

Multiple determinants of transfer of evaluative function after conditioning with free-operant schedules of reinforcement

CHARLOTTE DACK, PHIL REED, AND LOUISE MCHUGH

Swansea University, Swansea, Wales

The aim of the four present experiments was to explore how different schedules of reinforcement influence schedule-induced behavior, their impact on evaluative ratings given to conditioned stimuli associated with each schedule through evaluative conditioning, and the transfer of these evaluations through derived stimulus networks. Experiment 1 compared two contrasting response reinforcement rules (variable ratio [VR], variable interval [VI]). Experiment 2 varied the response to reinforcement rule between two schedules but equated the outcome to response rate (differential reinforcement of high rate [DRH] vs. VR). Experiment 3 compared molar and molecular aspects of contingencies of reinforcement (tandem VIVR vs. tandem VRVI). Finally, Experiment 4 employed schedules that induced low rates of responding to determine whether, under these circumstances, responses were more sensitive to the molecular aspects of a schedule (differential reinforcement of low rate [DRL] vs. VI). The findings suggest that the transfer of evaluative functions is determined mainly by differences in response rate between the schedules and the molar aspects of the schedules. However, when neither schedule was based on a strong response reinforcement rule, the transfer of evaluative judgments came under the control of the molecular aspects of the schedule.

The processes underlying how people learn to like or dislike certain products (e.g., Barnes-Holmes, Keane, Barnes-Holmes, & Smeets, 2000; Foxall, 1997; Foxall & Greenley, 1999; Foxall & Yani-de-Soriano, 2005; Smeets & Barnes-Holmes, 2003; Stuart, Shimp, & Engle, 1987), people (Hebl & Mannix, 2003; Walther, Nagengast, & Trasselli, 2005), and events (Dack, McHugh, & Reed, 2009) are important to understand, since preferences play an enormous role in all aspects of life. One approach to understanding the processes behind the development of preferences suggests that their acquisition emerges through associative learning, or conditioning, whereby a neutral stimulus and a preference become associated with one another. This transfer of valence through an association (i.e., the degree to which an individual likes or dislikes a stimulus) is commonly referred to as *evaluative conditioning* (Levey & Martin, 1975; see De Houwer, Thomas, & Baeyens, 2001, for a review of evaluative conditioning).

Evaluative conditioning occurs when the valence of a novel *target* (the *conditioned stimulus* [CS]) becomes altered by the valence of stimuli/events (the *unconditioned stimulus* [US]) that are presented with, or are dependent on, the CS. This phenomenon is often shown using the picture–picture paradigm. In a typical experiment, a picture that was previously rated as neutral is repeatedly paired with a picture that had been rated as pleasant (e.g., Baeyens, Eelen, Crombez, & Van den Bergh, 1992). As a consequence, the affect of the previously neutral picture

changes in the direction of the valence of the picture with which it was paired (e.g., its ratings will change in the direction of pleasantness). Evaluative conditioning may also play a role in the development of several clinical disorders (Hermans, 1998; Olatunji, 2006; Olatunji, Lohr, Smits, Sawchuck, & Patten, 2009; Schienle, Stark, & Vaitl, 2001). For example, in depression and anxiety disorders, certain types of negative evaluations, such as causal effectiveness and self-efficacy, are said to be characteristic of these conditions (Abramson, Metalsky, & Alloy, 1989; Beck, 1967; Beck, Rush, Shaw, & Emery, 1979; Clark, Beck, & Alford, 1999; Dohr, Rush, & Bernstein, 1989; Loeb, Feshbach, Beck, & Wolf, 1964; MacLeod & Byrne, 1996; Roth & Rehm, 1980; Seligman, 1975) and may be acquired through evaluative conditioning processes (see Fulcher, Mathews, Mackintosh, & Law, 2001; Walther et al., 2005).

However, a feature of many clinical disorders is the extent to which such evaluations generalize across many circumstances—for example, depressive attributional style (Walther et al., 2005) and generalized anxiety (Morgan & Banerjee, 2006). Potentially these negative expectations could lead to maladaptive behaviors such as social withdrawal, avoiding particular activities, and functional impairments (Simon, 2003). It is unlikely that similar negative evaluative judgments would be directly conditioned across such a wide range of circumstances, and the question remains regarding how such evaluations

L. McHugh, l.mchugh@swansea.ac.uk

are produced and generalized across so many aspects of a person's life. For example, previous models of emotional disorders (Beck, 1967, 1976) have not focused on the factors that produce negative evaluations and biases in the first instance, and conditioning and learning models (Fester, 1973; Lewinsohn, 1974; Mowrer, 1960) have failed to adequately explain the complexities of symptoms observed in these disorders or how evaluations can generalize to many stimuli without a direct conditioning event. Evidence for the indirect effect of learning was shown in a set of studies by Rozin, Markwith, and McCauley (1994). They demonstrated that an aversion toward a person with AIDS could be generalized to objects associated with this person (sweater, car, bed, etc.), even though it was perfectly clear that the objects to which the aversion generalized could not be affected.

Similarly, Walther (2002) showed that an attitude (positive or negative) toward one stimulus can transfer to stimuli that had been preassociated with that stimulus. In the baseline phase of this study, participants were asked to rate a number of pictures of male faces for their likeability. In the preconditioning phase, two neutral faces (N1, CS) were paired together. In the conditioning phase, one of the neutral faces (CS) was paired with a positively evaluated picture (Experiment 1) or a negatively evaluated picture (Experiment 2). In the test phase, the participants were asked to rate all the visual stimuli again for their likeability. The results showed not only that ratings of the directly conditioned stimulus increased (Experiment 1) or decreased (Experiment 2) in likeability, but also that this attitude transferred to the neutral face (N1) that was preassociated with the CS. The *spreading attitude* effect was also replicated using a second-order conditioning paradigm (Walther, 2002, Experiment 4), where the preconditioning and conditioning phases were reversed.

An additional explanation of the development of generalized evaluations and attitudes is produced through examining derived stimulus relations in equivalence classes, an account that suggests that stimuli can become associated with one another despite never having been directly paired and despite their lack of any shared physical properties (Sidman, 1971). The potential breadth of such classes could allow for widespread transfer of particular associations learned about one member of that class.

Equivalence classes can be established by training a minimal number of relations between individual stimuli in a group. For example, if a group of stimuli consisted of the letters A, B, and C, an equivalence class could be established by training two-term relations in-between: AB and BC. If a class has been established, many new emergent relations, which had not been taught directly, are formed between the stimuli (reflexive [i.e., $A \rightarrow A$]; symmetrical [i.e., $B \rightarrow A$]; transitive [i.e., $A \rightarrow C$]; and equivalence [i.e., $C \rightarrow A$]; see Bush, Sidman, & de Rose, 1989). If all of the emergent relations control responding, the group of stimuli can be said to function as an equivalence class (Sidman, Kirk, & Wilson-Morris, 1985) and the stimuli are fully substitutable for one another (Sidman, 1990, 1994).

Dack et al. (2009) investigated the manner in which various schedules of reinforcement are associated with

differential evaluations of causal efficacy (high vs. low ratings) and whether these ratings transfer to previously neutral stimuli (CSs) paired with schedules of reinforcement in an evaluative learning paradigm and to unrelated stimuli through derived equivalence relations. Participants were exposed to a number of trials of two alternating schedules of reinforcement (Experiment 1, differential reinforcement of high rate [DRH] and differential reinforcement of low rate [DRL] schedules; Experiment 2, DRH and variable interval [VI] schedules; Experiment 3, VI and variable ratio [VR] schedules) and were asked to rate how effective they felt their responses on a space bar were at producing outcomes/points on the computer screen (i.e., their causal efficacy) at the end of each trial. A CS was paired with each schedule (i.e., it appeared on the computer screen while the participants were pressing the space bar to earn points). The participants were also trained to learn 2 three-member equivalence classes, each containing two nonsense words and one affective word (the word *useless* in Class 1 and the word *good* in Class 2). The participants were finally asked to categorize each class member with one of the CS associated with each schedule, as well as to rate each schedule CS in terms of *good* or *useless*.

The experiments replicated previous research (Reed, 1994, 1999, 2001a, 2001b, 2003) in which causal efficacy ratings mirrored the pattern of rates of response, even when the rates of reinforcement (Experiment 2) and the probability of an outcome given a response (Experiment 3) were equated. That is, when response rates were high, ratings of causal efficacy were also high, and when response rates were low, ratings of causal efficacy were low. The results of all three experiments showed a general trend of evaluative learning in the schedules task. The conditioned stimuli paired with each schedule were shown to acquire the valence of the participants' causal efficacy ratings. The Class 1 members (*useless*) were categorized with the stimuli paired with the schedule that had lower causal efficacy ratings and lower response rates associated with it, and the Class 2 members (*good*) were categorized with the stimuli paired with the schedule that had higher causal efficacy ratings and higher rates of response associated with it. The participation of new stimuli (the schedule CS) in derived stimulus networks could be a possible mechanism behind the acquisition of evaluative judgments without direct experience of the event.

Despite these findings (Dack et al., 2009), it is still not clear which aspects of the schedule contingencies are responsible for controlling the efficacy ratings or their subsequent transfer. Dack et al. assumed that the transfer of stimulus function followed the ratings of causal efficacy that the schedule produced—an assumption strengthened by the fact that aspects of the contingencies, like the probability of an outcome given a response (Experiment 3), were controlled. However, schedules of reinforcement can differ from one another in many ways (e.g., the rule that relates responding to reinforcement, the reinforcement rate, and the response rate, as well as the probability of an outcome given a response), and it is not clear precisely which aspects of the schedules determine the evaluative function of the related stimuli, since these ratings of causal efficacy

have previously been found to covary with other aspects of the schedule (cf. Alloy & Abramson, 1979; Alloy & Tabachnick, 1984; Reed, 1993, 1999, 2001a, 2001b; Wasserman, Elek, Chatlosh, & Baker, 1993).

Previous research has examined how some of the factors mentioned above might produce certain efficacy ratings. Allan (1980) and Wasserman et al. (1993) have found, in contexts in which an outcome did not serve as a reinforcer (i.e., a triangle flash), that when a schedule's probability of an outcome given no response was kept constant, decreases in the probability of an outcome given a response (and, therefore, increases in the number of responses per outcome) led to lower ratings of efficacy. At higher probabilities of an outcome given a response, one response is likely to produce an outcome, which may lead to higher ratings of effectiveness. Schedules with a greater frequency of outcomes have also been found to be associated with higher ratings of causal effectiveness (Alloy & Abramson, 1979; Wasserman, 1990). These findings are very robust and have led to the development of many accounts of causal efficacy judgments specifically based on these probabilities (e.g., Allan, 1980). It seems likely that, under some circumstances, such probabilities will exert an effect on causal efficacy judgments.

In the context of situations in which the outcome may serve as a reinforcer (i.e., the delivery of points or hypothetical or real money), schedules that typically reinforce high rates of responding have been found to be associated with higher ratings of efficacy, as compared with schedules that reinforce low rates of responding, even when the rate of reinforcement (e.g., Reed, 2001a, 2001b) and probability of an outcome given a response (see, e.g., Dack et al., 2009) were equated across schedules. Two views have been put forward as explanations for this schedule-controlled behavior, and these suggest two further factors that may control the impact of schedules on judgments of causal efficacy.

The *molar* view (e.g., Baum, 1973) suggests that participants are sensitive to the overall relationship between responding and the rate of reinforcement and that this may lead to strong ratings of causal effectiveness (see Reed, 2001a, 2001b). In contrast, the *molecular* view stresses the importance of the pattern of responding emitted just prior to reinforcement (e.g., Morse, 1966; Peele, Casey, & Silberberg, 1984). For example, certain schedules differentially reinforce long interresponse times (IRTs) (e.g., VI schedules and DRL schedules). In contrast, on other schedules (e.g., VR schedules and DRH schedules), there is no such favoring of long IRTs, and reinforcement is more likely to occur after a short IRT than after a long IRT. It may be that the local context of responding at the time of reinforcement is a determinant of causal judgments. Some research (e.g., Reed, 2001a, 2001b) has shown that high ratings of causal efficacy are associated with schedules that typically reinforce high rates of responding just before reinforcement (e.g., VR and DRH schedules), and low ratings of causal efficacy are associated with schedules that typically reinforce low rates of responding just before reinforcement (e.g., VI and DRL schedules).

The aim of the present series of experiments was to extend the research of Dack et al. (2009) by exploring the various factors that may be involved in determining the evaluative ratings of the stimuli associated with each schedule of reinforcement and the subsequent transfer of this evaluative conditioning to unrelated stimuli through derived equivalence ratings. This was investigated by employing a range of schedules and yoking procedures that allowed these factors to be varied. However, it is important to note that these factors (response rate, reinforcement rate, etc.) are not easy to control, since schedules of reinforcement, by their nature, are *free operant*, and it is difficult to exert complete control over all factors simultaneously (see Jenkins, 1970, for a discussion). In spite of this difficulty, previous research (Dack et al., 2009) on this topic did employ such schedules of reinforcement, while controlling for rate of reinforcement (Experiment 2) and outcome probability (Experiment 3). However, all the schedules employed in this research produced different rates of response and had different response reinforcement rules. It was therefore thought to be important to explore more fully these effects by employing schedule manipulations previously used to examine the factors above (see Reed, 2001a, 2001b) and, in particular, to attempt to equate the response rate and the response reinforcement rule between schedules to investigate the impact of the evaluative ratings of the CS and its categorization. This was the primary aim of the present series of experiments.

A second aim was to address a potentially trivial explanation of the results found in the previous experiments (Dack et al., 2009). In the report by Dack et al., it is possible that the data may have been influenced by an interaction between the experimental procedure and possible task demands that were brought about by the use of similar labeling in the efficacy judgments made after each exposure to the schedule (e.g., *ineffective*, *effective*) and in the establishment of the equivalence classes (i.e., each class contained either the word *good* or *useless*). This similarity in the words employed in the two parts of the task may, in turn, have led to some semantic ambiguity in how the participants categorized the stimuli in the transfer-of-conditioning test, and, therefore, the results may not have reflected evaluative transfer based on the establishment of a stimulus network alone. This problem can be addressed by omitting the judgments of causal efficacy after each schedule trial. If transfer can occur spontaneously after experience of the different schedules, this would be evidence of genuine transfer.

EXPERIMENT 1

In the first experiment, we attempted to explore the impact of schedules of reinforcement on evaluative ratings of associated stimuli by employing VR 20 and VI schedules that were yoked for rate of reinforcement. These schedules might be expected to produce differences in response rate, despite being equated for rate of reinforcement, with there being higher rates of response for the VR than for the VI schedule (Baum, 1993; Peele et al., 1984; Reed, 2001a, 2001b). Dack et al. (2009) employed the same

schedules (Experiment 3) but used a VI 20 schedule and a VR schedule that was yoked in terms of the probability of an outcome given a response. Experiment 1 was designed to extend this research to determine which aspect of these schedules (e.g., the response reinforcement rule or the probability of an outcome given a response) in the absence of reinforcement rate differences (rather than in the absence of outcome probability difference; Dack et al., 2009, Experiment 3) drives the generation of evaluative judgments and the transfer of such evaluative function. Previous research (Reed, 2001a, 2001b) has shown that ratings of causal efficacy mirror the patterns of responding; that is, higher ratings of causal efficacy are given on VR, as compared with VI, schedules, despite the former schedule often having much lower outcome per response probabilities (since it generates a higher rate of response). Given this, it might be expected that there would be a positive transfer to the CS related to the VR schedule and a negative transfer to the CS related to the VI schedule. If a VR schedule generated a higher response rate than did a VI schedule, even when the rate of reinforcement was equated between the schedules, and such a pattern of transfer emerged despite the VR schedule having a lower probability of an outcome given a response, it would be evidence for the importance of the rate of response (and therefore, for the factors that control this aspect of schedule performance) in generating evaluative judgments and their transfer.

Method

Participants

Fourteen adults participated in Experiment 1 (1 of them male, 13 female). The ages of the participants ranged from 18 to 20 years, with a mean of 19.14 years ($SD = 0.66$). All the participants were undergraduates at Swansea University and were recruited by advertisements in the Psychology department. In return for their participation, they earned subject pool credit. Each participant had normal or corrected-to-normal eyesight and did not suffer from color blindness. All were naive as to the purpose of the experiment. The experiment was approved by the Department of Psychology, Swansea University, Ethics Committee.

Apparatus and Materials

The experiment was conducted in a quiet room free from distraction. The experimental room contained only a desk, a chair, and a personal computer with a 550-MHz processor, a 35-cm color monitor, and a standard computer mouse. All trial presentations and response recordings were controlled by programs written in Visual Basic 6.0. All responses involved either mouse clicking or the space bar key.

Derived stimulus relations training. The stimuli used as the members of the two equivalence classes were nonsense words and affective words, as shown in Table 1, along with their corresponding letter-number designation. All the stimuli were composed of Arial Bold characters in black, each of which occupied a certain

proportion of the screen (screen width/10 cm). Each stimulus was surrounded by a box, its width equaling 10 cm and its height equaling 2.5 cm, against a white background. The positioning of the comparison stimuli was as follows. The horizontal position of the left stimulus started 3.125 cm from the left edge, and the horizontal position of the right stimulus started 3.125 cm from the right edge of the screen. Their vertical positions were 19.375 cm down the screen. All feedback choices were in red surrounded by a box that was 15 cm wide and 3.75 cm in height, presented in the middle of the screen.

Evaluative conditioning task. Two differently colored circles were used to signal the two schedule conditions (VR and VI) and acted as the CSs. Three different combinations of colors were used (yellow/blue, blue/white, yellow/white), and these were counterbalanced across the schedules (2 participants had white [VR schedule] and blue [VI schedule], 2 participants had blue [VR] and white [VI], 2 participants had blue [VR] and yellow [VI], 3 participants had yellow [VR] and blue [VI], 2 participants had yellow [VR] and white [VI], and 3 participants had white [VR] and yellow [VI]), so that each participant had one of these color permutations for the stimuli used as the CSs for the VR and VI schedules of reinforcement. This was done to ensure that the postevaluative ratings of the conditioned stimuli were due to the conditioning task, and not to any preexisting preferences toward particular colors. The number of points earned throughout each schedule trial was shown in the corner of the computer screen.

Measures. The Beck's Depression Inventory (BDI; Beck, 1978; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) was administered to assess the level of depression in participants. Any participant scoring highly (i.e., >10) on this measure was excluded from the data analysis, as in previous studies (Dack et al., 2009; Reed & Antonova, 2007; Reed, Frasquillo, Colkin, Liemann, & Colbert, 2001). This was to ensure that the different schedules of reinforcement, and not any preexisting depressive condition, were responsible for the evaluative effects reported in this experiment.

A short questionnaire was designed to allow the participants to rate the evaluative function of a collection of stimuli, including the nonsense syllables used as members of the two equivalence classes and the CSs present during the conditioning task. For each stimulus, there was a 7-point scale: 1 being *useless*, 4 being *neutral*, and 7 being *good*. From this scale, the participants' evaluations for each item could be measured and converted into a number to test for any transfer of function.

A consent form was constructed to inform the participants of the aim of the experiment and to assure them of the confidentiality of the results.

Procedure

Each participant was taken into a quiet room and was given a consent form to read and sign and the BDI to answer before completing the experiment.

Derived stimulus relations training. All training and testing were conducted using a 2×3 match-to-sample, conditional discrimination paradigm, using stimuli that consisted of two nonsense words and an affective word, as shown in Table 1. Two three-member equivalence classes were established by training AB and BC relations in a linear/series structure. Each relation (A1–B1, A2–B2, B1–C1, and B2–C2) was presented at least three times during training in a quasi-random order in blocks of 12 relations. The criterion to proceed to the testing phase was 12 consecutively correct trials across all stages. There was no time limit for responding to individual trials.

Each trial started with the presentation of a sample (Sa) and two comparison stimuli. The positive comparison (Co+) stimulus was chosen from the same equivalence class as the Sa, and the negative comparison (Co–) stimulus was chosen from the other class. The stimuli were displayed in an isosceles triangle display on the monitor, with Sa at the vertex of the triangle and Co+ and Co– at the corners of the base. The following instructions were presented across the middle of the screen on the first trial only:

Table 1
Nonsense and Affective Words Used As Stimuli and
Their Assignment to Equivalence Classes

	A	B	C
Class 1	lewoly	gedeer	useless
Class 2	matser	rigund	good

Look at the Box Above and then Click on the Box Below that GOES WITH the one at the Top. Try Your Best NOT to Make Any Mistakes.

The participants chose a comparison by clicking on the left- or right-hand box containing the stimuli. When feedback was provided during training, choosing the Co+ produced a 1-sec display of the word "Correct," whereas choosing the Co- produced a 1-sec display of the word "Wrong." Both words were displayed in red across the middle of the computer screen.

Testing emergent relations. Once the criterion for the training session had been met, the test phase commenced, and the corrective feedback was terminated. On the first test trial, the following instructions were shown across the middle of the computer screen:

Look at the Box Above and then Click on the Box Below that GOES WITH the one at the Top. Try Your Best NOT to Make Any Mistakes. DURING THESE TRIALS THE COMPUTER WILL NOT GIVE YOU ANY FEEDBACK.

All tests for one-node transitivity ($A \rightarrow C$) and equivalence ($C \rightarrow A$) were presented in a single block. Each type of relation ($A1-C1$, $A2-C2$, $C1-A1$, and $C2-A2$) was presented nine times, with 36 trials in total. The mastery criterion for testing was at least 30 correct class-consistent selections across the block of 36 test trials, with no more than 2 incorrect trials of any one relation. If a participant failed to reach this criterion, he or she received further training and testing phases. If, after three training phases, a participant still did not meet the testing criterion on the third test phase, he or she was told that the experiment had finished, was debriefed by the experimenter, and was withdrawn from the experiment.

Evaluative conditioning task. The participants were presented with the following instructions on the computer screen:

You will shortly see a circle in the centre of the screen. The circle will be colored either [Color 1] or [Color 2].¹ When you see this circle you must press the space bar to earn points. Try to earn as many points as possible. You must work out what rate of pressing makes you the most points. Click here to proceed.

After presentation of these instructions, the participants were exposed to eight schedule trials, which consisted of two alternating schedules (VR and VI), with the VR schedule always presented first. Each schedule was, therefore, presented four times to each participant. For one schedule (VR), the participants responded on a VR 20 schedule. The participants were required to emit, on average, 20 responses in order to produce an outcome (i.e., receive a point). For a second schedule, the participants responded on a VI schedule, yoked to the temporal distribution of outcomes obtained in the previous VR condition for each participant separately. The yoking was done by dividing the total number of outcomes (points) per trial on the VR schedule by the total time of each trial (120 sec). This number was then used as the overall mean time interval between outcomes (points) for the yoked VI schedule. Points were received after a time interval that varied around this mean value. The VR trials always immediately preceded the VI trials yoked to those VR trials and were presented for 2 min. Each VI schedule trial lasted until the participants had received the same amount of points as with the previous VR schedule.

Transfer-of-conditioning test. The participants were presented with the following instructions on the computer screen:

Look at the Image Above and then Click on the Image Below that GOES WITH the one at the Top.

Each trial began with the presentation of one of the six stimuli from the two equivalence classes, in the middle at the top of the screen, as a sample and with the two colored circles from the evaluative conditioning task, in each corner at the bottom of the screen, as the comparisons. The participants chose one of the colored circles by clicking on it with the mouse, before moving on to the next trial. Each word from the derived stimulus relations training was shown

four times, giving a total of 24 trials. No feedback was presented during this phase of the experiment.

Eight participants received the derived stimulus relations task before the evaluative conditioning task, and the other 6 received the tasks in the opposite order. All the participants completed the transfer-of-conditioning test last. At the end of the experiment, each participant completed the transfer-of-function questionnaire, was thanked for participating, and was debriefed.

Results and Discussion

On the basis of the BDI scores (see Dack et al., 2009, and Reed et al., 2001), 3 participants were excluded from the analysis (BDI scores > 10). One of these participants also failed to reach the criterion of the equivalence test after three training phases. The mean score for the remaining participants on the BDI was 3.91 ($SD = 2.55$).

Derived Stimulus Relations Training and Testing

All the remaining participants passed the emergent relations test, with an overall mean of 95.45% (± 1.56) class-consistent responding. This was statistically significantly more correct responses than would be expected by chance [$t(10) = 29.20$]. Nine participants reached the passing criterion on the first test session. One participant needed a second training phase before reaching the test criterion (i.e., 30/36 correct), and 1 participant needed a third training phase. The mean number of trials received during the training phase(s) was 37.45 (range = 15–75, $SD = 21.59$) before the test criterion was met. The mean number of these trials that were correct was 29.27 (range = 14–57, $SD = 15.30$).

Evaluative Conditioning Task

The mean rates of responses were collapsed across the four trials for both schedules. Inspection of these data showed that the response rate was higher for the VR schedule (179.77 responses per minute, $SD = 57.27$), as compared with the VI schedule (123.35 responses per minute, $SD = 55.13$). A repeated measures ANOVA was performed on these data, with schedule (VR and VI) and trial as factors. A rejection criterion of $p < .05$ was adopted for this and all subsequent analyses. This analysis revealed a statistically significant main effect of schedule [$F(1,10) = 9.77$]. There was no statistically significant main effect of trial, nor was there a statistically significant interaction between the two factors ($F_s < 1$).

Since the VI schedule was yoked to the temporal distribution of outcomes obtained for the VR schedule, the number of outcomes was approximately the same for both schedules on each trial. The mean number of outcomes averaged across the four trials was the same for both schedules (16.41, $SD = 5.49$). An ANOVA showed no main effect of trial ($p > .20$). For all the participants, each VI schedule trial lasted until they had earned the same amount of points as on the yoked VR schedule trial. This meant that the VI trials were sometimes longer or shorter than 120 sec (the length of the VR trials). To ensure that the rate of reinforcement was still equated across the two schedules, the mean interreinforcer interval (IRI) for each schedule was calculated. The mean IRI was slightly longer for the VR

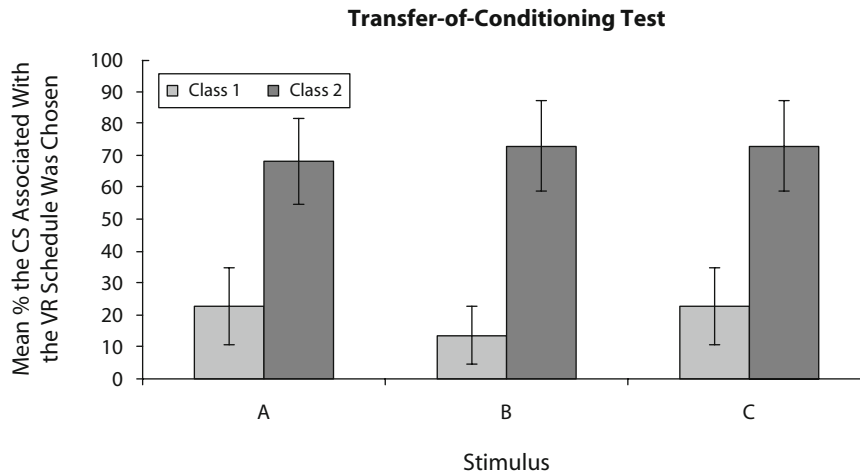


Figure 1. Results from Experiment 1: Mean percentages that the conditioned stimulus (CS) (variable ratio [VR] schedule) was chosen with each class member (Class 1 = *useless*, Class 2 = *good*). Error bars show mean standard errors.

schedule (11.25 sec, $SD = 11.53$) than for the yoked VI schedule (8.22 sec, $SD = 7.50$). This difference was not found to be statistically significant [$t(10) = 2.08$].

The mean probabilities of an outcome given a response were nearly twice as high with the VI schedule (.09, $SD = .003$) as with the VR schedule (.05, $SD = .047$). A repeated measures ANOVA (schedule \times trial) conducted on these data revealed a statistically significant main effect of schedule [$F(1,10) = 8.35$]. There was no statistically significant effect of trial and no statistically significant interaction between trial and schedule ($F_s < 1$).

Transfer-of-Conditioning Test

Figure 1 reveals that the participants chose the stimulus associated with the VR schedule that had a higher rate of response (but a lower probability of an outcome given a response) in the presence of the stimuli participating in the *good* equivalence class (Class 2), more than in the presence

of the stimuli participating in the *useless* equivalence class (Class 1). A within-subjects ANOVA (stimuli \times class) was performed on these data and revealed a statistically significant main effect of class [$F(1,10) = 5.35$]. There was no significant main effect of stimuli ($F < 1$) and no statistically significant interaction between the two factors ($p > .27$). Dependent t tests showed that the stimulus associated with the VR schedule was chosen significantly more in the presence of class member A2 as compared with A1 [$t(10) = 3.75$], B2 as compared with B1 [$t(10) = 5.99$], and C2 as compared with C1 [$t(10) = 4.15$].

Transfer-of-Function Questionnaire

The mean ratings (*useless–neutral–good*) for each nonsense stimulus in each class and for the two conditioned stimuli that signaled the different schedules can be seen in Figure 2. Inspection of these data shows that there were lower ratings (i.e., more *useless*) for the members

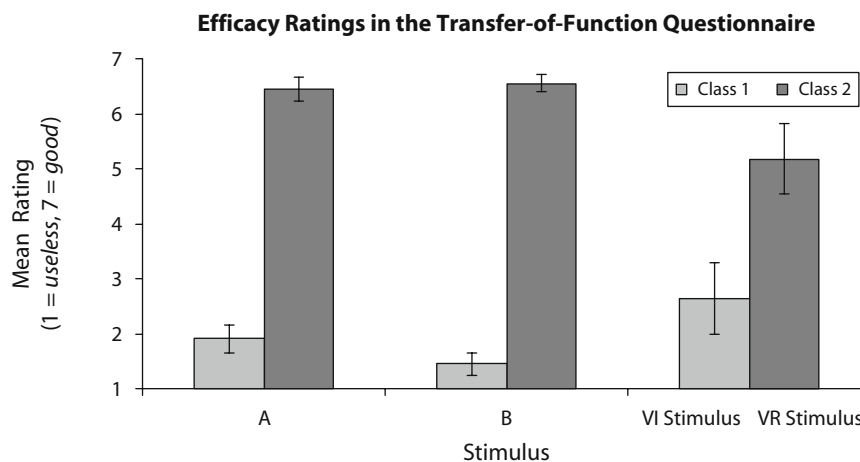


Figure 2. Results from Experiment 1: Mean ratings for the nonsense stimuli (Class 1 associated with *useless*; Class 2 associated with *good*) and for the conditioned stimuli that signaled the two different schedules (variable interval [VI], variable ratio [VR]), with mean standard error bars.

of Class 1 (associated with the term *useless*) than for the members of Class 2 (associated with *good*). To investigate whether the conditioned stimuli (colored circles) that signaled the two separate schedules (VR, VI) had joined the two existing equivalence classes, a two-factor repeated measures ANOVA was performed on the data, with class (1 and 2) and stimulus (A, B, and the CS) as the within-subjects factors. This revealed a statistically significant main effect of class [$F(1,10) = 61.92$], no statistically significant main effect of stimulus ($p > .30$), and no statistically significant interaction between class and stimulus [$F(2,20) = 3.45$].

In summary, response rates were higher with a VR schedule than with a VI schedule with the same frequency of reinforcement. This result mirrors previous findings (e.g., Ferster & Skinner, 1957; Reed, 2001a, 2001b). Under these circumstances, the participants rated the stimulus associated with the VR schedule (which was associated with higher rates of responding) as *good*, and the stimulus associated with the VI schedule as *useless*. In addition, in the transfer-of-conditioning test, the participants were shown to categorize the conditioned stimuli related to the schedules with the stimuli participating in the equivalence classes with the same evaluative function as the rating of the conditioned stimuli (*good*, *useless*). This pattern of data suggests that when two schedules of reinforcement with contrasting response reinforcement rules (VR, VI) are employed, it is response rate (and the factors that may be responsible for this response rate), rather than the rate of reinforcement (since this was equated across schedules), outcome probability (since this was actually significantly higher for the VI schedule than for the VR schedule), or the response reinforcement rule, that drives the transfer of evaluative function.

EXPERIMENT 2

In Experiment 1, when two schedules of reinforcement that differed in terms of their response reinforcement rule (VR vs. VI), but not in terms of their rate of reinforcement, were compared, the response rate and/or the response reinforcement rule were found to be important in determining the transfer of evaluative function to related stimuli. In Experiment 2, we attempted to further explore the factors responsible for this transfer by comparing two schedules that are likely to produce the same rate of response but have different response reinforcement rules. In this way, it was hoped to tease apart whether it is the rate of response or the response rule that drives evaluative conditioning on schedules of reinforcement.

It has been shown that a DRH schedule might have a stronger response-reinforcement-based rule than does a VR schedule (Reed, 1992). This is because the DRH contingency arranges an outcome only if a number of responses have been produced in close proximity. However, it is possible for participants to produce the same rate of response on both schedules (Reed, 2003). If it is the strength of the response reinforcement rule that drives evaluative conditioning and transfer of this function, rather than differences in response rate, it would be expected that the DRH sched-

ule would produce a positive transfer to the CS to which it was related, as compared with the VR schedule.

Method

Participants

Eighteen adults participated in Experiment 2 (2 of them male, 16 female). The ages of the participants ranged from 18 to 52 years, with a mean of 24.94 years (± 8.82). All the participants were either undergraduate or postgraduate students at Swansea University.

Apparatus and Materials

The apparatus and materials used were the same as those in Experiment 1, apart from those in the evaluative conditioning task. The only difference in this task was the different combinations of colored circles that were used to signal the two scheduling conditions. The two schedules were a DRH and a VR schedule. Three different color combinations were used (yellow/white, yellow/blue, and blue/white), and these combinations were counterbalanced across the participants (3 participants had yellow [DRH] and blue [VR], 3 participants had blue [DRH] and yellow [VR], 3 participants had white [DRH] and blue [VR], 3 participants had blue [DRH] and white [VR], 3 participants had yellow [DRH] and white [VR], and 3 participants had white [DRH] and yellow [VR]).

Procedure

Each participant was taken into a quiet room and was given a consent form to read and sign and the BDI to answer before completing the experiment.

Derived stimulus relations training. This was identical to that described in Experiment 1.

Evaluative conditioning task. The participants were presented with the following instructions on the computer screen.

You will shortly see a circle in the centre of the screen. The circle will be colored either [Color 1] or [Color 2] [see note 1]. When you see this circle you must press the space bar to earn points. Try to earn as many points as possible. You must work out what rate of pressing makes you the most points. Click here to proceed.

After presentation of these instructions, the participants were exposed to the schedules, which consisted of two alternating schedules (DRH and VR). Each schedule was presented four times to each participant. The DRH schedule was always presented first and lasted for 3 min. Each VR schedule lasted until the participants had received approximately the same number of points as that on the previous DRH schedule (e.g., the mean number of points per trial were the following: DRH Trial 1 = 51.23 [$SD = 5.48$], VR Trial 1 = 51.38 [$SD = 6.04$], DRH Trial 2 = 52.77 [$SD = 5.56$], VR Trial 2 = 51.69 [$SD = 6.28$], DRH Trial 3 = 53.00 [$SD = 4.73$], VR Trial 3 = 51.77 [$SD = 4.89$], DRH Trial 4 = 55.69 [$SD = 3.40$], VR Trial 4 = 55.15 [$SD = 4.22$]).

For one schedule condition, the participants responded on a DRH 5/2 schedule. The participants were required to emit five responses in 2 sec in order to produce an outcome (receive a point). In a second schedule condition, the participants responded on a VR schedule, yoked to the probability of an outcome given a response obtained on the DRH schedule. This was done by dividing the total number of outcomes (points) per trial on the DRH schedule by the total number of responses. This number was then used as the overall mean probability of receiving an outcome (point) given a response on the yoked VR schedule. Points were received with a probability that varied around this mean value.

Transfer-of-conditioning test. This was identical to that described for Experiment 1.

Half of the participants received the derived stimulus relations task before the evaluative conditioning task, and the remaining participants received the tasks in the reverse order. All the participants completed the transfer-of-conditioning test last. At the

end of the experiment, each participant completed the transfer-of-function questionnaire. They were thanked for participating and were debriefed.

Results and Discussion

Before the analyses were performed, 4 participants were excluded for high scores on the BDI (i.e., >10 ; see Reed et al., 2001). The mean score for the remaining participants on the BDI was $4.08 (\pm 1.75)$. One participant was excluded from all analyses, since the participant withdrew from the study after failing to reach the criterion of the emergent relations test after two training phases.

Derived Stimulus Relations Testing

All the remaining participants passed the emergent relations test with an overall mean of $98.08\% (\pm 0.66)$ class-consistent responding. This was statistically significantly more correct responses than would be expected by chance [$t(12) = 72.94$]. Eleven participants passed the equivalence test after one training phase; 2 participants needed a second training phase before reaching the test criterion (i.e., 30/36 correct) needed to continue. The mean number of trials received during the training phase(s) was 34.54 (range = 12–77, $SD = 20.82$) before the test criterion was met. The mean number of these trials that were correct was 26.54 (range = 12–54, $SD = 12.84$).

Evaluative Conditioning Task

The mean rate of response, averaged across all four trials emitted for the DRH schedule, was 249.41 ($SD = 43.46$), as compared with 247.67 ($SD = 45.96$) for the VR schedule. A two-factor ANOVA (schedule \times trial) was performed on these data and revealed no statistically significant main effect of trial ($p > .10$) or schedule ($F < 1$). There was no statistically significant interaction between trial and schedule ($F < 1$). Thus, as was hoped,

the two schedules produced rates of response similar to one another.

The mean number of outcomes per trial, averaged across all four trials, was highly similar on the DRH schedule (53.17 , $SD = 2.99$), as compared with the VR schedule (52.50 , $SD = 3.40$). However, despite the numerical similarity, a two-factor repeated measures ANOVA (schedule \times trial) revealed a statistically significant effect of schedule [$F(1,12) = 8.27$] but no statistically significant main effect of trial ($p > .13$) and no statistically significant interaction ($p > .08$). It is unclear whether this statistical significance, based on a very small numerical difference between the schedules (less than 2% higher reinforcement rate for the DRH schedule than for the VR schedule), would be discriminable to the participants.

Since the probability of an outcome given a response of the VR trials was yoked to the outcome probability of the DRH trials, this value was necessarily the same for both the DRH (.11) and VR (.11) schedules. An ANOVA (trial) showed no statistically significant main effect of trial [$F(3,36) = 1.16$].

Transfer-of-Conditioning Test

Inspection of Figure 3 reveals only a small difference between the participants' categorization of the stimulus associated with the DRH schedule across the two classes (Class 1, *useless*; Class 2, *good*). A repeated measures ANOVA (class \times stimuli) revealed no statistically significant main effect of class ($F < 1$) or stimuli ($p > .24$), nor was there a significant interaction between the two factors ($F < 1$).

Transfer-of-Function Questionnaire

The mean ratings for each nonsense stimulus in each class and for the conditioned stimuli that signaled the different schedules can be seen in Figure 4. An inspection

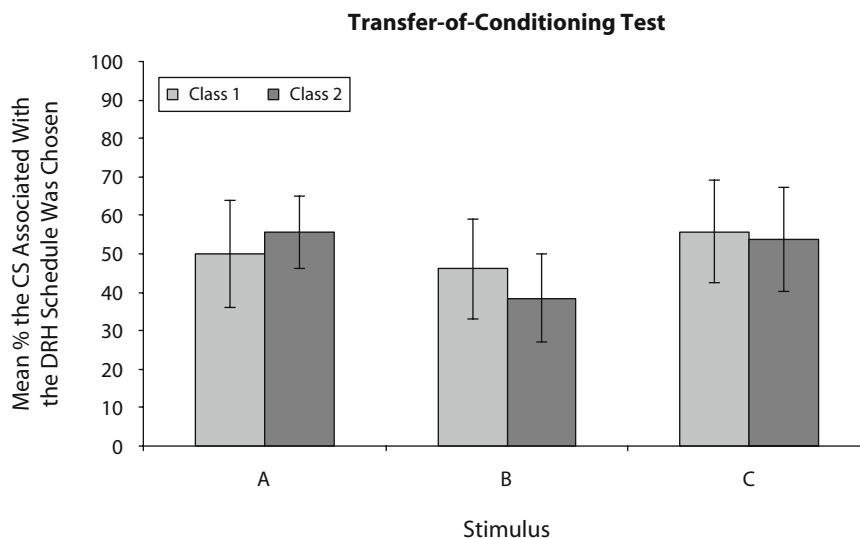


Figure 3. Results from Experiment 2: Mean percentages that the stimulus (colored circle) conditioned with the differential reinforcement of high rate (DRH) schedule was chosen with each stimulus from each equivalence class (Class 1 = *useless*; Class 2 = *good*).

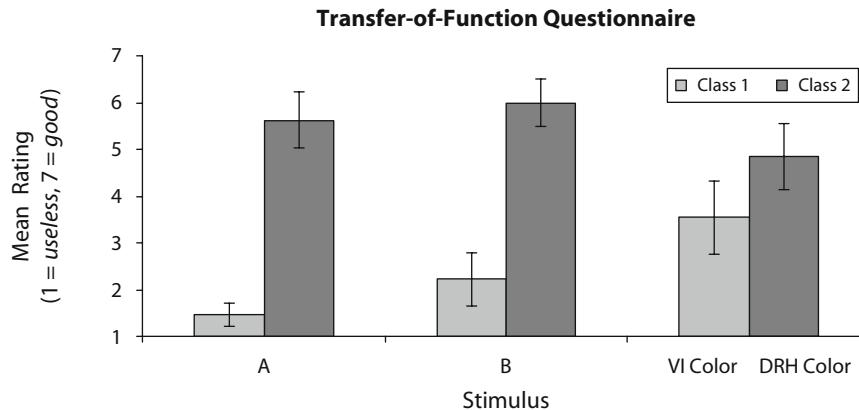


Figure 4. Results from Experiment 2: Mean rating for the nonsense stimuli (Class 1 associated with *useless*; Class 2 associated with *good*) and for the conditioned stimuli that signaled the two different schedules (differential reinforcement of high rate [DRH], variable ratio [VR]), with mean standard error bars.

of these data shows lower ratings (more *useless*) for the members of Class 1 (associated with the term *useless*) than for Class 2 (associated with the term *good*). To investigate whether the conditioned stimuli (colored circles) that signaled the two separate schedules (DRH, VR) had joined the two existing equivalence classes, a two-factor repeated measures ANOVA was performed on the data, with class (1 and 2) and stimulus (A, B, and the CS) as the within-subjects factors. This revealed a statistically significant main effect of class [$F(1,12) = 14.48$] and no statistically significant main effect of stimulus ($p > .30$). The interaction between class and stimulus approached statistical significance ($p > .07$). Simple effect analyses revealed an effect of class for Stimulus A [$F(1,12) = 28.73$] and for B [$F(1,13) = 16.02$], but not for the CS ($F < 1$).

In summary, when two schedules that had a similar rate of response and probability of an outcome given a response were compared with one another, there was very little differential evaluative transfer in the transfer-of-conditioning test, and the CSs were given similar evaluative ratings. These null effects may have occurred because the schedule manipulations and/or the amount of training the participants received for each schedule were insufficient to produce a strong effect on behavior. However, another plausible explanation of these results is that it is the difference in the rate of response between schedules that may be important in determining the evaluative transfer to the stimuli. This replicates the findings of previous research (Reed, 2003), which showed that participants gave higher ratings of causal efficacy to responses associated with a DRH schedule, as compared with a VR schedule, only when there was a significant difference in the rate of response between the two schedules (higher rate of response on the DRH schedule).

To investigate this suggestion further, the data were split into two groups: participants with a higher rate of response on the DRH schedule than on the VR schedule and participants with a higher rate of response on the VR schedule than on the DRH schedule. The differences between how the two groups categorized the DRH condi-

tioned stimuli can be seen in Figure 5. An inspection of Figure 5 shows that the participants who responded more on the DRH schedule categorized the DRH CS with the Class 2 stimuli (i.e., the *good* class) more than with the Class 1 stimuli. The opposite pattern was found within the participants who responded more on the VR schedule.

A three-way mixed ANOVA (group \times class \times stimuli) was performed on the data. This showed no statistically significant main effect of group [$F(1,11) = 4.09$, $p = .07$], class ($F < 1$), or stimuli ($p > .18$). There were also no statistically significant interactions between stimuli and group ($p > .25$), class and stimuli ($F < 1$), or group, class, and stimuli ($F < 1$). However, as can be predicted from this suggestion, there was an interaction between class and group that was significant [$F(2,22) = 7.85$].

To investigate the significant interaction between class and group, two separate one-factor ANOVAs (class) were performed for the VR higher rate of response group and for the DRH higher rate of response group. For the DRH higher rate of response group, there was a statistically significant main effect of class [$F(1,4) = 12.28$]. Dependent t tests revealed that the participants in this group chose the DRH stimulus with the Class 2 (*good*) stimuli significantly more than those with the Class 1 (*useless*) stimuli [A1–A2, $t(4) = 6.94$; B1–B2, $t(4) = 4.78$; C1–C2, $t(4) = 7.62$]. For the VR higher rate of response group, there was no statistically significant main effect of class [$F(1,7) = 2.38$, $p = .17$]. Although the statistical analysis and a numerical inspection of the figures suggest that rate of response may be the prime driver of the evaluative transfer, these analyses also suggest that this effect was stronger for the DRH schedule than for the VR schedule, constant with a potential secondary role for the response reinforcement rule.

It may also be possible that the overall reinforcement rate may play a role in how the stimuli paired with each schedule were categorized. When all the participants were analyzed, this was found to be significantly higher for the DRH schedule (53.17 outcomes) than for the VR schedule (52.50), yet no significant differences in the transfer-of-

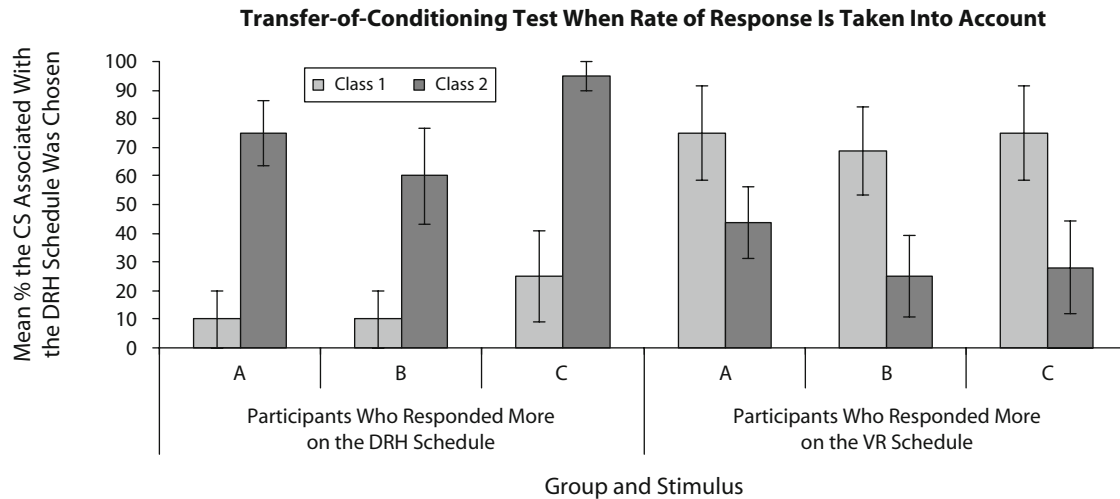


Figure 5. Results from Experiment 2: Mean percentages that the stimulus (colored circle) conditioned with the differential reinforcement of high rate (DRH) schedule was chosen with stimuli from each equivalence class (Class 1 = *useless*; Class 2 = *good*) for each group (higher rate of response on the DRH vs. higher rate of response on the variable ratio [VR] schedule).

conditioning test were seen, suggesting that reinforcement rate was not such a controlling factor as response rate. However, if the reinforcement rates correlate with the response rates within the two rate of response groups (i.e., a higher reinforcement rate on the VR schedule in the VR higher rate of response group and a higher rate of reinforcement on the DRH schedule in the DRH higher rate of response group), it could be the differences in reinforcement rate that produce the observed effect in the transfer-of-conditioning test.

To test this, the overall reinforcement rates for the DRH and VR schedules were compared within the two response rate groups. For the DRH higher rate of response group, there was no significant difference between participants' overall rate of reinforcement on the DRH schedule (53.40 points, $SD = 2.45$) and the VR schedule (52.75 points, $SD = 2.98$) [$t(4) = 1.73, p = .16$]. For the VR higher rate of response group, the difference between participants' overall rate of reinforcement on the DRH schedule (53.03, $SD = 3.43$) and the VR schedule (52.34, $SD = 3.83$) approached significance [$t(7) = 2.15, p = .068$]. Since the VR higher rate of response group gained more points on the DRH schedule but tended to categorize the stimulus associated with the DRH schedule with the Class 1 members (*useless*), this suggests that the evaluative transfer was driven more by the differences in response rates than by those in reinforcement rates. However, the fact that there were no significant differences in how the DRH stimulus was categorized between Class 1 and Class 2 in the VR higher rate of response group suggests that reinforcement rate does influence the evaluative transfer to some extent.

EXPERIMENT 3

In both of the previous experiments, it was shown that differences in the rate of response between the two sched-

ules used in the evaluative conditioning tasks seem to be an important factor in determining the transfer of affect (i.e., participants tend to prefer the stimulus that is related to a higher rate of response). This higher rate of response may also be related to the strength of the response reinforcement rule produced by the schedules. However, the schedules used in the previous two experiments reported here differ in two aspects concerning the response reinforcement rule: first, in terms of their overall response reinforcement feedback function (molar), and second, in terms of the pattern of responding typically emitted just prior to reinforcement (molecular). The feedback function relating the rate of response to the rate of reinforcement has been implicated in previous research (McDowell & Wixted, 1986; Reed, 2007) as an influencing factor on causal effectiveness. When there is a strong correlation between responding and reinforcement (e.g., on VR and DRH schedules), levels of responding and ratings of the causal efficacy of responses tend to be higher than when there is only a weak overall relationship between responding and reinforcement (e.g., on a VI or DRL schedule). It is important to note that the molecular aspects of a schedule have also been found to impact on response rates and causal efficacy ratings (Peele et al., 1984). If a high number of responses are required for reinforcement (e.g., on a VR or DRH schedule), the ratings of causal efficacy of those responses also tend to be higher (Reed, 2001a, 2001b, 2003).

In Experiment 3, we sought to examine whether it is the molar or the molecular aspects of the contingencies of reinforcement that are responsible for the subsequent transfer of affect to the stimuli related to each schedule. By comparing tandem VIVR and tandem VRVI schedules with one another, it was possible to study the impact of the pattern of responses emitted just before reinforcement on the transfer of affect to each stimulus, while simul-

taneously controlling for the overall response reinforcement feedback function. Participants are predicted to have higher rates of response just before reinforcement with the tandem VIVR schedule than with the tandem VRVI schedules. Such an effect has been noted numerous times in nonhumans (Ferster & Skinner, 1957; Peele et al., 1984) and in humans (Reed, 1993). However, the two schedules do not differ in their overall molar feedback function relating rate of response to rate of reinforcement. It was therefore predicted that if it is the feedback function relating response rate to rate of reinforcement that is solely responsible for the effects in the transfer-of-function test, then, given similar feedback functions, there should be no differences in how the stimuli associated with the VRVI and VIVR schedules are categorized or rated. However, if the molecular aspects of the schedules influence the evaluations of the conditioned stimuli, it would be predicted that the stimulus related to the VIVR schedule would be categorized more with the word *good*, as well as being rated more highly than the stimulus related to the VRVI schedule.

The previous experiments used a transfer-of-conditioning test that measures whether the CSs can join an existing equivalence class by virtue of sharing a common evaluative function (*good* or *useless*) with the members of that equivalence class. However, another possible procedure for testing for transfer is to train a novel function to one member of the equivalence class after it has been established and measure whether this function transfers to the other members of the equivalence class without further training (e.g., if A is paired with a positive or negative attribution, then the C [and B] stimulus will also acquire positive or negative attributions). This phenomenon is referred to as the *transfer*, or *transformation*, of function and has been demonstrated with a number of derived stimulus relations and behavioral functions, such as avoidance responses, preferences, self-discrimination, moods, and so forth (see Dymond & Rehfeldt, 2000, for a review). Given the potential of equivalence classes to explain widespread transfer of functions (see Dymond & Rehfeldt, 2000), a further aim of this experiment was to determine whether equivalence class members used as conditioned stimuli, presented during the evaluative conditioning task, could produce a transfer of that function (*good* or *useless*) to other members of an existing equivalence class that had not been directly conditioned to that evaluative function. If so, this would extend the generality of the previous effect (Experiment 1; see Dack et al., 2009).

Method

Participants

Nineteen adults, as described in Experiment 1, participated in the present experiment (6 of them male, 13 female). The age of the participants ranged from 19 to 26 years, with a mean of 20.56 years ($SD = 1.98$).

Apparatus and Materials

The apparatus and materials were identical to those employed in the previous experiments, except for the stimuli used as members of the two equivalence classes and the CSs presented during

each schedule of reinforcement. The A and B members of each class remained the same (see Table 1), but the C1 and C2 members of each class were changed to *casors* and *boceem*, respectively. These stimuli also acted as the CSs during the schedules of reinforcement; C1 was presented during the VRVI schedule of reinforcement, and C2 was presented during the VIVR schedule.

Procedure

Each participant was taken into a quiet room and was given a consent form to read and sign and the BDI to answer before completing the experiment.

Derived stimulus relations training. This was identical to that in the previous experiments, except that the C stimuli were *casors* and *boceem*, instead of *useless* and *good*.

Evaluative conditioning task. The participants were presented with the following instructions on the computer screen:

Try to earn as many points as possible. Click here to proceed.

After presentation of these instructions, the participants were exposed to the schedule trials, which consisted of two alternating schedules (a tandem VRVI and a tandem VIVR). Each schedule was presented four times to each participant. Each schedule lasted for 4 min. The tandem VRVI schedule was signaled by the presence of C1 (*casors*) on the computer screen during the task, and the tandem VIVR schedule was signaled by the presence of C2 (*boceem*) on the screen.

In one schedule condition, the participants responded on a tandem VR 10 VI 20-sec schedule. The participants were required to emit an average of 10 responses to produce access to the VI 20-sec schedule. The participants were unaware of when the component schedule changed from the VR to the VI. After completion of the VI component, an outcome (point) was received for a response on a space bar, using a formula in which outcomes were distributed around an average of 20 sec, with a range of 1–39 sec. After the point was received, the schedule then changed back to the VR 10 schedule (again, the participants were unaware of this change), and the contingency continued to cycle between the two schedules, in this manner, until the trial ended.

In a second schedule condition, the participants responded on a tandem VI 20-sec VR 10 schedule. The participants had to satisfy a VI 20-sec schedule in order to get access to the VR 10 schedule. They received an outcome (point) only once they had responded around an average of 10 times on the VR 10 schedule. After the delivery of a point, the VI schedule was then operative again. The component schedules were not differentially signaled.

Transfer-of-function test. The participants were presented with the following instructions on the computer screen:

Please click on the word that appears at the top of the screen, and drag it on top of one of the 3 words at the bottom of the screen. You should choose the word at the bottom of the screen which you feel best represents how you would perform on the previous space bar task if that word had appeared on the screen.

Each trial began with the presentation of one of the class members of either equivalence class (including the conditioned stimuli, C1 and C2) in the middle at the top of the screen and the words *good*, *average*, and *bad* at the bottom of the screen. The participants had to drag and drop the stimulus at the top of the screen onto whichever word (*good*, *average*, or *bad*) they thought reflected how they would perform if this was the conditioned stimulus in the space bar task. Each stimulus was shown four times, with a total of 24 trials. At no point during this test were the participants given any feedback for their responses.

Results and Discussion

Before the analyses were performed, 7 participants were excluded for high scores on the BDI (>10 ; see Reed et al.,

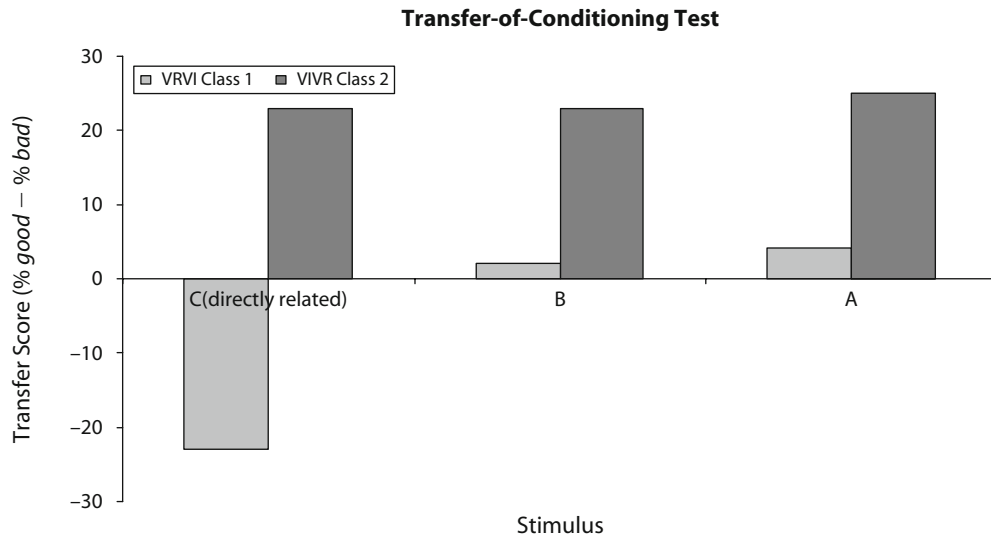


Figure 6. Results from Experiment 3: Amount of transfer to the stimulus conditioned with each schedule (VRVI, VIVR) and to the stimuli related to the conditioned stimulus through equivalence. VR, variable ratio; VI, variable interval.

2001). The mean score for the remaining participants on the BDI was 4.83 (± 2.59).

Derived Stimulus Relations Testing

All the participants passed the emergent relations test with an overall mean of 99.07% (± 0.71) class-consistent responding. This was statistically significantly more correct responses than would be expected by chance [$t(11) = 68.97$]. Ten participants passed the equivalence test after one training phase; 2 participants needed a second training phase before reaching the test criterion (i.e., 83% correct). The mean number of trials received during the training phase(s) was 37.00 (range = 14–73, $SD = 20.17$) before the test criterion was met. The mean number of these trials that were correct was 28.92 (range = 12–50, $SD = 14.63$).

Evaluative Conditioning Task

The mean rate of responding produced by participants, collapsed across all four trials, was higher for the tandem VIVR schedule (161.16, $SD = 116.85$) than for the tandem VRVI schedule (121.88, $SD = 121.88$). This finding replicated those of previous reports that have studied these two schedules (e.g., Peele et al., 1984; Reed, 1993). However, a two-factor ANOVA (schedule \times trial) revealed no statistically significant main effect of schedule [$F(1,11) = 2.37$] and no effect of trial or interaction between schedule and trial ($F_s < 1$).

The mean number of outcomes per minute, averaged across all four trials, was slightly higher in the tandem VIVR schedule (2.15, $SD = 1.05$) than in the tandem VRVI schedule (1.75, $SD = 0.93$). A two-factor repeated measures ANOVA (schedule \times trial) revealed no main effect of schedule [$F(1,11) = 4.33$, $p = .06$] or trial [$F(3,33) = 1.32$] and no statistically significant interaction ($F < 1$).

The probability of an outcome was approximately the same for each schedule, VRVI (.03) and VIVR (.02).

Transfer-of-Conditioning Test

To produce a transfer-of-conditioning score for each stimulus (A, B, and C), the percentage that each stimulus was categorized with the word *bad* was subtracted from the percentage that each stimulus was categorized with the word *good*. A negative score, therefore, reflected a negative evaluation (*dislike*) for that stimulus, whereas a positive score reflected a positive evaluation (*like*) for that stimulus. From Figure 6, it can be seen that Class 1 members, associated with the tandem VRVI schedule, had lower scores than did Class 2 members (associated with the tandem VIVR schedule). Stimulus C1 (VRVI) had a negative value, whereas stimulus C2 (VRVI) had a positive value. However, these differences were not statistically significant for each cross-class pair [$A1-A2$, $t(11) = 0.50$; $B1-B2$, $t(11) = 0.66$; $C1-C2$, $t(11) = 1.26$]. An ANOVA (stimuli \times class) was performed on these data and revealed no statistically significant effect of stimuli or class and no significant interaction between the two factors (all $F_s < 1$).

Transfer-of-Function Questionnaire

The mean ratings (*useless–neutral–good*) for each class member, including C1 and C2, that had been directly related to the two different schedules can be seen in Figure 7. An inspection of these data shows slightly lower ratings (more *useless*) for the members of Class 1 (associated with the VRVI schedule), as compared with Class 2 (associated with the VIVR schedule). These differences were more prominent for the C stimulus that had been presented during the evaluative conditioning task. However, these differences were not statistically significant for each cross-class pair [$A1-A2$, $t(11) = 0.21$; $B1-B2$, $t(11) = 0.13$; $C1-C2$, $t(11) = 1.33$]. A repeated measures ANOVA (stimuli \times class) showed no significant effect of class or stimulus, nor was there a significant interaction between the two factors (all $p_s > .20$).

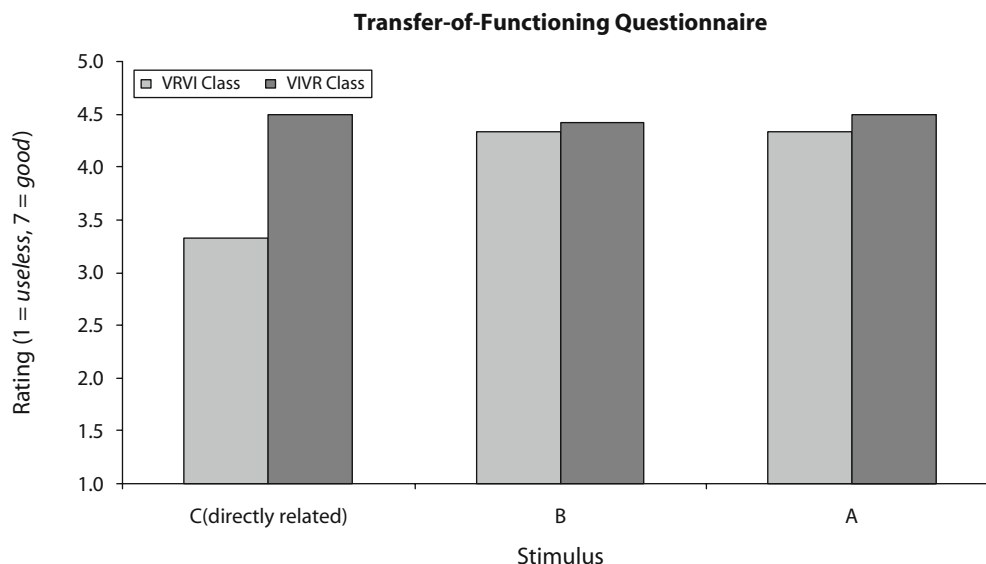


Figure 7. Results from Experiment 3: Mean rating for the stimuli in each class (stimuli C1 and C2 were directly related to the schedule), with standard error bars. VR, variable ratio; VI, variable interval.

In summary, the results replicate the findings of Experiment 2 showing that when there is no significant difference between the participants' rate of responding to both schedules, there is also no significant difference in the affective evaluations given to the stimuli associated with each schedule. It was therefore difficult to assess whether an evaluative function trained directly to an equivalence class member (i.e., the stimuli associated with each schedule) would transfer to other class members that had not been directly related to the schedules through derived stimulus relations. Further research in which schedules are used that do differ significantly in the behavior associated with them is needed to investigate this possibility. If the evaluative functions of causal efficacy were shown to transfer to stimuli unrelated to the schedules of reinforcement, this could potentially provide a mechanism for the development of problems experienced by individuals that can occur without direct learning experience, as are observed in some anxiety disorders and phobias.

The pattern of results suggests, since the overall feedback function between the rate of responding and reinforcement rate was equated between the schedules and the stimuli associated with the schedules did not differ particularly in their evaluative ratings, that it may be the molar characteristics of schedules of reinforcement that are associated with stronger differences in the transfer of valence. Another possibility is that the participants' behavior may have been insensitive to the changes between the schedules and that longer trial lengths may have been needed to produce a significant response rate difference. The changes in the procedure of the transfer-of-conditioning test may have also impacted on the results. Further research would be needed to investigate whether the apparent lack of an effect was a product of procedural variation or whether, in fact, there was

no difference between the rates of responding on the schedules.

EXPERIMENT 4

In the previous three experiments, schedules were compared that had relatively strong response reinforcement rules (e.g., VR, DRH). In these contexts, it has been shown that it is differences in the rate of response between each schedule that seem to determine the evaluative function of the stimuli associated with each schedule. In Experiment 4, we turned to examine the effects of schedules that specify only weak response-reinforcement-based rules on the transfer of valence.

Previous research (Reed, 2007) has shown that an increase in response rate leads to an increased sensitivity to the molar characteristics of the schedule in rats. It might therefore be expected that schedules that lead to low rates of responding (DRL, VI) may increase participants' sensitivity to the molecular aspects of the schedule.

To this end, a DRL schedule was compared with a VI schedule, yoked in terms of the rate of reinforcement. Under these conditions, it might be expected that the transfer of function would be influenced by the molecular aspects of the schedule. Since a low rate of responding is reinforced with both schedules, the probability that a response will produce an outcome should also be higher on both schedules, potentially highlighting this factor as salient for the participants. Since the DRL schedule actually specifies that reinforcement is delivered only for a response that is separated from the preceding response by a specified minimum time, it would be predicted that there would be a higher probability of an outcome given a response in the DRL schedule, as compared with the VI. If this factor is important in determining the transfer effect, it is predicted that the DRL stimulus will be categorized

with the stimuli participating in the *good* equivalence class and be given higher ratings than the VI stimulus.

Method

Participants

Twenty-two adults, as described in Experiment 1, participated in the present experiment (6 of them male, 16 female). The ages of the participants ranged from 18 to 39 years, with a mean of 22.43 years ($SD = 5.61$).

Apparatus and Materials

The apparatus and materials were the same as those in Experiment 1, apart from those in the evaluative conditioning task. The only difference in this task was the different combinations of colored circles that were used to signal the two scheduling conditions (DRL, VI). Four different combinations were used (yellow/red, yellow/blue, blue/green, green/yellow), and these were counterbalanced across the participants (3 participants had blue [DRL] and green [VI], 2 participants had green [DRL] and blue [VI], 2 participants had blue [DRL] and yellow [VI], 3 participants had yellow [DRL] and blue [VI], 3 participants had red [DRL] and yellow [VI], 3 participants had yellow [DRL] and red [VI], 3 participants had green [DRL] and yellow [VI], and 3 participants had yellow [DRL] and green [VI]).

Procedure

Each participant was taken into a quiet room and was given a consent form to read and sign and the BDI to answer before completing the experiment.

Derived stimulus relations training. This was the same as described in Experiment 1.

Evaluative conditioning task. The participants were presented with the following instructions on the computer screen:

You will shortly see a circle in the centre of the screen. The circle will be colored either [Color 1] and [Color 2] [see note 1]. When you see this circle you must press the space bar to earn points. Try to earn as many points as possible. You must work out what rate of pressing makes you the most points. Click here to proceed.

After presentation of these instructions, the participants were exposed to the schedule trials, which consisted of two alternating schedules (DRL and VI). Each schedule was presented three times to each participant. The DRL schedule was always presented first and lasted for 2 min. Each VI schedule lasted until the participants had received approximately the same number of points as in the previous DRL schedule.

In one schedule condition, participants responded on a DRL 5-sec schedule. In this schedule, an outcome (a point) was received immediately after a response on a space bar, provided that at least 5 sec had elapsed since the last response had been made or the last point had been received. A response before 5 sec had elapsed reset the DRL time requirement. In a second schedule (VI), participants responded to a VI schedule, yoked to the temporal distribution of outcomes obtained in the DRL condition. The yoking was done by dividing the total number of outcomes (points) per trial on the DRL schedule by the total time of each trial (120 sec). This number was then used as the overall mean time interval between outcomes (points) for the yoked VI schedule. Points were received after a time interval that varied around this mean value.

Transfer-of-conditioning test. This was the same procedure as that described in Experiment 1.

Overall, 12 of the participants received the derived stimulus relations task before the evaluative conditioning task, and the other 10 participants received the tasks in the opposite order. All the participants completed the transfer-of-conditioning test last. At the end of the experiment, each participant completed the transfer-of-

function questionnaire. They were thanked for participating and were debriefed.

Results and Discussion

On the basis of the BDI scores (see Reed et al., 2001), 5 participants were excluded from the analysis (BDI scores > 10). The mean score for the remaining participants on the BDI was 4.47 (± 0.74).

Derived Stimulus Relations Testing

Before the analyses were performed, 2 participants were excluded, since they withdrew from the study after failing to reach the pass criterion (83% correct class-consistent responding) after two training exposures. The remaining participants passed the emergent relations test with an overall mean of 96.48% (± 1.47) class-consistent responding. This was statistically significantly more correct responses than would be expected by chance [$t(14) = 31.58$]. Fourteen participants reached the passing criterion on the first test session, and 1 participant needed a second training phase before reaching the test criterion (i.e., 83% correct). The mean number of trials received during the training phase(s) was 49.20 (range = 13–178, $SD = 53.70$) before the test criterion was met. The mean number of these trials that were correct was 34.60 (range = 12–115, $SD = 30.88$).

Evaluative Conditioning Task

The mean rates of responses to both schedules are shown in Figure 8. This shows a slightly higher rate of response to the VI than to the DRL schedule. A repeated measures ANOVA (schedule \times trial) revealed a statistically significant main effect of trial [$F(2,28) = 4.77$]. There was no statistically significant effect of schedule [$F(1,14) = 1.19$] and no significant interaction between the two factors ($F < 1$).

Since the VI schedule was yoked to the temporal distribution of outcomes obtained in the DRL schedule, the total number of outcomes was the same for both conditions on each trial. The mean number of outcomes on

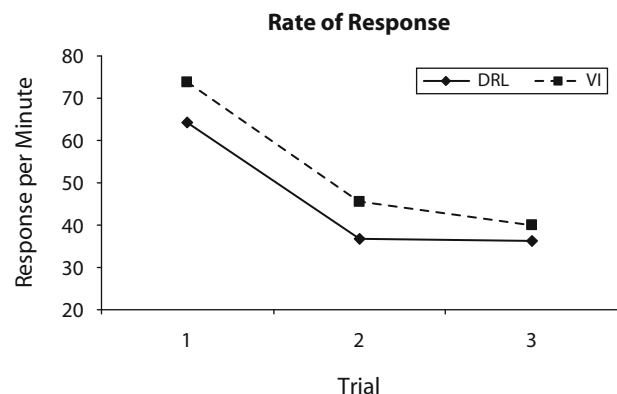


Figure 8. Results from Experiment 4: Mean responses per minute across each trial for each schedule. DRL, differential reinforcement of low rate; VI, variable interval.

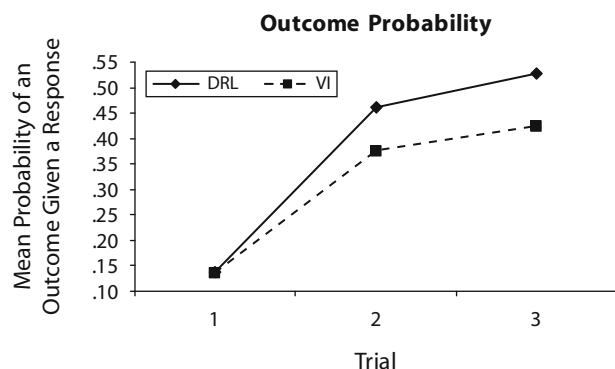


Figure 9. Results from Experiment 4: Mean probability of an outcome given a response across each trial for each schedule (differential reinforcement of low rate [DRL] and variable interval [VI]).

Trial 1 was 6.00 ($SD = 5.45$), on Trial 2, 10.80 ($SD = 7.80$), and on Trial 3, 11.00 ($SD = 7.84$). An ANOVA revealed a statistically significant main effect of trial [$F(2,28) = 5.62$].

The mean probabilities of an outcome given a response are shown in Figure 9. This shows an increase in outcome probability across trials, with a greater probability of an outcome given a response on the DRL, as compared with the VI, schedule by the end of training. A repeated measures ANOVA (schedule \times trial) revealed no statistically significant main effect of schedule [$F(1,14) = 2.73$] and no interaction between trial and schedule [$F(2,28) = 1.82$]. There was a significant main effect of trial [$F(2,28) = 10.44$].

Transfer-of-Conditioning Test

Figure 10 reveals that the participants chose the CS associated with the DRL schedule in the presence of the stimuli participating in Class 1 (*useless*) less than in the presence of the stimuli participating in Class 2 (*good*).

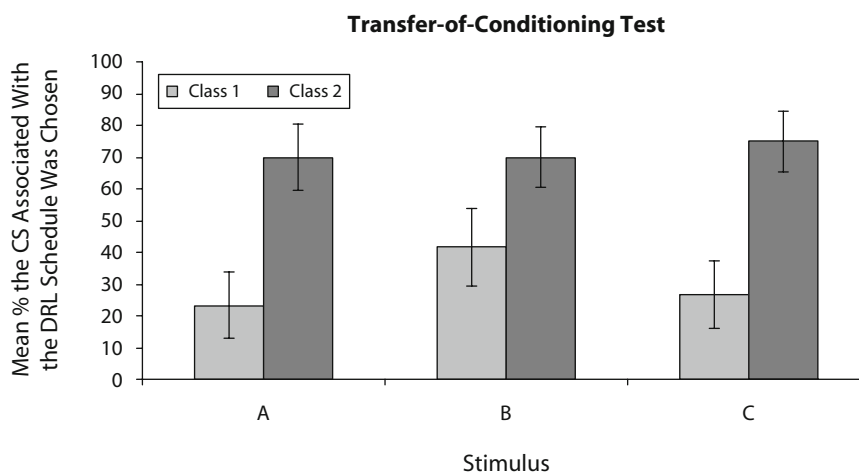


Figure 10. Results from Experiment 4: Mean percentages that the stimulus (colored circle) conditioned with the differential reinforcement of low rate (DRL) schedule was chosen with stimuli from each equivalence class (Class 1 = *useless*; Class 2 = *good*). Error bars represent mean standard errors.

A repeated measures ANOVA (class, stimuli) performed on the data showed a significant effect of class [$F(1,14) = 4.40$]. There was no significant effect of stimuli [$F(2,28) = 3.15$] and no significant interaction between the two factors [$F(2,28) = 2.81$]. Dependent t tests revealed that the stimulus associated with the DRL schedule was chosen significantly more with the A2 stimulus than with the A1 stimulus [$t(14) = 4.54$] and with the C2 stimulus than with the C1 stimulus [$t(14) = 4.90$]. No significant difference was found between the percentages that the DRL stimulus was chosen in the presence of the B1 and B2 stimuli [$t(14) = 2.74, p = .10$].

Transfer-of-Function Questionnaire

The mean ratings for each nonsense stimulus in each class and for the two CSs that signaled the different schedules can be seen in Figure 11. An inspection of these data again shows lower ratings (more *useless*) for the members of Class 1, as compared with Class 2. The ratings for the VI stimulus were similar to the ratings for Class 1 members (*useless*), whereas the ratings for the DRL stimulus were similar to the ratings for Class 2 members (*good*). A two-factor ANOVA (class \times stimuli) was performed on the data. This revealed a significant effect of class [$F(1,14) = 90.70$]. There was no main effect of stimuli ($F < 1$) and no interaction between the two factors [$F(2,28) = 2.21$].

In summary, the probability of getting an outcome given a response was higher with a DRL schedule than with a VI schedule with the same frequency of reinforcement (although this difference was not significant). The stimulus that was associated with the DRL schedule was rated as *good*, and the stimulus related to the VI schedule as *useless*. Participants also showed that they would categorize these CSs with existing equivalence classes with the same function (i.e., *good* or *useless*), despite the fact that they had never previously been directly related. This pattern of results suggests that when schedules are compared with weak response reinforcement rules, the molecular proper-

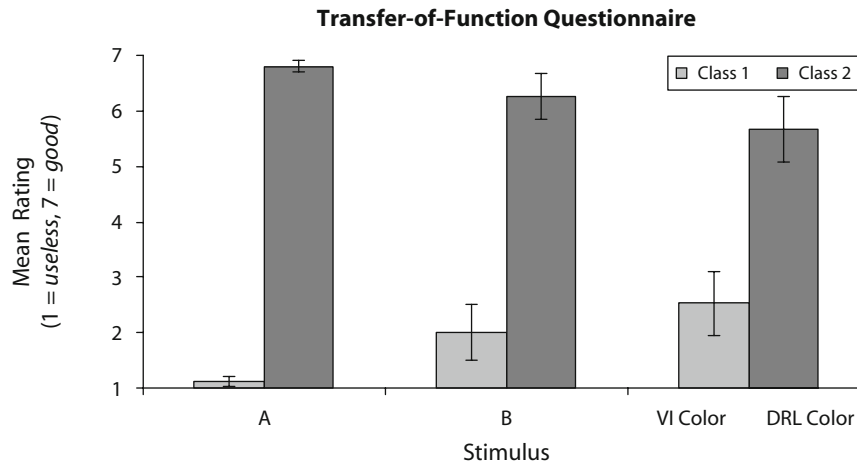


Figure 11. Results from Experiment 4: Mean ratings for the stimuli in each class (Class 1 = useless; Class 2 = good) and the conditioned stimuli that signaled each schedule, with standard error bars. VI, variable interval; DRL, differential reinforcement of low rate.

ties of the schedule may have more influence in determining the valence of the associated stimuli. In particular, it is the high probability that one response will produce an outcome as on the DRL schedule that produces higher ratings of efficacy in the transfer-of-function questionnaire. This replicates previous findings (Reed, 1992).

GENERAL DISCUSSION

The present series of experiments explored the numerous factors, such as response rate, response reinforcement rule, and probability of an outcome given a response, that can differ between schedules of reinforcement and their impact on the evaluative ratings (*good*, *useless*) given to CSs associated with each schedule. In addition, the aim of the experiments was to demonstrate whether these stimuli would join established equivalence classes, through derived stimulus networks, with the same evaluative function (*good*, *useless*), even though these stimuli had not been directly associated with either of the schedules (Experiments 1, 2, and 4). In addition, Experiment 3 was designed to test for direct transfer by using CSs during the schedule tasks that were already members of preexisting equivalence classes, so participants could categorize the related stimuli. It was hoped that this would extend the generality of the effects.

Overall, these findings replicate those of Dack et al. (2009), in that the transfer of valence from the schedule task to the related stimuli appears to be determined mainly by differences in response rate between the schedules (Experiments 1 and 2) and, to a lesser extent, by the response reinforcement rule and reinforcement rate (Experiment 2). However, when neither schedule was based on a strong response reinforcement rule (Experiment 4), although the differences in the probability of an outcome given a response between schedules was not found to be significant, participants rated the stimuli associated with a DRL schedule (higher outcome probability) significantly bet-

ter than the stimuli associated with a VI schedule (lower outcome probability). The DRL stimulus was also categorized significantly more times with the *good* equivalence class members than was the VI schedule stimulus. This suggests that in this context, even a small difference in outcome probability between schedules with a weak response reinforcement rule can produce differences in the valence of the associated stimuli.

Importantly, these results show how complex schedules of reinforcement are (see Jenkins, 1970) and that, in different contexts, different aspects of the schedule can determine how *good* or *bad* the CSs present during the task are rated. When two schedules are compared with very different response reinforcement rules (VR vs. VI), the stimulus associated with the schedule that produces higher response rates (VR) is categorized more with the *good* class members and is given higher ratings than is the stimulus associated with the schedule with lower rates of responding (despite an equal rate of reinforcement on both schedules and a higher outcome probability produced by the VI schedule). It seems that, in this context, it is the factors responsible for the differences in response rate (not reinforcement rate) that are responsible for the evaluative functions.

This interpretation was supported by Experiments 2 and 3, where the transfer effects were generally weaker than in Experiment 1. The rate of response was the same for both schedules in Experiment 2 (VR and DRH) and Experiment 3 (VRVI, VIVR), suggesting that when this aspect of the schedules is similar, there is little difference in how the stimuli associated with each schedule are rated and categorized. When participants were separated depending on which schedule they responded the most on in Experiment 2, an interaction between group and which class the schedule stimuli were categorized with was found. The stimulus associated with the DRH was categorized with the *good* equivalence class in participants who had a higher response rate on the DRH schedule (as

compared with the VR schedule). These effects suggest that the molar aspects of the schedule may play some role in determining the valence of the associated stimuli, in addition to the overt rate of response.

In Experiment 4, where neither schedule (DRL, VI) reinforced high rates of responding, participants rated and categorized the stimulus associated with the schedule that had a higher probability of an outcome given a response (DRL), despite the fact that this schedule had lower response rates. By employing schedules with weaker response–reinforcer correlations and the resulting low rates of responding, it seems that the participants paid more attention to the molecular aspects of the schedule and, in particular, the probability of an outcome given a response prior to reinforcement. In a DRL schedule, where a response has to be produced in isolation to be reinforced, there will be a higher probability that just one response will produce an outcome, in comparisons with a VI schedule. In this context, the stimulus associated with the high outcome probability (i.e., the perception that one response will produce an outcome) is evaluated positively. This effect was found despite the fact that the differences in outcome probability and response rates were not statistically significant.

It was thought that the results reported by Dack et al. (2009) may have been influenced by the interaction between the previous experimental procedure and the possible task demands brought about by common labeling between the efficacy judgments made after each exposure to the schedule (e.g., *ineffective*, *effective*) and each equivalence class (each class contained either the word *good* or *useless*). This, in turn, may have led to some semantic ambiguity in how participants categorized the related stimuli in the transfer-of-conditioning test. However, a transfer of valence was observed in Experiments 1 and 4, despite no judgment of causal efficacy being implemented after each schedule trial. Since there were still differences observed in the transfer-of-conditioning test and the transfer-of-function questionnaire after these judgments were removed from the procedure, Experiments 1 and 4 provide evidence of a genuine transfer effect.

Future studies are needed replicating these findings using a clinical population before the findings of these experiments can be generalized to such clinical populations. Previous research (Alloy & Abramson, 1979) has shown differences in judgments of contingencies between response and outcomes between depressed and nondepressed students. Depressed participants were found to judge the contingencies less favorably but more accurately, whereas the nondepressed participants overestimated the degree of contingency between their responses and outcomes. These results may be replicated in depressed and nondepressed participants' ratings of causal efficacy, where higher ratings may be found in a nondepressed population, as compared with a depressed population. There may also be differences in the transfer of causal efficacy functions between depressed and nondepressed groups. Depressed individuals would be expected to have more negative self-evaluations about their effectiveness at the schedule tasks,

and, therefore, more negative transfer might be expected. Further research repeating these experiments, but using a group of clinical participants, might shed some light on any differences between the factors of schedules of reinforcement that produce high or low causal efficacy ratings and, subsequently, any differences in the transfer of causal efficacy functions.

Importantly, the results of these experiments inform us about the types of contingencies within the environment, in a nondepressed population, that might produce negative and positive evaluations about the self and, through evaluative conditioning and derived stimulus relations, how these evaluations about the self might generalize to negative *schemas* about the environment and future events. In the majority of the experiments (see also Dack et al., 2009; cf. Experiment 4), the more participants responded in the task, the more they felt effective in the task. This might suggest that tasks that are effortful may lead to a sense of achievement, whereas tasks where a person is rewarded for doing very little, or even nothing at all, might possibly produce feelings of *guilt* or *fraudulence*. It has been shown that when participants imagine that they have worked hard on a project and succeeded, they expected more positive affective reactions than when they imagine that they succeeded with little effort (Jagacinski & Nicholls, 1984). To explore this issue, future studies could employ psychometric measures to investigate any changes in participants' mood after being exposed to the different schedules of reinforcement.

However, in people with high levels of depressive symptomatology, differences may be found in how they rate their causal efficacy under different contingencies. In particular, there is some evidence (Costello, 1972) that the delivery of points may not function as reinforcers in depressed individuals. Given this, it might be predicted that the patterns of causal efficacy would follow the results produced by outcome schedules, where higher ratings of efficacy are produced by schedules of reinforcement (such as DRL and VI schedules) that tend to reinforce low rates of responding, as compared with schedules (e.g., DRH or VR schedules) that reinforce high rates of responding (Reed, 1992, 1993). In these studies, the outcomes (i.e., a flashing light) used in the schedule tasks are also unlikely to act as reinforcers, and so similar patterns of causal efficacy might be expected. Further research would be needed to confirm this speculation.²

These experiments begin to inform us about the factors in the environment that produce positive and negative evaluations of self-performance in a task and how evaluative conditioning may model how certain stimuli paired with these evaluations come to elicit feelings of hopelessness, and helplessness, in a nonclinical population. In the transfer-of-conditioning tests, it was then shown that the CSs associated with the task would then join established stimulus equivalence classes, through derived stimulus networks, with the same evaluative function. This may be one possible learning-based approach to explaining the development of generalized evaluations (such as effectiveness, helplessness, etc.) without direct experience

of each stimulus that could potentially add to previous cognitive and learning models, which have difficulties in explaining where beliefs about the self (including causal efficacy) originate from and the precise mechanisms that cause them to transfer across a wide range of seemingly unrelated areas of a person's life.

AUTHOR NOTE

The research described here was supported in part by a research grant from the United Kingdom Economic and Social Research Council (ESRC) and the Leverhulme Trust. Correspondence concerning this article should be addressed to L. McHugh, Department of Psychology, Swansea University, Singleton Park, Swansea SA2 8PP, Wales (e-mail: l.mchugh@swansea.ac.uk).

REFERENCES

- ABRAMSON, L. Y., METALSKY, G. I., & ALLOY, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, **96**, 358-372.
- ALLAN, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, **15**, 147-149.
- ALLOY, L. B., & ABRAMSON, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, **108**, 441-485. doi:10.1037/0096-3445.108.4.441
- ALLOY, L. B., & TABACHNIK, M. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, **91**, 112-149.
- BAEYENS, F., EELEN, P., CROMBEZ, G., & VAN DEN BERGH, O. (1992). Human evaluative conditioning: Acquisition trials, presentation schedule, evaluative style and contingency awareness. *Behavior Research & Therapy*, **30**, 133-142. doi:10.1016/0005-7967(92)90136-5
- BARNES-HOLMES, D., KEANE, J., BARNES-HOLMES, Y., & SMEETS, P. M. (2000). A derived transfer of emotive functions as a means of establishing differential preferences for soft drinks. *Psychological Record*, **50**, 493-511.
- BAUM, W. M. (1973). The correlational-based law of effect. *Journal of the Experimental Analysis of Behavior*, **20**, 137-153. doi:10.1901/jeab.1973.20-137
- BAUM, W. M. (1993). Performances on ratio and interval schedules of reinforcement: Data and theory. *Journal of the Experimental Analysis of Behavior*, **59**, 245-264. doi:10.1901/jeab.1993.59-245
- BECK, A. T. (1979). *Depression: Clinical, experimental, and theoretical aspects*. New York: Hoeber.
- BECK, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.
- BECK, A. T. (1978). *The depression inventory*. Philadelphia: Center for Cognitive Therapy.
- BECK, A. T., RUSH, A. J., SHAW, B. F., & EMERY, G. (1979). *Cognitive therapy of depression*. New York: Guilford.
- BECK, A. T., WARD, C. H., MENDELSON, M., MOCK, J., & ERBAUGH, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, **4**, 561-571.
- BUSH, K. M., SIDMAN, M., & DE ROSE, T. (1989). Contextual control of emergent equivalence relations. *Journal of the Experimental Analysis of Behavior*, **51**, 29-46. doi:10.1901/jeab.1989.51-29
- CLARK, D. A., BECK, A. T., & ALFORD, B. A. (1999). *Scientific foundations of cognitive theory and therapy of depression*. New York: Wiley.
- COSTELLO, C. G. (1972). Depression: Loss of reinforcers or loss of reinforcer effectiveness. *Behavior Therapy*, **3**, 240-247.
- DACK, C., MCHUGH, L., & REED, P. (2009). Generalization of causal efficacy judgments after evaluative learning. *Learning & Behavior*, **37**, 336-348. doi:10.3758/LB.37.4.336
- DE HOUWER, J., THOMAS, S., & BAEYENS, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, **127**, 853-869.
- DOHR, K. B., RUSH, A. J., & BERNSTEIN, I. H. (1989). Cognitive biases and depression. *Journal of Abnormal Psychology*, **98**, 263-267.
- DYMOND, S., & REHFELDT, R. A. (2000). Understanding complex behavior: The transformation of stimulus functions. *Behavior Analyst*, **23**, 239-254.
- FERSTER, C. B. (1973). A functional analysis of depression. *American Psychologist*, **28**, 857-870.
- FERSTER, C. B., & SKINNER, B. F. (1957). *Schedules of reinforcement*. New York: Appleton-Century-Crofts.
- FOXALL, M. R. (1997). The emotional texture of consumer environments: A systematic approach to atmospherics. *Journal of Economic Psychology*, **5**, 505-523.
- FOXALL, M. R., & GREENLEY, G. E. (1999). Consumers' emotional responses to service environments. *Journal of Business Research*, **46**, 149-158.
- FOXALL, M. R., & YANI-DE-SORIANO, M. M. (2005). Situational influences on consumers' attitudes and behavior. *Journal of Business Research*, **58**, 518-525.
- FULCHER, E. P., MATHEWS, A., MACKINTOSH, B., & LAW, S. (2001). Evaluative learning and the allocation of attention to emotional stimuli. *Cognitive Therapy & Research*, **25**, 261-280. doi:10.1023/A:1010732328104
- HEBL, M. R., & MANNIX, L. M. (2003). The weight of obesity in evaluating others: A mere proximity effect. *Personality & Social Psychology Bulletin*, **29**, 28-38.
- HERMANS, D. (1998). Evaluative conditioning [Special issue]. *Gedragstherapie*, **31**, 3-6.
- JAGACINSKI, C. M., & NICHOLLS, J. G. (1984). Conceptions of ability and related affects in task involvement and ego involvement. *Journal of Educational Psychology*, **76**, 909-919. doi:10.1037/0022-0663.76.5.909
- JENKINS, H. M. (1970). Sequential organization in schedules of reinforcement. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 63-109). Englewood Cliffs, NJ: Prentice Hall.
- LEVAY, A. B., & MARTIN, I. (1975). Classical conditioning of human "evaluative" responses. *Behaviour Research & Therapy*, **13**, 221-226. doi:10.1016/0005-7967(75)90026-1
- LEWINSOHN, P. M. (1974). A behavioral approach to depression. In R. J. Friedman & M. M. Katz (Eds.), *Psychology of depression: Contemporary theory and research* (pp. 157-178). Washington, DC: Winston.
- LOEB, A., FESHBACH, S., BECK, A. T., & WOLF, A. (1964). Some effects of reward upon the social perception and motivation of psychiatric patients varying in depression. *Journal of Abnormal & Social Psychology*, **68**, 609-616.
- MACLEOD, A. K., & BYRNE, A. (1996). Anxiety, depression and the anticipation of future positive and future negative experiences. *Journal of Abnormal Psychology*, **105**, 286-289.
- MCDOWELL, J. J., & WIXTED, J. T. (1986). Variable-ratio schedules as variable-interval schedules with linear feedback loops. *Journal of the Experimental Analysis of Behavior*, **46**, 315-329. doi:10.1901/jeab.1986.46-315
- MORGAN, J., & BANERJEE, R. (2006). Social anxiety and self-evaluation of social performance in a nonclinical sample of children. *Journal of Clinical Child & Adolescent Psychology*, **35**, 292-301.
- MORSE, W. H. (1966). Intermittent reinforcement. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 52-108). Englewood Cliffs, NJ: Prentice Hall.
- MOWRER, O. H. (1960). *Learning theory and the symbolic processes*. New York: Wiley.
- OLATUNJI, B. O. (2006). Evaluative learning and emotional responding to fearful and disgusting stimuli in spider phobia. *Anxiety Disorders*, **20**, 858-876. doi:10.1016/j.janxdis.2006.01.005
- OLATUNJI, B. O., LOHR, J. M., SMITS, J. A. J., SAWCHUCK, C. N., & PATTEN, K. (2009). Evaluative conditioning of fear and disgust in blood-injection-injury phobia: Specificity and impact of individual differences in disgust sensitivity. *Journal of Anxiety Disorders*, **23**, 153-159. doi:10.1016/j.janxdis.2008.06.002
- PEELE, D. B., CASEY, J., & SILBERBERG, A. (1984). Primacy of inter-response time reinforcement in accounting for rates under variable-ratio and variable-interval schedules. *Journal of Experimental Psychology: Animal Behavior Processes*, **10**, 149-167. doi:10.1037/0097-7403.10.2.149
- REED, P. (1992). Effect of local context of responding on human judgment of causality. *Memory & Cognition*, **20**, 573-579.

- REED, P. (1993). Influence of the schedule of outcome presentation on causality judgements. *Quarterly Journal of Experimental Psychology*, **46A**, 327-345.
- REED, P. (1994). Influence of the cost of responding on human causality judgments. *Memory & Cognition*, **22**, 243-248.
- REED, P. (1999). Effect of perceived cost on judgments regarding the efficacy of investment. *Journal of Economic Psychology*, **20**, 657-676. doi:10.1016/S0167-4870(99)00029-X
- REED, P. (2001a). Human response rates and causality judgments on schedules of reinforcement. *Learning & Motivation*, **32**, 332-348. doi:10.1006/lmot.2001.1085
- REED, P. (2001b). Schedules of reinforcement as determinants of human causality judgments and response rates. *Journal of Experimental Psychology: Animal Behavior Processes*, **27**, 187-195.
- REED, P. (2003). Human causality judgments and response rates on DRL and DRH schedules of reinforcement. *Learning & Behavior*, **31**, 205-211.
- REED, P. (2007). Response rate and sensitivity to the molar feedback function relating response and reinforcement rate on VI+ schedules of reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, **33**, 428-439. doi:10.1037/0097-7403.33.4.428
- REED, P., & ANTONOVA, M. (2007). Interference with judgments of control and attentional shift as a result of prior exposure to controllable and uncontrollable feedback. *Learning & Motivation*, **38**, 229-241. doi:10.1016/j.lmot.2006.08.005
- REED, P., FRASQUILLO, F., COLKIN, C., LIEMANN, V., & COLBERT, S. (2001). Interference with judgements of control and learning as a result of prior exposure to controllable and uncontrollable feedback during concept-learning tasks. *Quarterly Journal of Experimental Psychology*, **54B**, 167-183. doi:10.1080/02724990042000155
- ROTH, D., & REHM, L. P. (1980). Relationships among self-monitoring processes, memory and depression. *Cognitive Therapy & Research*, **4**, 149-157. doi:10.1007/BF01173646
- ROZIN, P., MARKWITH, M., & MCCAULEY, C. (1994). Sensitivity to indirect contacts with other persons: AIDS aversion as a composite of aversion to strangers, infection, moral taint, and misfortune. *Journal of Abnormal Psychology*, **103**, 495-504. doi:10.1037/0021-843X.103.3.495
- SCHENLE, A., STARK, R., & VAITL, D. (2001). Evaluative conditioning: A possible explanation for the acquisition of disgust responses? *Learning & Motivation*, **32**, 65-83. doi:10.1006/lmot.2000.1067
- SELIGMAN, M. E. P. (1975). *Helplessness: On depression, development, and death*. San Francisco: Freeman.
- SIDMAN, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech & Hearing Research*, **14**, 5-13.
- SIDMAN, M. (1990). Equivalence relations: Where do they come from? In D. E. Blackman & H. Lejeune (Eds.), *Behavior analysis in theory and practice: Contributions and controversies* (pp. 93-114). Hillsdale, NJ: Erlbaum.
- SIDMAN, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.
- SIDMAN, M., KIRK, B., & WILSON-MORRIS, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. *Journal of the Experimental Analysis of Behavior*, **43**, 21-42. doi:10.1901/jeab.1985.43-21
- SIMON, G. E. (2003). Social and economic burden of mood disorders. *Biological Psychiatry*, **54**, 208-215.
- SMEETS, P. M., & BARNES-HOLMES, D. (2003). Children's emergent preferences for soft drinks: Stimulus-equivalence and transfer. *Journal of Economic Psychology*, **24**, 603-618. doi:10.1016/S0167-4870(03)00004-7
- STUART, E. W., SHIMP, T. A., & ENGLE, R. W. (1987). Classical conditioning of consumer attitudes: Four experiments in an advertising context. *Journal of Consumer Research*, **14**, 334-351.
- WALTHER, E. (2002). Guilty by mere association: Evaluative conditioning and the spreading attitude effect. *Journal of Personality & Social Psychology*, **82**, 919-934.
- WALTHER, E., NAGENGAST, B., & TRASELLI, C. (2005). Evaluative conditioning in social psychology: Facts and speculations. *Cognition & Emotion*, **19**, 175-196. doi:10.1080/02699930441000274
- WASSERMAN, E. A. (1990). Detecting response-outcome relations: Toward an understanding of the causal texture of the environment. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 27-82). London: Academic Press.
- WASSERMAN, E. A., ELEK, S. M., CHATLOSH, D. L., & BAKER, A. G. (1993). Rating causal relations: Role of probability in judgments of response-outcome contingency. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 174-188.

NOTES

1. In the instructions that each participant saw, Color 1 and Color 2 represent whichever color combination the participant was given to signal each schedule of reinforcement.

2. Provisional analyses comparing the high BDI participants with the low BDI participants suggest that this might be the case. Please contact the authors if you would like more information about these data.

(Manuscript received November 5, 2009;
revision accepted for publication March 4, 2010.)