinforcement delay as a tandem or chained schedule; responses in the initial link (the operative programmed schedule) being maintained by entry into the terminal link (the delay). For example, a random-interval (RI) 60-s schedule with a 3-s delay would be a tandem RI 60 s, fixed-time (FT) 3-s schedule if the delay were not signaled; but a chain RI 60 s, FT 3 s if the delay were signaled. It has been well established that behavior other than the target can be fortuitously reinforced at the end of a delay period, and can either interfere with response rates (Sizemore & Lattal, 1978), or produce uncertainty or shifts in attributions (e.g., was it a response to the keyboard or something else?) regarding the cause of an outcome (Buehner, 2005). This would be the case for the tandem schedule, where the absence of a change in the stimulus makes it unclear that one schedule (i.e., RI 60 s) has been satisfied, and the other (i.e., FT 3 s) has begun. Any learning accruing in the FT part of the schedule will generalize across the entire tandem schedule. Ferster (1953) suggested that in the chain schedule, the delay signal serves a discriminative function; that is, it signals the change in contingency between RI 60 s (which has been satisfied), and the FT 3 s, in a way that does not happen in the tandem schedule. The presence of a discriminative stimulus demarking both elements of the schedule (RI and FT) would restrict the effect of fortuitously reinforced (competing behaviors) during the FT 3 s (delay) component to that component (see also Thomas et al., 1989, for a similar interpretation). Although involving verbal information given to participants, Buehner and May (2003) demonstrated the detrimental effects of an outcome delay can be reduced if information about the potential delay is provided. It may be that the discriminative or informational functions of a delay stimulus are critical in human ratings of causal effectiveness.

In addition to ratings of causality, several researchers have noted the role of delayed outcomes in causality detection regarding reinforcement; accuracy of detecting the response–reinforcer dependency is higher the greater the temporal contiguity of the reinforcer with a response (Keely, 1999; Killeen, 1978; Kuroda, 2012; Lattal, 1975). A conditional discrimination task (the "causality-detection" procedure) has examined nonhuman detection of response–reinforcer dependencies (Killeen, 1978; Lattal, 1975). In this causality-detection task, a trial comprises sample and choice components. Completion of the schedule in the sample component is followed by the choice component. In the choice component, subjects choose one of two responses, and the only reliable cue for a correct choice response is the response–reinforcer relation arranged in the sample component.

Killeen (1978) arranged a random ratio 20 schedule for pigeons to operate on one central key during the sample component. As pigeons pecked, computer-generated "pseudopecks" occurred at the same rate as real key pecks, and the sample component could end either in a response-dependent or response-independent manner. In the choice phase, two keys were presented; one was correct if the sample phase ended response-dependently, and the other key was correct if the sample phase had ended response-independently. Pigeons correctly discriminated the relation between responding and the termination of the sample component, but only when the time between the last key peck in the sample component and the onset of the choice component was brief. Keely (1999) replicated and extended these results by presenting pigeons with a concurrent variable interval (VI) 120-s VI 120-s schedule in the sample component. The schedule could be completed either by a response to either element (left or right) of the concurrent schedule. In the choice component, one response was correct if a left-side response ended the sample component, the other was correct if a right-side response ended the sample component. On some trials, a variable-time (VT) 240-s schedule was added to the sample component, meaning the sample could end independently of responding. On the trials where the choice component was produced by the VT schedule, pigeons tended to peck the choice key associated with the last side-key pecked.

Kuroda (2012) investigated the impact of unsignaled and signaled delays in the causality-detection task developed by Keely (1999). In a sample component, a concurrent VI extinction (Ext) schedule operated, with assignment of the VI to the left or right element being random. The key peck that ended the VI schedule could immediately produce the choice component of the trial or could initiate a delay that was unsignaled, fully signaled, or have a stimulus present during the first second of the delay. Responses in the subsequent choice component were deemed correct depending on whether responses to the left or right elements in the sample had terminated that component; one alternative was correct if the sample response had been a left-key response, and the other alternative was correct if the sample had been a right-key response. Detection accuracy was high with a full delay signal, slightly lower with a partial delay signal, and lowest without a signal. Thus, responses producing delayed reinforcers could be detected, but only when a stimulus change accompanied the response. Such effects of an outcome delay have not been investigated in human causality judgment procedures while, additionally, using a causal detection paradigm.

The current series of experiments examined the potential mechanisms underlying both causal ratings and causal detection in humans, using a novel paradigm incorporating features from both types of study. Experiments 1 to 3 explored whether stimuli presented during an outcome delay acted similarly on causal ratings and causal detection. They also manipulated the relationship between the response, delay signal, and outcome, to explore various predictions from the marking, perceptual catalysis, and discriminative function views. In particular, they assessed whether brief stimuli, not filling the delay period, would enhance ratings and detection as effectively as stimuli fully filling the delay. This is predicted by the conditioned reinforcement, marking, and discriminative function views, but not necessarily by perceptual catalysis, which would expect a weakened effect (Experiments 2 and 3). Also, they tested whether stimuli presented following both successful and unsuccessful responses would enhance ratings and detection as predicted by marking and perceptual catalysis, but not conditioned reinforcement or discriminative function views (Experiments 3 and 4).

The conjoint use of a causality-detection procedure would allow further investigation of the role of a delay stimulus in determining detection of causal relationships. The conditioned reinforcement, marking, and perceptual catalysis views would all predict similar impacts of delay stimuli on causal judgments and causal detection. The discriminative function also typically suggests that there would be similar impacts, albeit through different mechanisms. The current Experiments 1 to 3 examined this prediction. However, there is one circumstance in which a signaling manipulation may be expected to improve causal detection, but to reduce causal ratings, and this effect is predicted by the discriminative function view but not the other views. When the availability of reinforcement is signaled (as opposed to the response that produced the reinforcement, as with a signaled delay), then most views would suggest that ratings and detection of causal judgments may be expected to be reduced by the reinforcement-availability signaling stimulus overshadowing the response through better predictive validity. The discriminative function view does not rule out this effect for causal judgments. However, such a stimulus may enhance causal detection, as it demarks the effective part of the trial (i.e., the interval schedule) from the outcome-delay period that follows it satisfaction. This may allow the response–outcome relationship that is actually operative in the schedule to be better perceived, by preventing interference with this judgment by the degraded response–outcome relationship that occurs as a result of the delay period. That is, a stimulus serves a discriminative function to signal the end of the operation of effective component, and the start of a second (delay) component, of a chain schedule (Ferster, 1953). Such a stimulus may restrict the effect of any fortuitous response–outcome relationship (or perceived lack of relationship) occurring during the delay period from generalizing to the whole schedule (as would be the case on a tandem schedule where the delay is not differentially signaled) and stop interference with causal detection. Thus, this stimulus effectively produces a chain RI RT schedule, with the components differentially signaled, so that responses emitted fortuitously (or not) close to the outcome in the delay (RT) period would not interfere with judgments about responses in the RI element. The final study examined this prediction.

Experiment 1

Experiment 1 was a systematic replication, using human participants, of the experiments reported from pigeons by Kuroda (2012). In an initial part of training, participants were exposed to a concurrent RI, Ext schedule in a sample part of a trial. Responses to one of two colored circles produced an immediate triangle flash on completion of the schedule, whereas responses to the other colored circle did not. Following the outcome, participants rated the causal effectiveness of their responses. Following this part of the trial, two different colored stimuli would appear, and a response to one of these stimuli would make the triangle flash, depending on the successful color (i.e., that associated with the RI 10 s schedule) in the previous, sample, part of the trial. Following this initial training, a second phase occurred, during which each participant received three types of trial: an immediate outcome in the sample component; a 3-s unsignaled delayed triangle flash in the sample component; and a 3-s signaled delayed outcome in the sample component. Based on previous work with ratings or causal effectiveness (Reed, 1992a, 1992b), ratings of effectiveness should be highest after immediate outcome trials, then signaled delayed outcome trials, and lowest after unsignaled delayed outcome trials. Based on previous work with pigeons (Killeen, 1978; Kuroda, 2012), responses in the choice component should be most accurate following immediate outcome trials, then signaled delayed outcome trials, and least accurate after unsignaled delayed outcome trials. The degree to which responses in the choice component were based on the last response made before the sample triangle flash was also assessed.

Method

Transparency and Openness

We report how we determined sample size, data exclusions, and manipulations. Data are available (including those for the additional online materials) at the Open Science Framework (OSF) page (https://osf.io/c7m45/). Data were analyzed using SPSS Version 26. The study was not preregistered.

Participants

A sample of 70 volunteers (25 male and 45 female) was recruited through the psychology department subject pool. Participants were aged between 18 and 32 years ($M = 21.61 \pm 3.77$ *SD*). Inclusion criteria were that subjects had to be 18 years of age or above and have English as a first language. Participants received no financial remuneration for their participation, but earned credits allowing them to use the subject pool. G-Power calculation implied that for 95% power, with a $p < .05$ criteria, and a medium effect size ($f = .25$), 68 participants would be required for a repeated-measures analysis of variance (ANOVA), with minimal correlations between measures. As a result, it was decided to recruit 70 participants per group. In total 185 individuals started the experiment, but 115 did not complete, and their data were discarded. These individuals all started the study, but they ceased the study prior to completion. They were not asked for reasons, but it may have been the complexity of the task and its length (around 30 min) contributed. Those participants who did not complete, tended not to be learning the task requirements. Ethical approval was obtained through the University psychology Ethics Committee.

Apparatus

The experiment was programmed using Visual Basic 6.0. Participants made responses using a standard PC keyboard and mouse. The screen was white, and four 5 cm diameter circles could be presented: Only two were presented at a time, which could be either two circles equally spaced from each other, with their centers one-third down the screen (for the sample component); or two circles equally spaced from each other, with their centers one-third up from the bottom of the screen (for the choice component). The top two circles could be filled with either blue or orange colors, and the bottom two circles could be filled with either green or red colors. To respond, participants had to move the cursor with the mouse so that the cursor was pointing at a circle, and then click the mouse. Clicking the mouse while the cursor was pointing at either of the two circles shown would sometimes produce a 250-ms white triangle flash in the center of the screen.

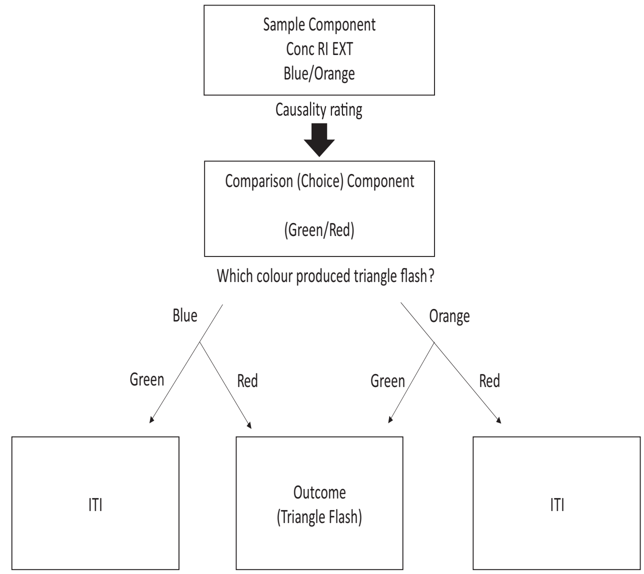
Procedure

Participants were presented with instructions on the screen saying that they would receive many trials, with each trial comprising two parts. In the first part of each trial, their task was to discover the causal effectiveness of a response in making a triangle appear on the monitor and to give a judgment of the extent to which their response was related to the outcome occurring (see Reed, 1999). In the second part of each trial, they had to make the triangle flash by responding to one of two stimuli. The response that produced the triangle flash in the second half of the trial depended on what happened in the first part of the trial, and participants were to learn this by trial and error. When they had read the instructions, participants were instructed to press return, and the trials began. All trials started after a 2-s presentation of a blank screen (intertrial interval [ITI]). A schematic of the overall procedure is shown in Figure 1.

Initial Training. After the 2-s ITI, both of the top two circles were illuminated: one in blue and one in orange (this is termed the "sample" part of the trial). The left-right assignment of colors was quasi-random across trials. During a trial, an RI 10 s schedule was assigned to one of the colors, and an Ext schedule was assigned to the other; that is, a concurrent RI 10 s, Ext schedule was in operation. During each block of eight trials, the RI 10 s schedule was assigned to the left circle for four trials and to the right circle for four trials. The RI 10 s schedule was also assigned to the blue circle for four trials and to the orange circle for four trials. Thus, in any given eight-trial block, the RI schedule was assigned to each color-position combination twice. In each trial, completion of the RI 10 s schedule was followed immediately by a 250-ms triangle flash. Responses to the Ext component had no programmed effects.

Figure 1

Schematic Representation of the Trial Structure for All of the Experiments Reported in This Series



Note. Based on Figure 3 reported by Kuroda (2012). Conc RI EXT = concurrent random interval, extinction; ITI = intertrial interval.

Immediately following the triangle flash (outcome), the screen went blank, and then participants were presented with a 10-cm line on the screen. Above the line appeared the instructions: "Overall, to what extent do you feel responding controlled the triangle lighting up?" Over one end of the line (left), was the text: "not effective at all," and above the other end of the line (right) was the text "completely effective." There was no other text associated with the line. Participants were told, via instructions on the screen, to move a cursor on the screen along the line, using the mouse, to indicate how effective their responses had been in the previous part of the trial. The degree to which participants moved the cursor along the line to the right (i.e., toward "completely effective") was recorded, and the length of movement converted into a score between 0 (no movement) and 100 (the complete 10 cm). If the cursor were moved 5 cm, this would give a score of 50; if it were moved 2.5 cm, this would give a score of 25, etc. When the participants had moved the cursor to the point they wanted, they were instructed to press "return," which initiated the comparison part of the trial.

Following the participants giving their rating of causal effectiveness, the bottom two circles appeared: one green, and one red (for the "comparison" part of the trial). The left-right position of the colors was quasi-random for each trial, with the restriction that the same configuration could not occur on more than three consecutive trials, and equal number of left-right presentations each occurred within a session. A choice was taken as the first response that was made to the circles. This response made both of the circles disappear and could make the triangle flash depending on whether the choice was "correct." A "correct" choice depended on which colored circle produced the triangle in the previous sample part of the trial (i.e., the RI 10 s associated circle). For half of the participants, the red circle was correct if responses to the orange circle had produced the flash in the sample component (i.e., the orange circle had been associated with the RI 10 s schedule); the green circle was correct if responses to the blue circle had produced the triangle flash in the sample component (i.e., the blue circle has been associated with the RI 10 s schedule). The remaining participants received the reverse configuration. A correct choice response resulted in the two circles disappearing, and a 250-ms triangle flash, followed by the 2-s ITI, and the next trial. An incorrect choice resulted in 2,250 ms ITI, followed by the next trial.

This training continued for at least 40 trials (five blocks of eight trials), or until performance in the choice components was at 87% (i.e., 7/8 in each block of eight trials) for at least two successive blocks. Following this training, the test phase commenced.

Test Phase. Participants were not informed about the introduction of the test phase, and training continued as above, with the following alterations. As in the training phase (above), in each trial, two circles were initially presented in the sample phase (orange and blue). One of these circles was associated with an RI 10 s schedule (with different delay contingencies associated with the outcome), and the other was associated with an Ext schedule. There could be three types of RI trial: (a) no delay, (b) unsignaled delayed outcome, and (c) signaled delayed outcome. There were four blocks of 12 trials (i.e., 48 trials in total). In each block of 12 trials, there were four no delay, four unsignaled delays, and four signaled delays. That is, there were 16 presentations of each of the three conditions. Each of the three conditions was associated with each of the four color-position assignments in the sample phase. The participants were required to give a causal rating in the sample phase in all conditions.

In the no delay trials, each trial proceeded as described during the initial training phase; that is, there was a sample component (RI 10 s Ext) and then a comparison component. The unsignaled delay trials were the same, except the schedule was altered to a conc (tand RI *i*-s, RT *j*-s), Ext schedule. Here the mean value of *j* was 3 s, and, for each trial, *i* was 10 s minus *j*. Following this period, during which the circles remained present so as not to provide a signal for the start of a delay period by their removal, a 250-ms triangle flash would occur, as described above. The signaled delayed outcome condition was the same, except that immediately following the response that satisfied the RI *i*-s part of the tandem schedule, a row of six "Xs appeared at the bottom of the screen and remained visible until the triangle flash."

Data Analysis

The data were analyzed in three ways. Firstly, response rates to the two components (RI and Ext) in the first (sample) part of the trial were examined. Secondly, the causal ratings (judgments given after the sample part of the trial) were analyzed using a repeated-measures ANOVA to compare these judgments given after the three trial types (immediate, delayed, and signaled). Similarly, the causal accuracy responses made in the second part of the trials were analyzed by repeated-measures ANOVA to compare these judgments given after the three trial types (immediate, delayed, and signaled). Finally, the impact of response–outcome contiguity on causal accuracy was examined by a repeated-measures ANOVA comparing the degree to which choices in the second part of the trials followed the component in which the last response was made prior to the delivery of the outcome (which could be either RI 10 s or Ext in trials with outcome delays), for the three trial types. For all sets of analysis, in addition to the ANOVA, the effect size (and 95% confidence intervals [CIs]) and the appropriate Bayes statistic (using a uniform noninformative prior; Fienberg, 2006) were calculated. All data are available at the OSF page (https://osf.io/c7m45/).

Results and Discussion

Initial Training

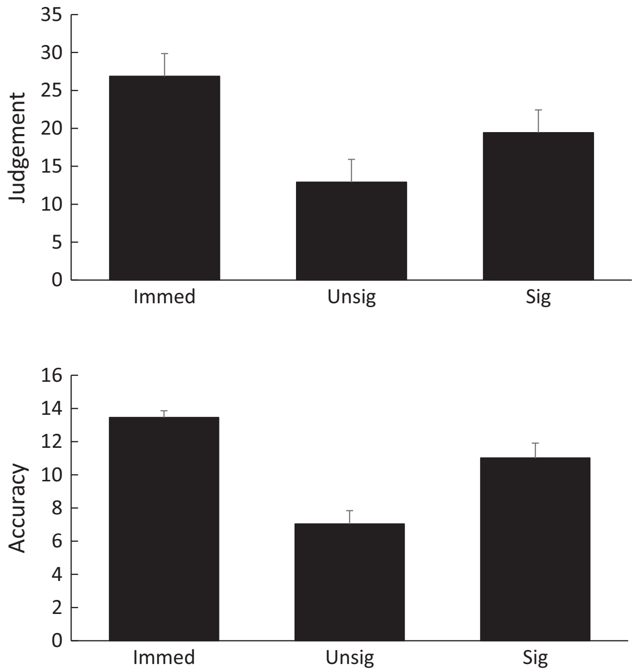
Participants took a mean of 46.17 ($\pm$7.10) trials to reach criterion in the initial training. Over the last two blocks of their initial training prior to reaching criterion, they emitted a mean response rate of 19.64 ($\pm$6.64) responses per min to the ultimately successful (RI) color, and 19.77 ($\pm$5.31) responses per minute to the unsuccessful (Ext) color. It is worth noting that rates of response were not discriminated across the two circles. This is because participants never knew which color would be associated with the RI 10 s schedule on any given trial (the color associated with this schedule was random). They received a mean of 4.68 ($\pm$1.19) outcomes per min. There was a mean probability of an outcome given a response of .24 ($\pm$.07) for the successful (RI) color. The mean judgment of causal effectiveness given after the sample part of the trial (across all three trial types) was 29.42 ($\pm$11.62), and the mean accuracy in the choice component was 15.04 ($\pm$.78) out of a possible 16. The mean rates of response per min, outcomes per min, probability of an outcome given a response for each component, and the actual delay to outcome experienced in each condition during the test trials are given in the additional online materials at the OSF page.

Causal Ratings and Causal Detection

The top panel of Figure 2 shows the mean judgments of causal effectiveness for the three conditions, averaged over the test trials. There were higher ratings given following immediate outcomes than signaled delayed outcomes. Unsignaled delayed outcomes received the lowest ratings. These data confirm previous studies of the effect of unsignaled and signaled outcome delays (Greville & Buehner, 2010; Young & Falmier, 2008). A repeated measures ANOVA, along with effect size and appropriate Bayesian value, revealed a significant effect of condition, $F(2, 138) = 61.71$, $p < .001$, $\eta_p^2 = .472$, 95% CI [0.348, 0.560], $p(H_1/D) = .999$. Paired tests with a Bonferroni correction ($p = .05/3 = .016$) revealed a significant difference between: immediate and unsignaled delay conditions, $t(69) = 11.44$, $p < .001$, $d = 1.36$, $p(H_1/D) = .999$; immediate and signaled delay conditions, $t(69) = 6.27$, $p < .001$, $d = 0.75$, $p(H_1/D) = .999$; and signaled delay and unsignaled delay conditions, $t(69) = 4.79$, $p < .001$, $d = 0.57$, $p(H_1/D) = .998$.

The bottom panel of Figure 2 shows the mean accuracy of detection scores for the conditions over the test trials (possible total = 16). The highest accuracy was following the immediate outcomes, followed by the signaled delay outcomes, and lowest following the unsignaled delayed outcomes. These data demonstrate that for humans in a causal-detection task, detection of a response–outcome relationship is damaged by a delayed outcome (Keely, 1999; Killeen, 1978), but that a signal presented during the delay partly restores this detection (Kuroda, 2012). A repeated-measures ANOVA revealed a significant effect of condition, $F(2, 138) = 99.32$, $p < .001$, $\eta_p^2 = .590$, 95% CI [0.482, 0.662], $p(H_1/D) = .999$. Paired tests with a Bonferroni correction ($p = .05/3 = .016$) revealed

Figure 2
Experiment 1



Note. Top panel shows the mean judgment of causality ratings for the three conditions during test. Bottom panel shows the mean accuracy of detection scores for the three conditions over test. Error bars denote 95% confidence limits. Immed = immediate outcome; unsig = unsignaled outcome delay; sig = signaled outcome delay.

a significant difference between: immediate and unsignaled delay conditions, $t(69) = 7.11$, $p < .001$, $d = 2.21$, $p(H_1/D) = .999$; immediate and the signaled delay conditions, $t(69) = 5.30$, $p < .001$, $d = 0.63$, $p(H_1/D) = .999$; and signaled delay and unsignaled delay conditions, $t(69) = 7.24$, $p < .001$, $d = 0.85$, $p(H_1/D) = .999$.

These data demonstrated that both judgments of causal effectiveness and detection of a response–outcome relationship were reduced by inserting a delay between the response and presentation of the outcome (Reed, 1992a, 1992b) and that this attenuation was partly ameliorated by the insertion of a signal during the delay (Reed, 1992b; Young & Falmier, 2008). These data replicate, for humans in a causality judgment task, findings concerning detection of response–reinforcer relationships in nonhumans (Keely, 1999; Kuroda, 2012) and further suggest correspondences between results from condition and causal judgment studies.

Effect of Response–Outcome Contiguity

The percentage of choice-component responses that corresponded to the previous sample component in which the last response was made before the outcome was presented was calculated. It was possible for responses to be made to either component during the delay prior to the outcome). These data presented a different pattern compared to the percentage of choice component responses that were accurate shown in Figure 2. Percentage choices corresponding to the component in which the last response was made in the sample were necessarily 100% for the immediate condition. However, more choices in the comparison part of the trials corresponding to the last response made in the sample part were made in the unsignaled outcome delay (75.27 ± 15.76) condition than in the signaled delay (61.84 ± 26.28) condition, $t(69) = 3.75$, $p < .001$, $d = 0.448$, $p(H_1/D) = .980$. For the immediate condition, necessarily 16/16 trials had the last response prior to the outcome emitted in the RI schedule. For the unsignaled delay condition, the last response prior to the outcome was emitted on 9/16 RI trials, and this was 8/16 trials for the signaled delay condition. The rates of response during the delay for the various components of the various conditions are shown and analyzed in the additional online materials at the OSF page and did not differ from one another (except for the immediate RI condition, which was zero).

In sum, these data corroborate the patterns of judgment of causal effectiveness noted in previous studies of unsignaled and signaled outcome delays (Greville & Buehner, 2010; Reed 1992b; Young & Falmier, 2008). They also demonstrate that participants' ability to detect a causal relationship (irrespective of how strong they believe that relationship to be) is impaired by a delay and partially restored when that outcome delay is filled by a signal. These data suggest that similar findings are seen, in this respect, to those noted in studies of nonhuman conditioning (Keely, 1999; Kuroda, 2012).

In terms of the signal function, the data suggest that participants were making a choice concerning the response made in closest temporal contiguity to the outcome in the immediate and unsignaled delay condition, suggestive of a correspondence with an analysis of the detrimental impact of a delayed reinforcement by Sizemore and Lattal (1978; see also Ferster, 1953; Schaal & Branch, 1988), in terms of the reinforcement of competing behaviors. This did not occur in the signaled condition, where the last response made during the signaled delay was not necessarily to the RI component circle (but could have been made to the Ext component); however, these responses made during the delay did not interfere with the detection of causality. Possibly this was a consequence of the signal giving information that what follows it may be of limited relevance to the outcome, that is, the signal marked the cause of the event or gave information that the outcome was to be expected.

One aspect of the data is worth some comment, which was the high number of participants who did not complete the task. Informal feedback suggested that this was, in part, due to the length of the procedure, and the degree of potential confusion generated by having three different trial types during test. As a consequence, subsequent studies adopted a between-subject approach.

Experiment 2

Experiment 2 sought to replicate the findings from Experiment 1 with respect to the effects of unsignaled and signaled outcome delays, but also explored whether a brief stimulus would produce higher ratings of causal effectiveness and promote causal perception, compared to an unsignaled delayed outcome. Additionally, it analyzed whether the brief delay signal would ameliorate the interfering impact of responses made after the termination of the effective schedule (RI). It has been suggested that the signal between the response and outcome does not have to fill the entire delay period in order to be effective (Reed, 1992a, 1992b; Young & Falmier, 2008). This has previously been noted in studies of nonhuman conditioning with a delay of reinforcement (Lieberman et al., 1979;

Williams, 1991), and it has been suggested to show that the response–outcome delay does not have to be perceptually bridged in order for the stimulus to be effective (Williams, 1991). This result would still be compatible with conditioned reinforcement (Williams, 1991) and marking (Lieberman et al., 1979) views. Perceptual catalysis would suggest that a stimulus only filling part of the outcome delay period would not be as effective as one completely bridging the period (Rescorla, 1982).

Method

Participants and Apparatus

A sample of 120 volunteers (53 male and 67 female), aged between 18 and 32 years ($M = 21.39 \pm 3.63$), were recruited, with the experiment being presented through an online delivery platform. G-Power calculation implied that for 95% power, with a $p < .05$ criteria, and a large effect size ($f' = .40$), that 112 participants would be required for a between-subject ANOVA. In total 182 individuals started the experiment, but 62 did not complete, and their data were discarded. Group allocation was random until 30 participants had been allocated to any one group, then new participants were randomly allocated to the remaining groups. Ethical approval was obtained through the University Psychology Ethics Committee.

Procedure

The same procedures were used as in Experiment 1, except that in the present study any one group received only one of the treatments. In the test phase, the outcomes following completion of the RI 10 s depended on group allocation and were either: (a) no delay, (b) unsignaled delayed outcome, and (c) signaled delayed outcome, as described in Experiment 1. For the fourth group, there was a briefly signaled delayed outcome trials. This group experienced the same contingency as in the signaled delay group, except that the row of six "Xs remained visible for 1 s, only instead of filling the delay period, and there then followed a 2-s unsignaled delay until the triangle flash outcome." This procedure was adopted so that the brief delay signal was not contiguous with the outcome (only with the response). However, this meant that the scheduled outcome delay in this group was always 3 s, whereas the scheduled delay in the other groups had a mean of 3 s, but was variable. The data were analyzed as described in Experiment 1, except that between-subject ANOVAs were employed.

Results and Discussion

Initial Training

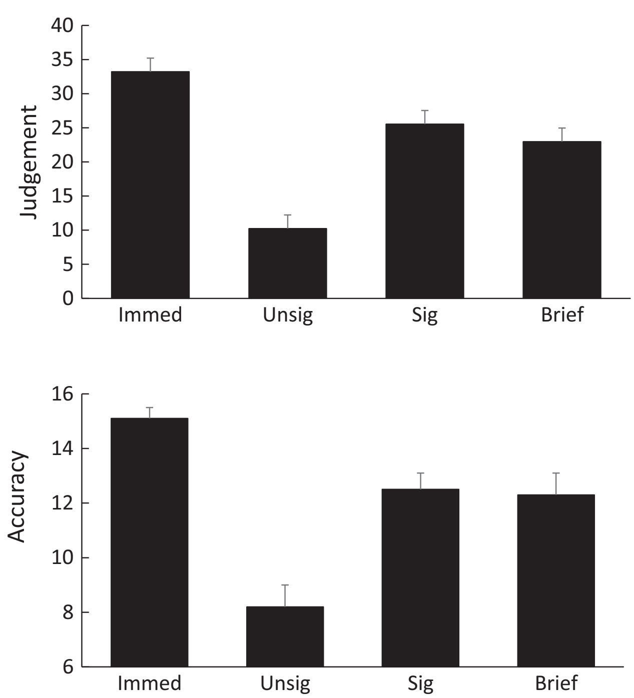
Participants took a mean of 44.40 ($\pm$5.47) trials to reach criterion. In the last two blocks of their training prior to reaching criterion, they emitted a mean of 25.68 ($\pm$10.38) responses per min to the ultimately successful (RI) color, and 25.58 ($\pm$10.56) responses per minute to the unsuccessful (Ext) color. Response rates were not different across the two circles as participants never knew which color would be associated with the RI 10 s schedule on any given trial (the color associated with this schedule was random). They received a mean 5.00 ($\pm$1.31) outcomes per min, with a mean of 0.23 ($\pm$0.12) probability of an outcome given a response to the successful (RI) color. The mean judgment of causal effectiveness following the sample components across all trial types was 32.14 ($\pm$14.70), and the mean accuracy in the choice component across all trial types was 15.34 ($\pm$.75) out of 16. The additional online materials at the OSF page present the mean rates of response per min, outcomes per min, probability of an outcome given a response for each component, and the actual delay to outcome experienced in each condition during the test trials.

Causal Ratings and Detection

The top panel of Figure 3 shows mean judgments of causal effectiveness for the four groups, averaged over the test trials. Higher ratings were given by the immediate outcomes group than by the signaled, or briefly-signaled, delayed outcome groups, with the unsignaled delayed outcome group producing the lowest rating. These data confirm previous studies using similar designs (Reed, 1992a, 1992b; Young & Falmier, 2008). A between-subject ANOVA revealed a significant effect of group, $F(3, 116) = 103.54$, $p < .001$, $\eta_p^2 = .728$, 95% CI [0.638, 0.778], $p(H_1/D) = .999$. Tukey's Honestly Significant Difference (HSD) tests revealed a significant difference between the immediate and each of the other groups, and between both the signaled and brief signal groups and the unsignaled delay, all ps $< .05$. There was no difference between the signaled and brief signal groups.

The bottom panel of Figure 3 shows the mean accuracy of detection scores for the conditions over the test trials (possible total = 16).

Figure 3
Experiment 2



Note. Top panel shows the mean judgment of causality ratings for the four groups during test. Bottom panel shows the mean accuracy of detection scores for the four groups over test. Error bars denote 95% confidence limits. Immed = immediate outcome; unsig = unsignaled outcome delay; sig = signaled outcome delay; brief = brief signal for outcome delay.

Highest accuracy was in the immediate group, followed by the signaled and briefly signaled delay groups, and lowest in the unsignaled delay. A between-subject ANOVA revealed a significant effect of condition, $F(3, 116) = 91.34$, $p < .001$, $\eta_p^2 = .702$, 95% CI [0.605, 0.757], $p(H_1/D) = .999$. Tukey's HSD tests revealed a significant difference between the immediate and each of the other groups, and between both the signaled and brief signal groups and the unsignaled delay, all $ps < .05$. There was no difference between the signaled and brief signal groups.

These data from replicated the findings from the current Experiment 1, and they extended the effect of a delay signal to the action of a brief signal presented following a response but not temporally contiguous with the outcome. This suggests that the signal does not have to bridge the response and outcome to be effective (Williams, 1991). Such findings have previously been reported for human causality ratings (Reed, 1992a, 1992b; Young & Falmier, 2008), and also for nonhuman conditioning (Lieberman et al., 1979; Williams, 1991). The effects noted were similar for both ratings and accuracy of causal perception, and suggest that a signals for an outcome delay have similar effects for human ratings of causal effectiveness as they do for response–reinforcer detection in conditioning procedures for nonhumans (Kuroda, 2012).

Effect of Response–Outcome Contiguity

The percentage of choice-component responses that corresponded to the previous sample component in which the last response was made before the outcome was presented were calculated. Percentage choices corresponding to the sample component in which last response was made (i.e., considering responses made during any outcome delay), for the immediate condition were necessarily 100%. In the other three groups, this choice was greater for the unsignaled outcome delay group (72.79 ± 12.97), than for the brief signaled delay group (64.79 ± 12.55), and was lowest in the signaled delay group (53.12 ± 10.48). A between-subject ANOVA for all groups except the immediate outcome group revealed a significant effect of group, $F(2, 87) = 20.04$, $p < .001$, $\eta_p^2 = .315$, 95% CI [0.312, 0.613], $p(H_1/D) = .999$. Tukey's HSD tests revealed a significant difference between all pairs of groups, all $ps < .05$. For the immediate condition, 16/16 trials had the last response prior to the outcome emitted in the RI schedule. For the unsignaled and brief signaled groups, the last response prior to the outcome was emitted on 7/16 RI trials, and this was 8/16 trials for the signaled delay condition. The rates of response during the delay for the various components of the various conditions are shown and analyzed in the additional online materials at the OSF page and did not differ from one another (except for the immediate RI component condition).

As with Experiment 1, these data for causal judgments and causal perceptions follow those reported previously for nonhuman studies of instrumental conditioning (Williams, 1991), as well as those from previous investigations of human judgments of causal ratings with delayed outcomes (Greville & Buehner, 2010; Reed, 1999). They also extend the findings reported in Experiment 1 to show that a brief stimulus presented following a response will serve to promote otherwise reduced ratings and perceptions of causality (Reed, 1999). This suggests that the signal does not have to bridge the temporal gap between response and outcome to be effective (Lieberman et al., 1979; Schaal & Branch, 1988). These findings leave open several alternative theoretical accounts of these effects, including versions of conditioned reinforcement (Williams, 1991) and marking (Lieberman et al., 1979; Reed, 1999). However, the data are more ambiguous for perceptual catalysis, this view would suggest that a brief stimulus not filling the delay period would be less effective than a stimulus providing full temporal bridging, and it was not. However, it could always be assumed that catalysis could occur in memory as opposed to actually bridging a temporal gap.

The current results also are compatible with the signal serving a discriminative function to give information about the fulfilment of the schedule (Ferster, 1953). That is, the delay stimulus actually serves as a discriminative stimulus to signal the operation of a second component of the chain schedule. This may restrict the effect of any fortuitously response (or no response) outcome pairings occurring at the end of the delay period generalizing to the whole schedule (as would be the case on a tandem schedule where the delay is not differentially signaled) and stop them interfering with judgments of effectiveness regarding the target response. The stimulus effectively produces a chain RI RT schedule, with the components differentially signaled, so that responses emitted fortuitously close to the outcome in the delay (RT) period would not interfere with judgments about responses in the RI element. That responses (or no responses) emitted during the outcome delay period appeared to interfere with ratings and perception of causality in the unsignaled delay condition, but not in the signaled and briefly signaled conditions, supports this latter discriminative view.

Experiment 3

Experiment 3 extended the investigation to explore whether a stimulus presented during a delay period would facilitate judgments and detection of causal relationships when it was presented following responses to both successful (RI) and unsuccessful (Ext) components (Lieberman et al., 1979; Rescorla, 1982; Williams, 1991). Under these situations, as suggested by Rescorla (1982; see also Williams, 1991), both marking and perceptual catalysis views still predict that learning about the target response (in this case, those made on the RI 10 s component) would be enhanced. If the stimulus is not thought of as an outcome that can support ratings in its own right (see Buehner, 2005, for such an argument), this view would predict higher ratings for the RI 10 s with the stimulus compared to in its absence. The conditioned reinforcement view suggests that the impact of the stimulus may be reduced somewhat, as it is not always followed by an outcome, but its partial association with the outcome would still allow some elevation of ratings.

In contrast, the discriminative function view derived from Ferster (1953) would predict that such a stimulus would not enhance ratings or perceptions of causality. In situations where the stimulus occurs in both components, it would not serve to signal the operation of the outcome-delay period, which even brief stimuli can do effectively (Neuringer & Chung, 1967), as it could come after responses to the Ext component not followed by an outcome. As its association with the termination of one component would be degraded, then its discriminative function would be undermined. Thus, a tandem RI FT schedule would be in operation, in which the impact of the final component would generalize back across the entire schedule; in contrast to where the outcome-delay period is reliably signaled, producing a chain RI FT schedule and learning in the latter (delay) portion would not generalize to the initial part of the schedule. If this were the case, then the stimulus would not impede the interfering effects responses made in closer temporal contiguity to the outcome in the delay (RT) period, and judgments of the effectiveness and accuracy of responses to the RI schedule would not be enhanced relative to a no delay stimulus condition.

Method

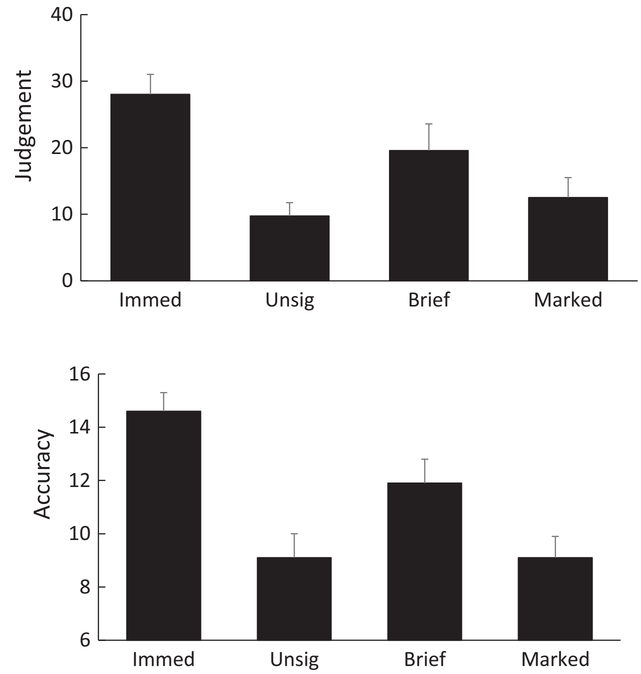
Participants and Apparatus

A sample of 120 volunteers (56 male and 64 female), aged between 18 and 31 years ($M = 21.15 \pm 3.31$), were recruited as described in Experiment 1. In total, 165 individuals started the experiment, but 45 did not complete, and their data were discarded. The apparatus was as described in Experiment 1.

Procedure

The initial training procedure was as described for Experiment 2. In the next phase, training was as described as in Experiment 2 for the no delay, unsignaled delayed outcome, and briefly signaled delayed outcome groups. For the "marking" group, the 6X stimulus was presented following the response that satisfied the RI requirement, as in the briefly signaled delay group, but it was also presented on an RI *i*-s schedule in the Ext schedule. There was a 6X stimulus presented on every trial for the RI component, but it was also presented once a trial in the Ext component. The data were analyzed as described in Experiment 1, except that between-subject factors were employed in mixed-model designs were appropriate.

Figure 4
Experiment 3



Note. Top panel shows the mean judgment of causality ratings for the four groups during test. Bottom panel shows the mean accuracy of detection scores for the four groups over test. Error bars denote 95% confidence limits. Immed = immediate outcome; unsig = unsignaled outcome delay; brief = brief signal for outcome delay; marked = stimulus presented for responses in both components.

Results and Discussion

Initial Training

Participants took a mean of 43.33 ($\pm$6.92) trials to reach criterion. During the last two blocks of training prior to criterion, they emitted a mean of 23.25 ($\pm$8.78) responses per min to the ultimately successful (RI) color, and 21.89 ($\pm$8.48) responses per minute to the unsuccessful (Ext) color. Response rates were not different across the two circles as participants never knew which color would be associated with the RI 10 s schedule on any given trial (the color associated with this schedule was random). They received a mean of 5.86 ($\pm$1.67) outcomes per min, with a mean probability of an outcome given a response to the successful (RI) color of 0.28 ($\pm$0.13). The mean judgment of causal effectiveness across all trials was 28.57 ($\pm$11.10), and the mean accuracy was 15.04 ($\pm$0.77) out of 16. See the additional online materials at the OSF page for mean response rate, outcome rate, outcome/response probability, and the actual delay to outcome experienced in each condition during the test trials.

Causal Ratings and Detection

The top panel of Figure 4 shows mean judgments of causal effectiveness for the four groups averaged over all test trials. Higher ratings were given by the immediate outcome group than for the briefly signaled delayed outcome groups, and the unsignaled delayed outcome, groups produced a low rating. The marking group also gave a low rating of the causal effectiveness of a response. This is a similar result to that noted for learning in rats reported by Williams (1991) and suggests that there is no marking function for humans as there is in nonhuman studies of delayed reinforcement (Lieberman et al., 1979). A between-subject ANOVA corroborated these impressions, revealing a significant effect of group, $F(3, 116) = 33.41$, $p < .001$, $\eta_p^2 = .463$, 95% CI [0.321, 0.556], $p(H_1/D) = .999$. Tukey's HSD tests revealed a significant difference between the immediate and each of the other groups, and between the briefly signaled and both the marked and the unsignaled delay, all *p*s $< .05$. There was no difference between the unsignaled and marked groups.

The bottom panel of Figure 4 shows the mean accuracy of detection scores for the conditions over the test trials (possible total = 16). There was highest accuracy in the immediate group, followed by the briefly signaled delay groups, and lowest in the marked and unsignaled delay. A between-subject ANOVA revealed a significant effect of group, $F(3, 116) = 44.47$, $p < .001$, $\eta_p^2 = .534$, 95% CI [0.401, 0.617], $p(H_1/D) = .999$. Tukey's HSD tests revealed a significant difference between the immediate and each of the other groups, and between the briefly signaled and both the marked and the unsignaled delay, all *p*s $< .05$. There was no difference between the unsignaled and marked groups.

Effect of Response–Outcome Contiguity

The percentage of choice-component responses that corresponded to the previous sample-component in which the last response was made before the outcome was necessarily 100% for the immediate group. In the other three groups, this choice was greater for the unsignaled (79.79 ± 20.41), and marked (73.54 ± 19.11), delayed outcome groups compared to the brief-signaled delay group (62.71 ± 21.93). A between-subject ANOVA for all groups except the immediate outcome group, revealed a significant effect of group, $F(2, 87) = 5.32$, $p = .007$, $\eta_p^2 = .109$, 95% CI [0.004, 298], $p(H_1/D) = .779$. Tukey's HSD tests revealed a significant difference both the unsignaled, and marked outcome, delay groups and the brief signaled outcome delay group, both ps < .05, but not between the unsignaled and marked groups. For the immediate condition, 16/16 trials had the last response prior to the outcome emitted in the RI schedule. For the brief signaled group, the last response prior to the outcome was emitted on 7/16 RI trials, and this was 8/16 trials for the unsignaled and marked groups. The rates of response during the delay for the various components of the various conditions are shown and analyzed in the additional online materials at the OSF page and did not differ from one another (except for the immediate RI component condition).

Taken together, these data corroborate those for Experiments 1 and 2 for the immediate, unsignaled delay, and brief signaled delay groups. Both ratings of causal effectiveness, and the detection of the causal relationship, were improved by a stimulus filling an outcome delay, and a stimulus filling only the first part of an outcome delay, as they were in Experiment 2 (see also Kuroda, 2012; Reed, 1992a, 1992b), and as it is for nonhuman instrumental conditioning (Schaal & Branch, 1988). This was not the case for a stimulus following some responses on both RI and Ext schedules. Although stimuli presented following correct and incorrect responses have been found to facilitate learning in some simple conditioning situations with delayed reinforcement (Lieberman et al., 1979), it has been shown that, in two-choice discrimination procedures, such stimuli do not serve to improve learning (Williams, 1991).

That the stimulus following successful and unsuccessful responses did not improve causal judgment or accuracy in the current experiment is not compatible with a pure marking interpretation (Lieberman et al., 1979), which would have predicted this effect. Similarly, if the Xs served as another outcome (Buehner, 2005), then causal judgments should have been high in both RI and Ext components. The implications of this result for conditioned reinforcement views are less clear; that is, the stimuli did not elevate ratings may be compatible with a conditioned reinforcement view, as presentations after incorrect responses would have reduced the effectiveness of the "Xs stimulus as a condition reinforcer. The results are consistent with the discriminative function view" (Ferster, 1953; Pliskoff et al., 1964), as the stimulus in the marking group could not serve to signal the operation of a chain RI RT schedule, as its association with the termination of one component would be degraded. As a consequence, it could not impede the interfering effects on judgments of responses made in closer temporal contiguity to the outcome in the delay (RT) period.

Experiment 4

The final experiment in this series further examined the discriminative role of a signal for a delay. This study explored whether a stimulus presented to indicate when reinforcement is available for collection by the next response would facilitate detection and judgments of causality. Such a stimulus is presented on timing out of the RI schedule, but prior to the successful response, and remains present until the outcome delivery. This procedure has been termed a "signaled reinforcement" procedure (Ferster & Skinner, 1957). When such a stimulus is presented, it tends to lower rates of responding on VI schedules, but it makes responding more optimal with regard to the interval controlling delivery of reinforcement (McCoy et al., 1976; Tarpy et al., 1986). According to the marking hypothesis, a stimulus that signals the availability of an outcome for a response should serve to mark the response as effectively as a stimulus occurring after the response (Lieberman et al., 1985). There is evidence to suggest that this will occur in some circumstances with humans (Grindle & Remington, 2005). If this is the case, then both judgments and detection should be improved with signaled reinforcement procedures; however, that it has been found to reduce response rates on VI schedules in instrumental conditioning may suggest this will not be the case (Tarpy et al., 1986). In contrast, a discriminative view of a delay stimulus suggests that a signaled reinforcement procedure may have differential effects on causal judgments and causal detection. The stimulus would serve to reduce the ratings of causal effectiveness, as it serves to reduce response rates (McCoy et al., 1976). This reduction could be due either to improved temporal discrimination of when to respond (Ferster & Skinner, 1957), or overshadowing of the response–outcome relationship by the stimulus–outcome association (see Schachtman & Reilly, 1987; Tarpy et al., 1986). However, the stimulus in a signaled reinforcement procedure could simultaneously improve detection of causal accuracy, as it may demark the delay (RT) period from the effective (RI) part of a schedule.

Method

Participants and Apparatus

A sample of 120 volunteers (44 male and 78 female), aged between 18 and 32 years ($M = 21.25 \pm 3.45$) were recruited as described in Experiment 1. In total, 183 individuals started the experiment, but 63 did not complete, and their data were discarded. The apparatus was as described in Experiment 1.

Procedure

The initial training procedure was as described for Experiment 1. The test phase was as described in Experiment 1 for the no delay, unsignaled delayed outcome, and signaled delayed outcome groups. For the final group (signaled reinforcement), when the RI *i*-s schedule had timed out, and the next response to the RI component circle would be reinforced, the row of 6Xs appeared on the screen. The 6Xs remained present until a response had been made to the RI component circle, when it disappeared, and the outcome was delivered 3 s later. See Figure 5 for a schematic of the different conditions. The following "accuracy" part of the trial was as described for the other groups, with the triangle flash being delivered if the circle associated with the preceding RI component was responded to. The data were analyzed as described in Experiment 1, except that between-subject factors were employed in mixed-model designs were appropriate.

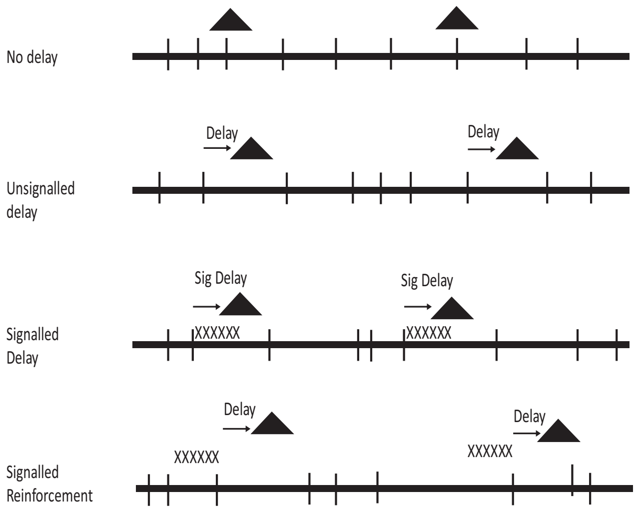
Results and Discussion

Initial Training

Participants took a mean of 44.80 (±6.08) trials to reach criterion. During the last two blocks of training prior to reaching criterion, they

Figure 5

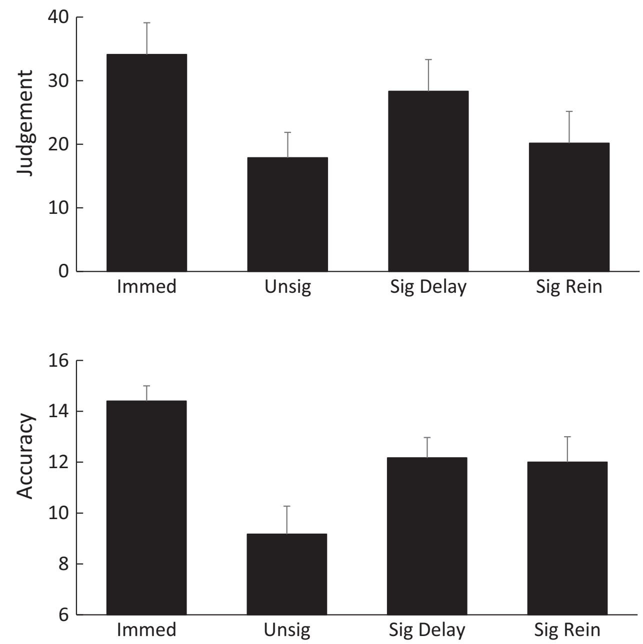
Schematic Representation of Experiment 4



Note. Schematic of the response–signal–outcome relationship in the four groups. No delay = no response–outcome delay. Unsignaled delay = 3 s delay from response to outcome. Signaled delay = 3 s delay to outcome filled by row of 6Xs. Signaled reinforcement = row of 6Xs signals when the next response can produce an outcome, which is delayed by 3 s from the response. Sig = signalled.

Figure 6

Experiment 4 Results



Note. Top panel shows the mean judgment of causality ratings for the four groups during test. Bottom panel shows the mean accuracy of detection scores for the four groups over test. Immed = immediate outcome; unsig = unsignaled outcome delay; sig delay = signal for outcome delay; sig rein = stimulus presented when response will produce an outcome and remained on during delay = 95% confidence limits.

emitted a mean of 19.98 (±6.66) responses per min to the ultimately successful (RI) color, and 19.85 (±7.74) responses per minute to the unsuccessful (Ext) color. Response rates were not different across the two circles as participants never knew which color would be associated with the RI 10 s schedule on any given trial (the color associated with this schedule was random). They received a mean of 5.31 (±1.56) outcomes per min, with a mean probability of an outcome given a response to the successful color of 0.29 (±0.11), with a probability of 0.14 (±0.06) overall. The mean judgment of causal effectiveness across all test trials was 33.57 (±11.22), and the mean accuracy in the choice phase was 14.99 (±0.83) out of 16. The mean rates of response per min, outcomes per min, probability of an outcome given a response for each component, and the actual delay to outcome experienced in each condition during the test trials are given in the additional online materials at the OSF page.

Causal Ratings and Detection

The top panel of Figure 6 shows mean judgments of causal effectiveness for the four groups averaged over all test trials for the target trial type. There were higher ratings made by the immediate outcome and signaled delayed outcome groups compared to the other two groups. A between-subject ANOVA revealed a significant effect of group, $F(3, 116) = 11.15$, $p < .001$, $\eta_p^2 = .223$, 95% CI [0.090, 0.332], $p(H_1/D) = .984$. Tukey's HSD tests revealed a significant difference between both the immediate and the signaled delay and each of the other two groups, all $ps < .05$. There were no other pairwise differences. These data replicate the previous findings using the current procedure for unsignaled and signaled outcome delays (see also Greville & Buehner, 2010; Young & Falmier, 2008). They also extend the functional equivalence between instrumental conditioning and human causality judgment results, by showing that a signaled reinforcement (i.e., presenting a stimulus when reinforcement becomes available to collect) procedure reduces the ratings of causal effectiveness relative to an immediate outcome, just as it reduces response rates (McCoy et al., 1976; Tarpy et al., 1986). This reduction may reflect improved temporal discrimination of responding (Ferster & Skinner, 1957), or overshadowing of the response–outcome relationship by the stimulus–outcome association (Schachtman & Reilly, 1987; Tarpy et al., 1986).

The bottom panel of Figure 6 shows the mean accuracy of detection scores for the conditions over the test trials (possible total = 16). Inspection of these data shows highest accuracy in the immediate group, followed by the signaled delay and signaled reinforcement groups, and lowest in the unsignaled delay group. A between-subject ANOVA revealed a significant effect of group, $F(3, 116) = 24.19$, $p < .001$, $\eta_p^2 = .384$, 95% CI [0.237, 0.485], $p(H_1/D) = .999$. Tukey's HSD tests revealed a significant difference between the immediate and each of the other groups, and between both the signaled delay and signaled reinforcement groups and the unsignaled delay group, all $ps < .05$. There was no difference between the signaled delay and signaled reinforcement groups. In contrast to the effect of a signaled reinforcement on ratings of causal effectiveness, this procedure improved detection of causal relationships relative to an unsignaled delay. That the signaled reinforcement procedure improved detection of causal accuracy can be explained by the discriminative view (Ferster, 1953), as the stimulus would demark the delay (RT) period from the effective (RI) part of a schedule.

Effect of Response–Outcome Contiguity

Percentage choices corresponding to the sample component in which last response was made were 100% for the immediate group. In the other three groups, this choice was greater for the unsignaled group (79.79 $\pm$ 8.94), than for the signaled delay (60.21 $\pm$ 12.33) and the signaled reinforcement (60.42 $\pm$ 14.61) groups. A between-subject ANOVA for all groups except the immediate outcome group, revealed a significant effect of group, $F(2, 87) = 25.53$, $p < .001$, $\eta_p^2 = .370$, 95% CI [0.123, 0.592], $p(H_1/D) = .999$. Tukey's HSD tests revealed a significant difference between the unsignaled group and both the signaled delay and signaled reinforcement groups, both ps $< .05$, but not between the signaled delay and signaled reinforcement groups. For the immediate condition, 16/16 trials had the last response prior to the outcome emitted in the RI schedule. For the signaled delay group, the last response prior to the outcome was emitted on 7/16 RI trials, and this was 8/16 trials for the unsignaled delay and signaled reinforcement groups. The rates of response during the delay for the various components of the various conditions are shown and analyzed in the additional online materials at the OSF page and did not differ from one another (except for the immediate RI component condition).

This signaled reinforcement procedure lowered judgments of causality (see McCoy et al., 1976; Tarpy et al., 1986, for similar effects on response rates in instrumental conditioning), but improved detection of the causal relationship. These results taken together would not be predicted by the marking hypothesis, which suggests such a cue should mark the response as effective and improve both ratings and detection of causality (Grindle & Remington, 2005; Lieberman et al., 1985). In contrast, the discriminative view of a delay stimulus (Ferster, 1953) implies that the signal for reinforcement would improve detection of causal accuracy, as it would demark the delay period from the effective part of a schedule, and prevent interference from delay responses. This is consistent with the present results obtained from a comparison of the choice responses based on the last response made. It should be noted that the dissociation in performance between the judgment and accuracy aspects of the tasks may imply that these two responses are tapping into different processes. It may be that the rating measure is assessing causal judgment about the relation between response and outcomes, but the "accuracy" measure is assessing learning of the biconditional stimulus–stimulus rules. If this were the case, then it may not be so surprising that the signaled reinforcement condition would have different effects on the two outcome measures. However, the data from the contingency analysis do suggest that what is most important is the degree to which responses are contiguous to an outcome, and the degree to which the impact of this contiguity is restricted to one particular stimulus. At the very least, this implies some contribution of the response–outcome association to both of these measures.

General Discussion

The current series of experiments explored the effect of signals presented during outcome delays on human ratings and detection of causal relationships, using a novel paradigm for this sort of task. Participants responding on a conc RI Ext schedule for a triangle flash in a sample component, then matched responses in a subsequent choice component to those causing the flash in the previous sample component. Immediate outcomes were associated with higher causal ratings and better causal detection than outcomes delayed by 3 s (Experiments 1 and 2; Michotte, 1946; Reed, 1992a, 1992b; van Elk et al., 2014; Young & Nguyen, 2009). In Experiments 1, 2, and 4, the presentation of a stimulus filling the outcome delay period alleviated these deficits (Greville & Buehner, 2010; Reed, 1999; Young & Falmier, 2008). A signal during the delay improved ratings and detection even when it was contiguous with the response but not the reinforcement (Experiments 2 and 3; Lieberman et al., 1979; Reed, 1992a, 1992b; Schaal & Branch, 1988). However, stimuli associated with both components (marking cues) did not impact judgments or detection (Experiment 3), as has been found in studies of two-choice discrimination in rats (Williams, 1991). Stimuli signaling the availability of an outcome if a response were made (signaled reinforcement) did not improve causal judgments (cf., McCoy et al., 1976; Tarpy et al., 1986), but they did improve detection of stimuli associated with the outcome (Experiment 4).

Many of the aspects of these results mirror those noted in studies on nonhuman instrumental conditioning and extend the suggestion that similar processes may control both instrumental learning and human judgments of causal effectiveness (Allan, 1993; Dickinson & Balleine, 2000). The novel empirical extension was to causal-detection paradigms, previously used to show that reinforcement delays impact nonhumans ability to detect response–reinforcer contingencies (Killeen, 1978; Kuroda, 2012; Lattal, 1975). The current results also provide some insight into which mechanisms may operate when the delay period is signaled.

Conditioned reinforcement is traditionally the account of the function of a stimulus signaling a delay (Royalty et al., 1987; Williams, 1991). Many aspects of the results reported here, such as the action of a stimulus signaling the entire outcome delay period, are compatible with such a view (see Experiments 1, 2, and 4). However, there are some aspects of these current data which are harder for this view to accommodate. For example, the briefly signaled outcome delays which facilitated judgments and casual detection (Experiments 2 and 3), were not contiguous with the outcome, which would reduce the conditioned reinforcing strength of the stimulus. Of course, a stimulus does not need to be contiguous with a reinforcer to acquire conditioned reinforcing properties, but can acquire such properties by signaling its relative imminent arrival (Fantino, 1977). The dissociation between causal judgment and causal detection in Experiment 4 also could be difficult for a straightforward version of conditioned reinforcement to accommodate, if it is assumed that its impact on one aspect of behavior should be the same as its impact on other aspects. Of course, causal judgment and causal detection may be based on different processes, and conditioned reinforcement may impact one to a greater extent than the other.

The "marking hypothesis" (Lieberman et al., 1979, 1985) could explain the facilitating action on causal judgment and detection of the brief stimuli in the current Experiments 2 and 3, as well as the action of the stimulus presented contiguously with a response initiating the delay but not with the outcome. This would be analogous to the explanations offered in instrumental responding (Schaal & Branch, 1988), and nonhuman causal detection (Kuroda, 2012). However, this effect only occurred when the brief stimulus was associated with the successful outcome (Experiments 2 and 3), and not when it followed both, and only under these situations can a marking or perceptual catalysis view be accepted over a conditioned reinforcing view of the action of the delay stimulus (Rescorla, 1982). These data also correspond to those reported for nonhuman conditioning in a two-choice discrimination procedures, where Williams (1991) noted that presentations of a stimulus following only correct choices facilitated a discrimination by rats (see also Reed, 1996), and there was little effect of stimulus duration relative to the delay period (see also Kuroda, 2012). It may also be noted that marking effects have proved elusive in other human procedures (Grindle & Remington, 2005; Reed, 1998).

Perceptual catalysis (Rescorla, 1982; Young & Nguyen, 2009) has some difficulty with results from the current experiments. According to this view, it would be expected that, when the delay is completely filled, the target and outcome should be linked together perceptually. A stimulus only filling part of the delay would not be as effective in producing this bridging. However, the delay period did not need to be filled for the stimulus to be effective (Experiments 2 and 3), which suggests more than a perceptual bridging effect is operating (Lieberman et al., 1985; Schaal & Branch, 1988; Williams, 1991). Of course, it may be that the bridging occurs in memory, and, so long as the representation of the response and the outcome occur together, then they may be processed as jointly occurring. However, if memory processes are invoked the effect is not perceptual but memorial, and is functionally identical to marking (Lieberman et al., 1985). Moreover, perceptual catalysis has been challenged also by a number of previous findings. For example, Williams (1991) noted that presentations of a stimulus following only correct choices (not both correct and incorrect) facilitated a discrimination by rats, and there was little effect of the tone's duration relative to the delay period (see also Kuroda, 2012).

Thus, while the empirical similarities between instrumental conditioning with delayed reinforcement, and human judgments of causal effectiveness with delayed outcomes are clear, the mechanisms often suggested for instrumental conditioning do not easily accommodate the effects noted. However, this may not mean that the two systems are driven by different mechanisms, rather just that the often posited mechanisms of conditioned reinforcement, marking, and perceptual catalysis, are not ubiquitous. In fact, such views do not accommodate all effects seen in instrumental conditioning with signaled delays (Dinsmoor, 1983; Fantino, 1977; Williams, 1991). That facilitating effects of delay signals are not always reducible to these mechanisms may be important in the context of human judgments of causality (Buehner & May, 2003; Young, 1995). Pavlov (1927) highlighted the potential importance of a "second signaling system" for humans, especially, and Keller and Schoenfeld (1950; Pliskoff et al., 1964; Wyckoff, 1952) suggested that "conditioned reinforcement" is dependent on the discriminative function of a stimulus, a concept of importance as "informational" cues are of prime significance to the judgment and detection of causal relationships (Buehner & May, 2003; Young, 1995).

A discriminative function of a delay stimulus was highlighted by Ferster (1953). This account suggests that a delay stimulus serves to signal the operation of different components of the contingency, and acts as an occasion setter for responses in either element of the schedule. Behavior other than the target can become associated with the outcome, and its reinforcement can interfere with response rates (Sizemore & Lattal, 1978), or produce uncertainty or shifts in attributions regarding the cause of an outcome (Buehner, 2005). It may be that the discriminative or informational functions of a delay stimulus are critical in human ratings of causal effectiveness (Buehner & May, 2003; Young, 1995).

In support of this view, the current experiments noted that responses made during delay outcomes interfered with detection of the actual relationship with unsignaled delays (Experiments 1–4). However, this was not the case with fully or briefly signaled delays (Experiments 2–4), or with signaled reinforcement (Experiment 4). These results suggest a delay stimulus serves to signal the response has been successful, and demark the delay period by serving a discriminative function. In Experiment 3, a marking stimulus did not serve to signal the operation of a chain RI RT schedule, as its association with the termination of one component would be degraded. As a consequence, it could not impede the interfering effects on judgments of responses made in closer temporal contiguity to the outcome in the delay (RT) period. That a stimulus serving as a reinforcement signal (Experiment 4) served to reduce the impact of interfering delay responses and elevate causal detection, while simultaneously reducing causal ratings (see Tarpy et al., 1986), demonstrates the separability of the facilitating and discriminative function of delay cues. These findings mirror those seen in nonhuman conditioning, and they are consistent with the discriminative function view (Ferster, 1953; Pliskoff et al., 1964).

Beyond considerations regarding the mechanisms of stimuli signaling an outcome delay in human causal judgment procedures, the current results may suggest that the discriminative function view (Dinsmoor, 1983; Ferster, 1953) may provide a workable hypothesis to explain conditioned reinforcement effects for humans. Although this view has been challenged for nonhumans, the current results suggest that a stimulus, which is not the target outcome, only appears to gain control over human behavior when it also serves a discriminative function. Such a view has been suggested in other forms in the context of human causal judgments (Buehner & May, 2003; Young, 1995), and this hypothesis may well be worth future investigation to see its generality across other human learning paradigms.

It is also worth noting that participants did not differentiate between the components of the first part of the trial in terms of their overall response rates. That is, in all experiments, reported here, response rates to the RI 10 s and Ext component were highly similar to one another. This was because the participants did not know, on any given trial, which colored stimulus would be associated with the RI 10 s schedule. It could be argued that, in some studies involving the delay stimulus, the participants were processing the task in terms of the relationships between the sample and delay stimulus periods. For instance, they could have learned the rules: "if the 'Xs appear in orange,' then choose red; but, if they appear in in blue, choose green. This would suggest that the choices, and ratings, depended not on the responses per se, but to the stimuli and their relationships."

In summary, the current findings extend the range of procedures under which human causal judgments and instrumental responding are influenced by similar factors to one another. They highlight that delay signals may serve a greater variety of functions than conditioned reinforcement and suggest that, in human causal judgments, one such function may be discriminative. This may begin to allow a mechanism to be added to the oft suggested view that information, often verbal, is critical in altering human judgments of causality. The current set of studies also developed a procedure in which two aspects of the judgment task—both "causal ratings" and "accuracy"—can be assessed. How these two measures might covary and/or whether they reflect distinct processes from one another is an issue that can be studies in future research and may help to illuminate the process underlying human causal judgments more fully.

References

Allan, L. G. (1993). Human contingency judgments: Rule based or associative? *Psychological Bulletin*, *114*(3), 435–448. https://doi.org/10.1037/0033-2909.114.3.435

Buehner, M. J. (2005). Contiguity and covariation in human causal inference. *Animal Learning & Behavior*, *33*(2), 230–238. https://doi.org/10.3758/BF03196065

Buehner, M. J., & May, J. (2003). Rethinking temporal contiguity and the judgement of causality: Effects of prior knowledge, experience, and reinforcement procedure. *The Quarterly Journal of Experimental Psychology Section A*, *56A*(5), 865–890. https://doi.org/10.1080/02724980244000675

Dickinson, A., & Balleine, B. W. (2000). Causal cognition and goal-directed action. In C. Heyes & L. Huber (Eds.), *The evolution of cognition* (pp. 185–204). The MIT Press.

Dinsmoor, J. A. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Sciences*, *6*(04), 693–728. https://doi.org/10.1017/S0140525X00017969

Ferster, C. B. (1953). Sustained behavior under delayed reinforcement. *Journal of Experimental Psychology*, *45*(4), 218–224. https://doi.org/10.1037/h0062158

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. Appleton.

Fienberg, S. E. (2006). When did Bayesian inference become "Bayesian"? *Bayesian Analysis*, *1*(1), 1–40. https://doi.org/10.1214/06-BA101

Greville, W. J., & Buehner, M. J. (2010). Temporal predictability facilitates causal learning. *Journal of Experimental Psychology: General*, *139*(4), 756–771. https://doi.org/10.1037/a0020976

Grindle, C. F., & Remington, B. (2005). Teaching children with autism when reward is delayed. The effects of two kinds of marking stimuli. *Journal of Autism and Developmental Disorders*, *35*(6), 839–850. https://doi.org/10.1007/s10803-005-0029-2

Kaplan, P. S., & Hearst, E. (1982). Bridging temporal gaps between CS and US in autoshaping: Insertion of other stimuli before, during, and after CS. *Journal of Experimental Psychology: Animal Behavior Processes*, *8*(2), 187–203. https://doi.org/10.1037/0097-7403.8.2.187

Keely, J. P. (1999). *Assessing the discriminative properties of response-reinforcer relations using concurrent schedules of reinforcement*. West Virginia University.

Kelleher, R. T., & Gollub, L. R. (1962). A review of positive conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, *5*(S4), 543–597. https://doi.org/10.1901/jeab.1962.5-s543

Keller, F. S., & Schoenfeld, W. N. (1950/2014). *Principles of psychology: A systematic text in the science of behavior*. BF Skinner Foundation.

Killeen, P. R. (1978). Superstition: A matter of bias, not detectability. *Science*, *199*(4324), 88–90. https://doi.org/10.1126/science.199.4324.88

Kuroda, T. (2012). *Discrimination of temporally remote causal relations by pigeons: Effects of signals that mediate the temporal gaps*. West Virginia University.

Lattal, K. A. (1975). Reinforcement contingencies as discriminative stimuli. *Journal of the Experimental Analysis of Behavior*, *23*(2), 241–246. https://doi.org/10.1901/jeab.1975.23-241

Lattal, K. A. (1984). Signal functions in delayed reinforcement. *Journal of the Experimental Analysis of Behavior*, *42*(2), 239–253. https://doi.org/10.1901/jeab.1984.42-239

Lieberman, D. A., Davidson, F. H., & Thomas, G. V. (1985). Marking in pigeons: The role of memory in delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(4), 611–624. https://doi.org/10.1037/0097-7403.11.4.611

Lieberman, D. A., McIntosh, D. C., & Thomas, G. V. (1979). Learning when reward is delayed: A marking hypothesis. *Journal of Experimental Psychology: Animal Behavior Processes*, *5*(3), 224–242. https://doi.org/10.1037/0097-7403.5.3.224

McCoy, J. F., Parker, F. C., & McGee, G. (1976). Signalled reinforcement and dimensional stimulus control. *The Journal of Psychology*, *93*(1), 121–128. https://doi.org/10.1080/00223980.1976.9921382

Michotte, A. E. (1963/1946). *The perception of causality* (T. R. Miles, Trans.). Methuen.

Neuringer, A. J., & Chung, S. H. (1967). Quasi-reinforcement: Control of responding by a percentage-reinforcement schedule. *Journal of the Experimental Analysis of Behavior*, *10*(1), 45–54. https://doi.org/10.1901/jeab.1967.10-45

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex*. Oxford University Press.

Pliskoff, S. S., Hawkins, T. D., & Wright, J. E. (1964). Some observations on the discriminative stimulus hypothesis and rewarding electrical stimulation of the brain. *The Psychological Record*, *14*(2), 179–184. https://doi.org/10.1007/BF03393576

Rachlin, H. (1976). *Behavior and learning*. W. H. Freeman.

Reed, P. (1989). Marking effects in instrumental performance on DRH schedules. *Quarterly Journal of Experimental Psychology*, *41*(4-B), 337–353. https://psycnet.apa.org/record/1990-11578-001

Reed, P. (1992a). Effect of a signalled delay between an action and outcome on human judgement of causality. *Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology*, *44B*(2), 81–100. https://doi.org/10.1080/02724999208250604

Reed, P. (1992b). Signalled delay of reward: Overshadowing versus sign-tracking explanations. *Learning and Motivation*, *23*(1), 27–42. https://doi.org/10.1016/0023-9690(92)90021-D

Reed, P. (1996). No evidence for blocking in human judgments of causality by stimuli presented during an outcome delay. *Learning and Motivation*, *27*(3), 317–333. https://doi.org/10.1006/lmot.1996.0018

Reed, P. (1998). Absence of backward-scan marking mechanism triggered by salient stimuli in human serial list learning. *Learning and Motivation*, *29*(2), 133–151. https://doi.org/10.1006/lmot.1997.0997

Reed, P. (1999). Role of a stimulus filling an action-outcome delay in human judgments of causal effectiveness. *Journal of Experimental Psychology: Animal Behavior Processes*, *25*(1), 92–102. https://doi.org/10.1037/0097-7403.25.1.92

Reed, P., & Hall, G. (1989). The quasi-reinforcement effect: The influence of brief stimuli uncorrelated with reinforcement on variable ratio schedules. *Learning and Motivation*, *20*(3), 242–261. https://doi.org/10.1016/0023-9690(89)90007-6

Renner, K. E. (1964). Delay of reinforcement: A historical review. *Psychological Bulletin*, *61*(5), 341–361. https://doi.org/10.1037/h0048335

Rescorla, R. A. (1982). Effect of a stimulus intervening between CS and US in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes*, *8*(2), 131–141. https://doi.org/10.1037/0097-7403.8.2.131

Richards, R. W., & Marcattilio, A. J. (1978). Stimulus control and delayed reinforcement. *Learning and Motivation*, *9*(1), 54–68. https://doi.org/10.1016/0023-9690(78)90026-7

Royalty, P., Williams, B. A., & Fantino, E. (1987). Effects of delayed conditioned reinforcement in chain schedules. *Journal of the Experimental Analysis of Behavior*, *47*(1), 41–56. https://doi.org/10.1901/jeab.1987.47-41

Schaal, D. W., & Branch, M. N. (1988). Responding of pigeons under variable-interval schedules of unsignaled, briefly signaled, and completely signaled delays to reinforcement. *Journal of the Experimental Analysis of Behavior*, *50*(1), 33–54. https://doi.org/10.1901/jeab.1988.50-33

Schachtman, T. R., & Reilly, S. (1987). The role of local context in autoshaping. *Learning and Motivation*, *18*(4), 343–355. https://doi.org/10.1016/0023-9690(87)90002-6

Spence, K. W. (1947). The role of secondary reinforcement in delayed reward learning. *Psychological Review*, *54*(1), 1–8. https://doi.org/10.1037/h0056533

Staddon, J. E. R. (1983/2016). *Adaptive behavior and learning*. Cambridge University Press.

Tarpy, R. M., Claire-Smith, R. S., & Roberts, J. E. (1986). The effect of informational stimuli on instrumental response rate: Signalling reward versus signalling the availability of reward. *The Quarterly Journal of Experimental Psychology Section B*, *38*(2b), 173–189. https://doi.org/10.1080/14640748608402228

Tarpy, R. M., & Sawabini, F. L. (1974). Reinforcement delay: A selective review of the last decade. *Psychological Bulletin*, *81*(12), 984–997. https://doi.org/10.1037/h0037428

Thomas, G. V., Robertson, D., Cunniffe, G., & Lieberman, D. A. (1989). Facilitation of responding in a filled-delay trace autoshaping procedure: An occasion-setting hypothesis. *Learning and Motivation*, *20*(3), 225–241. https://doi.org/10.1016/0023-9690(89)90006-4

van Elk, M., Salomon, R., Kannape, O., & Blanke, O. (2014). Suppression of the N1 auditory evoked potential for sounds generated by the upper and lower limbs. *Biological Psychology*, *102*, 108–117. https://doi.org/10.1016/j.biopsycho.2014.06.007

Williams, B. A. (1991). Marking and bridging versus conditioned reinforcement. *Animal Learning & Behavior*, *19*(3), 264–269. https://doi.org/10.3758/BF03197885

Williams, B. A. (1994). Conditioned reinforcement: Experimental and theoretical issues. *Behavior Analyst*, *17*(2), 261–285. https://doi.org/10.1007/BF03392675

Wyckoff, L. B., Jr. (1952). The role of observing responses in discrimination learning. Part I. *Psychological Review*, *59*(6), 431–442. https://doi.org/10.1037/h0053932Young, M. E. (1995). On the origin of personal causal theories. *Psychonomic Bulletin & Review*, *2*(1), 83–104. https://doi.org/10.3758/BF03214413

Young, M. E., & Falmier, O. (2008). Color change as a causal agent revisited. *The American Journal of Psychology*, *121*(1), 129–156. https://doi.org/10.2307/20445447

Young, M. E., & Nguyen, N. (2009). The problem of delayed causation in a video game: Constant, varied, and filled delays. *Learning and Motivation*, *40*(3), 298–312. https://doi.org/10.1016/j.lmot.2009.02.002

Received July 27, 2022
Revision received May 1, 2024
Accepted May 2, 2024 ■