

Failure to obtain value enhancement by within-trial contrast in simultaneous and successive discriminations

JOANA ARANTES

University of Canterbury, Christchurch, New Zealand
and University of Minho, Braga, Portugal

AND

RANDOLPH C. GRACE

University of Canterbury, Christchurch, New Zealand

The present research tested the generality of the “work ethic” effect described by Clement, Feltus, Kaiser, and Zentall (2000). In Experiment 1, we trained 10 pigeons on a pair of either simultaneous or successive discriminations. One discrimination followed a high-effort requirement (20 pecks to the center key) and the other followed a low-effort requirement (1 peck). Contrary to Clement et al.’s results, we found that preferences between the S^+ and S^- stimuli in transfer tests depended on the event that initiated the trial: Pigeons preferred the stimulus from the baseline discrimination whose initiating event was most dissimilar from that preceding the test trial. Preferences were similar but less extreme in the successive condition. In Experiment 2, we investigated whether test preferences depended on the amount of training. A total of 12 pigeons were trained on a pair of simultaneous discriminations, except that test sessions were scheduled after every three baseline sessions. Preferences increased across test sessions but were similar to those in Experiment 1. Together with Vasconcelos, Urciuoli, and Lionello-DeNolf (2007a), our study represents a second failure to replicate Clement et al.’s work ethic effect. The finding that preference depends on the event that initiates the test trial suggests that choice probes may not provide unambiguous assessments of stimulus value.

A common view among researchers in learning and conditioning is that the associative strength or value of a stimulus varies inversely with respect to the context of reinforcement in which it occurs. This view provides an intuitively appealing account of a wide range of phenomena known collectively as *contrast effects* (Flaherty, 1996; Williams, 2002). For example, a stimulus that is correlated with a given rate of reinforcement will maintain a higher response rate when the reinforcement rate correlated with an alternative stimulus is relatively low than when it is relatively high (Reynolds, 1961; Williams, 1983). The elevation in response rate presumably reflects the increased value of the stimulus due to its occurrence in an overall lean context of reinforcement. Yet results of studies in which preference tests have been used to assess value independently of response rate have sometimes obtained paradoxical results (e.g., that positive contrast can be accompanied by decreased stimulus value in preference tests; Williams, 1991, 1992), and unraveling the complexities of these findings has proven to be a complex undertaking (Williams & McDevitt, 2001; see Williams, 2002, for review).

The focus of the present research is a particular type of contrast effect, defined as *within-trial contrast* by Zentall

(2005). The initial demonstration of this effect was a report by Clement, Feltus, Kaiser, and Zentall (2000). In their experiment, pigeons were trained on two simultaneous discriminations that differed in terms of the effort required. One discrimination required relatively little effort—a single response (FR1) to a center key—before the S^+ and S^- stimuli appeared (red and yellow illumination of the side keys), whereas the other discrimination required relatively high effort (20 center-key responses; FR20) before the S^+ and S^- stimuli appeared (green and blue illumination of the side keys). After pigeons were performing both discriminations to a high degree of accuracy, Clement et al. arranged a transfer test in which both S^+ or both S^- stimuli appeared (i.e., S_{FR1}^+ vs. S_{FR20}^+ or S_{FR1}^- vs. S_{FR20}^-). They found that the pigeons preferred both the S^+ and the S^- stimuli from the FR20 discrimination, and moreover, that this preference did not depend on the event that initiated the test trial (FR1 or FR20 response requirements on the center key, or no response requirement).

To explain these results, Clement et al. (2000) suggested that the value of a stimulus depended on the relative change in the animal’s hedonic state. Because the FR20 was a more effortful requirement, the appearance of the S_{FR20}^+ signaled

a relatively greater improvement in hedonic state than did the appearance of the S_{FR1}^+ . Thus, the value of the S_{FR20}^+ was enhanced by the relatively effortful response requirement that preceded it, leading to that stimulus being preferred over the S_{FR1}^+ during test. The corresponding preference for the S_{FR20}^- over S_{FR1}^- was explained as a process of value transfer from the S_{FR20}^+ to the S_{FR20}^- (see also von Fersen, Wynne, Delius, & Staddon, 1991; Zentall & Sherburne, 1994). Clement et al. described their results as a “work ethic” effect, because the pigeons preferred stimuli that were associated with increased response effort. Subsequent research has found evidence for similar value enhancement by within-trial contrast effects for stimuli that follow a delay versus those that do not follow a delay (DiGian, Friedrich, & Zentall, 2004), for stimuli that are associated with a low probability of alternative reinforcement versus those that are associated with a high probability of alternative reinforcement (Clement & Zentall, 2002), and for stimuli that follow nonreinforcement versus those that follow reinforcement (Friedrich, Clement, & Zentall, 2005).

The results of Clement et al. (2000) and subsequent studies are provocative, and Zentall (2005) suggested that within-trial contrast effects may be relevant for understanding a range of social psychological phenomena in humans (e.g., cognitive dissonance; Festinger, 1957). However, there are reasons to doubt whether Clement et al.’s original findings are reliable. Recently, Vasconcelos, Urcuioli, and Lionello-DeNolf (2007a) reported a series of unsuccessful attempts to replicate the preference for the stimulus preceded by a high-effort response requirement. They conducted five experiments that were modeled closely on Clement et al.’s design. The experiments differed in terms of the specific stimuli used for the S^+/S^- pairs; the center-key response requirement; whether the high-effort requirement was FR20 or FR40; whether the same Kodak Wratten filters used by Clement et al. were used to produce the stimuli; and whether the pigeons were experimentally naive. In each experiment, Vasconcelos et al. (2007a) found that preference for S_{FR20}^+ versus S_{FR1}^+ failed to deviate significantly from chance. Their results from S^- trials also generally failed to differ from chance levels, although Vasconcelos et al. (2007a) noted a tendency for the S_{FR1}^- to be preferred on test trials preceded by FR20, and for the S_{FR20}^- to be preferred on test trials preceded by FR1. Their consistent failure to replicate the preference for stimuli preceded by the high-effort response requirement led them to conclude that Clement et al.’s result may have derived from a Type I error.

In a commentary on Vasconcelos et al.’s (2007a) article, Zentall and Singer (2007) argued that the work ethic effect may only emerge slowly with extended training. Citing the results of Singer, Berry, and Zentall (2007), Friedrich and Zentall (2004), and their own unpublished data, they suggested that the work ethic effect could only be obtained reliably when pigeons were given sufficient training sessions. Specifically, they argued that pigeons might need 30 or more sessions of additional training once the baseline discriminations had been learned. In contrast, Vasconcelos et al.’s (2007a) pigeons only received 20 sessions of additional training prior to testing (note, however, that

in Clement et al.’s [2000] original demonstration, pigeons also received 20 sessions of overtraining). Thus, the status of the work ethic effect in pigeons—and hence the generality of within-trial contrast—remains controversial (see also Vasconcelos, Urcuioli, & Lionello-DeNolf, 2007b).

The possibility that within-trial contrast effects can modulate stimulus value may have implications for other areas of research, in particular the study of choice and conditioned reinforcement. Research in this area has often employed the concurrent-chains procedure, in which subjects respond during a choice phase (“initial links”) to produce access to one of two mutually exclusive outcomes (“terminal links”). Response allocation in the initial links is viewed as a measure of the relative value of the terminal-link stimuli as conditioned reinforcers (Grace, 2002). One of the most prominent accounts of choice and conditioned reinforcement is delay reduction theory (DRT; Fantino, 1969). According to DRT, the value of a stimulus as a conditioned reinforcer depends on the reduction in delay to reinforcement signaled by that stimulus relative to the overall average delay between reinforcers in the situation (see Fantino, 2001; Fantino, Preston, & Dunn, 1993; Fantino & Romanowich, 2007, for reviews). Specifically, a stimulus that signals a given delay to reinforcement has greater value if the duration of the preceding initial link is long rather than short. As Clement et al. (2000) noted, the work ethic effect is consistent with DRT if the FR20 response requirement is viewed as a relatively long delay.

Although DRT makes correct ordinal predictions for a variety of choice-related phenomena, subsequent models derived from the matching law (Herrnstein, 1961) have provided more accurate quantitative accounts of concurrent-chains performance (Grace, 1994; Mazur, 2001). In contrast to DRT, these models assume that the learned or conditioned value of a terminal-link stimulus is determined by the reinforcer delay signaled by that stimulus. Grace and Savastano (2000; see also Grace & Savastano, 1997) attempted to test directly whether the conditioned value of terminal-link stimuli depends on context by training pigeons on two different concurrent-chains procedures in each session. After baseline training, the relative value of terminal-link stimuli was assessed using different types of transfer tests. Grace and Savastano (2000) found that choice in the transfer tests was accurately predicted by the assumption that stimulus value was determined by the reinforcer delay signaled by the stimulus and unaffected by the duration of the prior initial link, contrary to DRT.

If reliable, the results of Clement et al. (2000) pose a dilemma: Why would stimulus value depend on the prior response requirement in a simultaneous discrimination, but apparently not depend on the duration of the initial link in concurrent chains? One major difference between the simultaneous discrimination task used in studies of within-trial contrast by Clement et al. and others and concurrent chains is that the discriminative stimuli are presented concurrently in the former, but successively in the latter. Note that although the initial-link stimuli occur simultaneously in concurrent chains, only one terminal-link stimulus is encountered on a given trial. In terms of how

the discriminative stimuli are presented, a concurrent-chains task resembles a successive rather than a simultaneous discrimination.

The initial goal of the present research was to test whether the work ethic effect reported by Clement et al. (2000) would be obtained regardless of whether pigeons were trained on simultaneous or successive discriminations in baseline. If within-trial contrast is a general phenomenon, then a similar work ethic effect should be obtained in both procedures. However, if the apparent discrepancy between the results of Clement et al. and Grace and Savastano (2000) is related to a fundamental difference in what is learned in simultaneous and successive discrimination training, then results might depend on the type of discrimination. Specifically, we anticipated that the work ethic effect would be replicated with simultaneous training, but that pigeons might be indifferent between S_{FR1}^+ and S_{FR20}^+ after exposure to successive discriminations. However, when, to our surprise, we failed to replicate the work ethic effect in Experiment 1 (we were unaware of Vasconcelos et al.'s [2007a] results until after the experiments reported here were completed), we conducted a second experiment to investigate whether preference on test trials might depend on amount of training.

EXPERIMENT 1

Experiment 1 consisted of two conditions, in which pigeons were trained on a pair of simultaneous or successive discriminations. The simultaneous condition was intended as a replication of Clement et al. (2000): Following completion of an FR1 or FR20 response requirement on the center key, the side keys were illuminated with different colors, and a single response to the designated S^+ produced reinforcement, whereas a response to S^- resulted in no reinforcement. The successive condition was similar, except that only one stimulus— S^+ or S^- , with equal probability—was presented following the center-key response requirement. Responses to S^+ produced reinforcement, whereas those to S^- produced none. After baseline training in each condition, transfer test trials were conducted in which both S^+ or both S^- stimuli were presented. As in Clement et al., transfer tri-

als were equally likely to be preceded by an FR1, FR20, or no response requirement on the center key.

Method

Subjects

Ten pigeons (*Columba livia*) were maintained at 85% of their free-feeding weights ± 15 g through appropriate postsession feeding. Pigeons were housed individually in a vivarium with a 12:12 h light:dark cycle (lights on at 07:00 h). All pigeons had prior experimental histories; 6 had previously served in experiments using concurrent-chains schedules, and 4 in experiments using delayed matching to sample. However, the pigeons had not been exposed to the white cross stimulus used for the initial response requirement. Water and grit were freely available in the home cages.

Apparatus

Four standard three-key operant chambers, 350 mm deep \times 360 mm wide \times 350 mm high, were used. The keys were 260 mm above the floor and arranged in a row. The side keys could be illuminated from behind with red, green, white, or blue lights, and the center key could be illuminated with a white cross superimposed on a black background. In each chamber, a houselight was located 70 mm above the center key and a grain magazine with an aperture (60 \times 50 mm) was 130 mm below the center key. The magazine was illuminated when wheat was made available. A force of approximately 0.10 N was necessary to operate each key, and effective responses produced an audible feedback click. Each chamber was enclosed in a sound-attenuating box, and ventilation and white noise were provided by an attached fan. Event scheduling and data recording were controlled with a MEDSTATE notation program and a MED-PC system interfaced to an IBM-compatible microcomputer located in an adjacent room.

Procedure

Sessions were conducted 7 days a week at approximately the same time of day, with few exceptions. Because all pigeons had prior experimental histories, training began immediately in the first condition. However, some pigeons did not respond reliably to the center-key stimulus (white cross on black background). These birds were trained to peck the center key using a modified autoshaping procedure, and then began the first condition.

The experiment consisted of two conditions, in which each pigeon was trained on two simultaneous and on two successive discriminations. Half of the pigeons completed the simultaneous condition first, followed by the successive condition, whereas the order was reversed for the other half of the pigeons. Because the birds that initially did not peck the white cross required different amounts of pretraining, the number of baseline sessions in the first condition

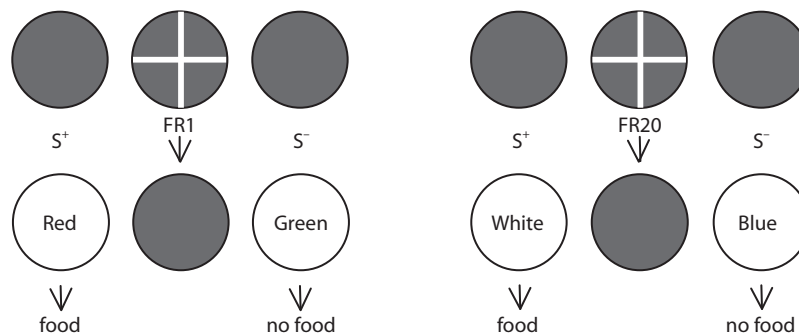


Figure 1. Structure of training trials in the simultaneous condition of Experiment 1. On each trial, 1 peck (FR1) or 20 pecks (FR20) to the center key produced the discriminative stimuli. In the arrangement shown here, a red-key choice is rewarded following the FR1, and a white-key choice following the FR20.

varied across subjects but was at least 31 sessions. After baseline training had been completed, 5 transfer test sessions were conducted (described below). Baseline training in the second condition then began, again followed by 5 transfer test sessions. All pigeons received at least 31 baseline sessions in the second condition. The average number of baseline sessions completed in both conditions was 37.2 (simultaneous range, 31–55; successive range, 31–55).

Simultaneous condition: Baseline training. Sessions consisted of 96 trials. Each trial was preceded by an intertrial interval (ITI) of 10 sec, during which the houselight was on and all keylights were dark. After the ITI, the houselight was turned off and the center key was illuminated with a white cross superimposed on a dark background. There were two types of trials, as shown in Figure 1. Whether red and white served as the S^+ stimuli (and green and blue as the S^-) and whether the red/green or white/blue discrimination was preceded by FR1 were counterbalanced across pigeons. However, for clarity, we describe the procedure as if all birds were trained according to Figure 1. On half of the trials, one response (FR1) to the center key extinguished its key light and lighted the side keys red and green (S^+_{FR1} vs. S^-_{FR1}); on the remaining trials, 20 responses (FR20) were required before the center key was turned off, and the side keys were lighted white and blue (S^+_{FR20} vs. S^-_{FR20}). After the side keys were lighted, a single response to either one turned off the alternate key and began a

6-sec delay, during which responses to the selected key were recorded. If the red or white key was chosen (S^+), the hopper was activated for 3 sec at the end of the delay, followed by the next ITI. If the green or blue key was chosen (S^-), the ITI started immediately at the end of the 6-sec duration. For each trial, whether the response requirement for the center key was FR1 or FR20, and whether the red and green (or blue and white) stimuli were located on the left and right keys or vice versa, was determined randomly, subject to the constraints that equal numbers of FR1 and FR20 trials occurred in each session and each color appeared equally often on each side key. The trials were canceled and omitted from subsequent analyses if the pigeon did not complete the FR1 or FR20 requirement within 2 min, or if it failed to respond to one of the lighted side keys within 10 sec.

Successive condition: Baseline training. All details of the procedure remained the same as in the simultaneous condition, with the following exceptions. After the center key response requirement was completed, only one side key was lighted. The stimulus assignments for all pigeons were the same as in the simultaneous condition. On half of the trials, the S^+ was presented, whereas the S^- was presented on the other half. The S^+ and S^- stimuli appeared equally often on the left and the right side. Thus, completion of the FR1 response requirement was followed by red illumination of the left or right key on 24 trials and green illumination of the left or right

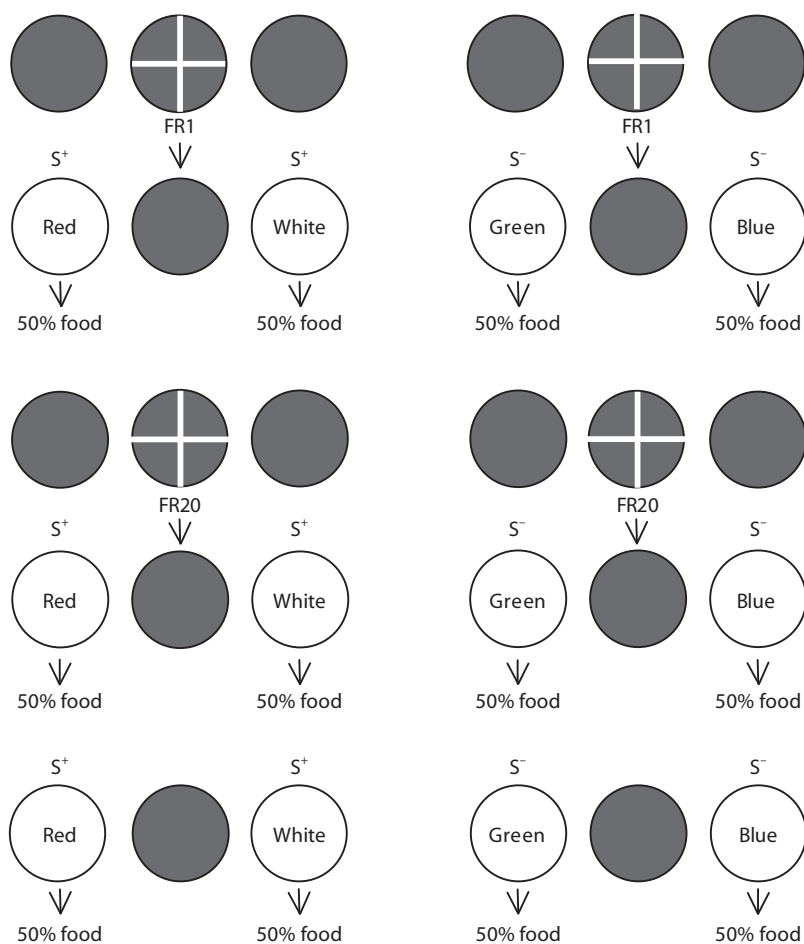


Figure 2. Structure of test trials in the simultaneous condition of Experiment 1. There were six types of test trials: those requiring 1 peck at the center key (FR1) followed by choice between both S^+ or both S^- stimuli (upper panels); those requiring 20 pecks (FR20) followed by choice between both S^+ or both S^- stimuli (middle panels); and those with no prior stimuli before choice between both S^+ or both S^- stimuli (lower panels).

key on 24 trials (cf. Figure 1). Similarly, completion of the FR20 requirement led to white or blue illumination of a side key on 24 trials each. If the stimulus was an S⁺, the first side-key response produced reinforcement (3-sec access to food) after a 6-sec delay. If the stimulus was an S⁻, the trial ended after 6 sec independently of the pigeons' behavior, and no food was presented.

Transfer test sessions. Each test session included 36 test trials randomly intermixed with 60 baseline trials. The contingencies on baseline trials remained exactly the same as during the baseline sessions preceding the test. There were six types of test trials, as shown in Figure 2, depending on the initial event that preceded the choice stimuli (FR1, FR20, or no response requirement) and on whether both S⁺ stimuli (S⁺_{FR1} and S⁺_{FR20}) or both S⁻ stimuli (S⁻_{FR1} and S⁻_{FR20}) were presented during the choice phase. On test trials with FR1 or FR20 response requirements, the white cross was presented on the center key, and after the appropriate number of responses had been made, the side keys were illuminated. For test trials with no response requirement, the side keys were illuminated directly after the ITI. Choice responses during test trials were reinforced randomly. After the first response to either side key, both side keys were turned off, and either 3-sec access to food was presented or the next ITI began, with 50% probability. Each type of test trial occurred six times during test sessions, and the locations of the stimuli were counterbalanced.

Results and Discussion

Baseline Training

All pigeons learned the basic discriminations, S⁺_{FR1} versus S⁻_{FR1} and S⁺_{FR20} versus S⁻_{FR20}, in both simultaneous and successive conditions. Fewer sessions were required overall to reach criterion (90% correct responses) in the simultaneous condition ($M = 2.1$ sessions) than in the successive condition (criterion of 90% discrimination ratio, $M = 9.8$ sessions) [$t(9) = -3.44, p < .01$]. A repeated measures ANOVA with condition and response requirement (FR1 or FR20) as factors found that asymptotic performance (defined as the average proportion correct or discrimination ratio during the last 5 baseline sessions) was greater in the simultaneous condition ($M = 99.6%$) than in the successive condition ($M = 94.6%$) [$F(1,9) = 14.58, p < .01$]. The effect of response requirement and the condition \times response requirement interaction were not significant. Thus, the simultaneous discriminations were learned more rapidly and to a higher degree of accuracy than were the successive discriminations, and the response requirement of a trial had no systematic effect on performance.

Transfer Testing

Figure 3 shows the results from the transfer test sessions for both conditions. Individual data from Figure 3 were entered into a repeated measures ANOVA with condition, stimulus type (S⁺ or S⁻), and initiating event (no response requirement, FR1, or FR20) as within-subjects factors and order of condition (simultaneous–successive or vice versa) as a between-subjects factor. There was a significant effect of initiating event [$F(2,16) = 19.91, p < .001$], indicating that choice for the FR20 stimulus was different depending on whether the test trial began with no response requirement ($M = 58.9%$) or an FR1 ($M = 71.5%$) or FR20 ($M = 39.1%$) on the center key. The effect of stimulus type was significant [$F(1,8) = 6.72, p < .05$], showing that choice for the FR20 stimulus was greater on

S⁻ trials ($M = 62.4%$) than on S⁺ trials ($M = 50.6%$). The interaction between condition and initiating event was also significant [$F(2,16) = 6.47, p < .01$]. This interaction resulted from the preference for the FR20 stimulus being greater in the simultaneous than in the successive condition for trials initiated by FR1 ($M_s = 75.0%$ and $67.9%$, respectively), but less in the former condition for trials initiated by FR20 ($M_s = 33.6%$ and $44.7%$, respectively). Thus, preference for the FR20 or the FR1 stimulus in the trials preceded by FR1 and FR20, respectively, was more extreme in the simultaneous than in the successive condition. The main effects of condition (simultaneous or successive) and order of condition, and all the remaining interactions, failed to reach significance (all $p_s > .10$).

Next, we assessed whether the preferences for the FR20 stimulus deviated significantly from chance (i.e., 50%). For the simultaneous condition, there was a significant preference for S⁻_{FR20} over S⁻_{FR1} for trials preceded by FR1 and by no response requirement [$M_s = 63.1%$ and $81.2%$; $t(10)s = 2.73$ and $5.96, p_s < .05$ and $.001$, respectively], but not for trials preceded by FR20 ($M = 40.7%, p < .05$).

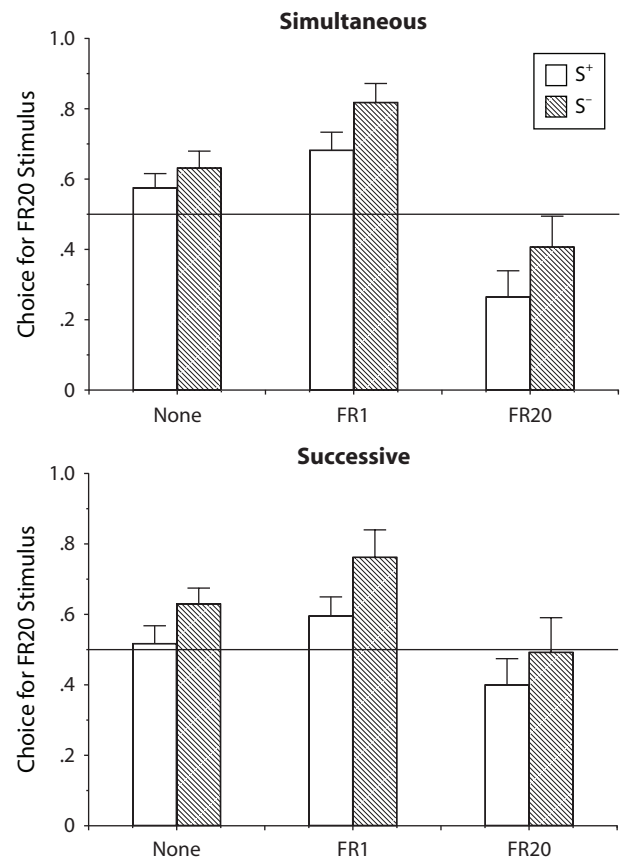


Figure 3. Mean proportions of choices for the FR20 stimuli during test in Experiment 1, when there was no initial stimulus (None) or the test trial required 1 peck (FR1) or 20 pecks (FR20) to be initiated. Unfilled bars show the results from the S⁺ test trials; filled bars show the results from the S⁻ test trials. Error bars indicate standard errors of the means for each condition. Upper panel: Simultaneous condition. Lower panel: Successive condition.

There was a significant preference for the S_{FR20}^+ over S_{FR1}^+ during trials preceded by FR1 [$M = 68.2\%$; $t(9) = 3.58$, $p < .01$], and during trials preceded by no requirement that preference approached significance [$M = 57.6\%$; $t(9) = 1.91$, $p < .09$]. However, there was a significant preference for S_{FR1}^+ over S_{FR20}^+ during trials preceded by FR20 [$M = 26.5\%$; $t(9) = -3.17$, $p < .05$].

Preferences for the FR20 stimulus in the successive condition were always in the same direction as in the simultaneous condition, but generally less extreme. The only results that were significantly different from 50% were the preferences for S_{FR20}^- over S_{FR1}^- on trials preceded by FR1 or by no response requirement [$M_s = 76.2\%$ and 63.0% ; $t(10)s = 3.39$ and 2.92 , $p_s < .01$ and $.05$, respectively].

These results may be summarized as follows. During transfer testing with S^+ and S^- stimuli from different baseline discriminations, pigeons generally preferred the stimulus associated with the discrimination whose initiating event was most dissimilar from the event that preceded the test trial. That is, stimuli that were preceded by FR20 in baseline were generally preferred on test trials that were initiated by no response or by FR1, whereas stimuli that were preceded by FR1 in baseline were preferred on test trials initiated by FR20. In addition, preferences deviated more from chance for S^- than for S^+ trials, and for the simultaneous as compared with the successive condition.

Our results from Experiment 1 differ in several ways from those reported by Clement et al. (2000) and Vasconcelos et al. (2007a). First, an overall "work ethic" effect was not obtained: We found significant preferences for both the FR20 and the FR1 stimuli, depending on the initiating event, and that the preference depended on stimulus type (S^+ or S^-), whereas Clement et al. found a significant preference for the FR20 stimulus regardless of the initiating event or stimulus type. Thus, together with Vasconcelos et al. (2007a), our results constitute a second independent failure to replicate the work ethic effect. Vasconcelos et al. (2007a), on the other hand, also failed to find significant effects of initiating event on preference during test. However, their data showed some tendency for the FR20 stimulus preference to be stronger on S^- trials (see their Figures 2 and 3, upper panel), although the results did not reach significance. Similarly, Clement et al. reported that preference for the FR20 stimulus was greater on S^- trials, although not significantly greater than on S^+ trials.

One likely reason that trends obtained by Clement et al. (2000) and Vasconcelos et al. (2007a) were significant in our study was the overall greater number of baseline training sessions that we conducted. In their commentary to Vasconcelos et al.'s (2007a) article, Zentall and Singer (2007) noted that the work ethic effect may require a substantial number of training sessions before it can be observed reliably, and they suggested that Vasconcelos et al. (2007a) may have failed to observe the effect because they only provided 20 additional baseline sessions after the pigeons had acquired the initial discriminations. Because Vasconcelos et al.'s (2007a) pigeons generally required between 2 and 3 sessions to reach criterion, the total amount of baseline training was only 22–23 sessions

(similar to the conditions in Clement et al., 2000). In contrast, our pigeons acquired the discriminations about as rapidly as Vasconcelos et al.'s (2007a) pigeons, but they received at least 31 baseline sessions prior to testing.

The original goal of Experiment 1 was to determine whether the work ethic effect depended on whether the pigeons had been baseline trained on simultaneous or successive discriminations. However, our failure to obtain a consistent preference for the S^+ or S^- preceded by FR20 renders that question moot. Nevertheless, the results are still informative in terms of providing a comparison between the two discrimination types. Overall, results were similar between the two conditions, although preferences for one stimulus or the other were generally more extreme in the simultaneous condition. This suggests that stimulus interactions may have played a role in probe preference, because such interactions are stronger for stimuli that are presented simultaneously rather than successively (see, e.g., Herrnstein, 1970). Our results were also similar regardless of condition order. This is important because the status of a stimulus as S^+ or S^- remained the same in both conditions. If what the pigeons learned about the stimuli were different for the two discrimination types, then one might expect that the results would depend on condition order. The fact that they did not, coupled with the overall consistency of the results in both conditions, suggests that the content of learning was similar for both simultaneous and successive discriminations.

EXPERIMENT 2

Despite our failure to replicate Clement et al.'s (2000) work ethic effect, the results of Experiment 1 suggested that preferences on test trials might depend on the initiating event. In Experiment 2, we explored whether test preference depended on the amount of training. Pigeons were exposed to the same baseline procedure as in the simultaneous condition in Experiment 1, but after every block of 3 baseline sessions a test session was conducted. All pigeons completed a total of 9 test sessions. If performance depends on amount of training, we should then observe a systematic change in preference across test sessions. In addition, a subgroup of 4 pigeons was tested again after extended baseline training (up to 112 sessions of cumulative exposure to the procedure).

Method

Subjects

Twelve pigeons (*Columba livia*) served in Experiment 2. Of these, 4 had also served in Experiment 1, although their participation in Experiment 2 was separated from the earlier experiment by a period of 4 months, during which they were involved in an unrelated experiment using a concurrent-chains procedure. Of the remaining 8 pigeons, 4 had previously served in an experiment using concurrent chains, and 4 (the extended training subgroup) had trained on a counting-related task. The 8 pigeons that had not served in Experiment 1 had not been previously exposed to the white cross stimulus used to signal the center-key response requirement. All pigeons were housed and maintained according to the same conditions as in Experiment 1.

Apparatus

The apparatus was the same as in Experiment 1.

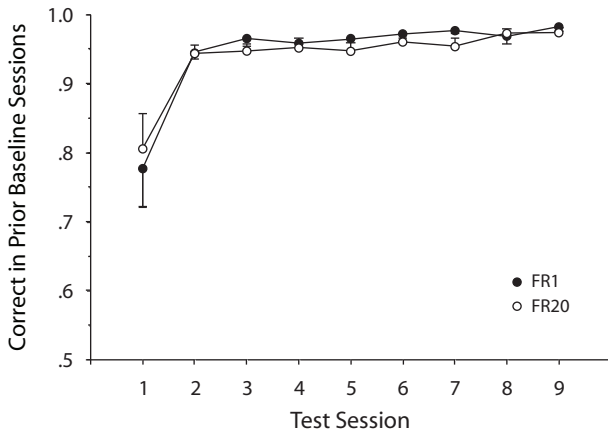


Figure 4. Mean proportions of correct responses for the discriminations preceded by FR1 and by FR20 for the three training sessions that immediately preceded each test session in Experiment 2. Error bars show the standard errors of the means for each condition.

Procedure

Because all pigeons had prior experimental histories, they were placed immediately in the final procedure. Some did not respond reliably and were given pretraining similar to that in Experiment 1. All details of the baseline procedure and the stimulus counterbalancing were the same as in the simultaneous condition of Experiment 1. For the 4 pigeons that had served in Experiment 1, the colors were assigned such that both the S⁺/S⁻ designation within each pair (red/green and white/blue) and the response requirement for each pair were reversed in comparison with their previous assignments.

With few exceptions, test sessions were conducted after every third baseline session. All pigeons completed at least 9 test sessions. Test sessions were identical to those used in the simultaneous condition in Experiment 1. Pigeons received at least 27 intervening baseline sessions, for a minimum of 36 sessions total.

A randomly selected subgroup of pigeons continued to receive baseline training after the 9 test sessions had been completed. These pigeons were then given a block of 5 consecutive test sessions. The cumulative number of sessions these pigeons had received prior to the final block was 76, 76, 102, and 112 sessions. These varied because the birds required different amounts of pretraining before they reliably pecked the center-key stimulus.

Results and Discussion

We checked whether performance differed between the pigeons that had participated in Experiment 1 and those that had not. The results were similar for the two groups, and the outcomes of all statistical tests were unchanged when the data from the pigeons that had participated in Experiment 1 were omitted. Thus, only data from the full group of 12 pigeons are reported below.

Figure 4 shows the proportions of correct responses for the discriminations preceded by both FR20 and FR1 center-key response requirements for the three baseline sessions immediately preceding each test session. Accuracy was near 80% for both discriminations prior to the first test session, and thereafter remained stable at levels between 95% and 98%. A repeated measures ANOVA found a significant effect of test session [$F(8,88) = 11.64, p < .001$], but the effect of response requirement and the test session \times response requirement interaction were not

significant. Pigeons learned both discriminations rapidly, and high levels of performance were maintained across the experiment.

Results from the test sessions are displayed in Figure 5. The upper panel shows data from S⁺ trials, and the bottom panel shows data from S⁻ trials. A repeated measures ANOVA found a significant effect of stimulus type (S⁺ or S⁻) [$F(1,11) = 7.32, p < .05$], indicating that preference for the FR20 stimulus was greater overall on S⁻ trials than on S⁺ trials. The effect of initiating event (i.e., whether the test trial was preceded by no response requirement, FR1, or FR20) was significant [$F(2,22) = 13.73, p < .001$], confirming that preference for the FR20 stimulus was greater on trials preceded by no response requirement or FR1. There was also a significant interaction between test session number and initiating event [$F(16,176) = 3.80, p < .001$]. This interaction resulted because preference for the FR20 stimulus decreased on trials preceded by FR20, but increased or remained approximately constant for trials preceded by no requirement or FR1. Thus, the

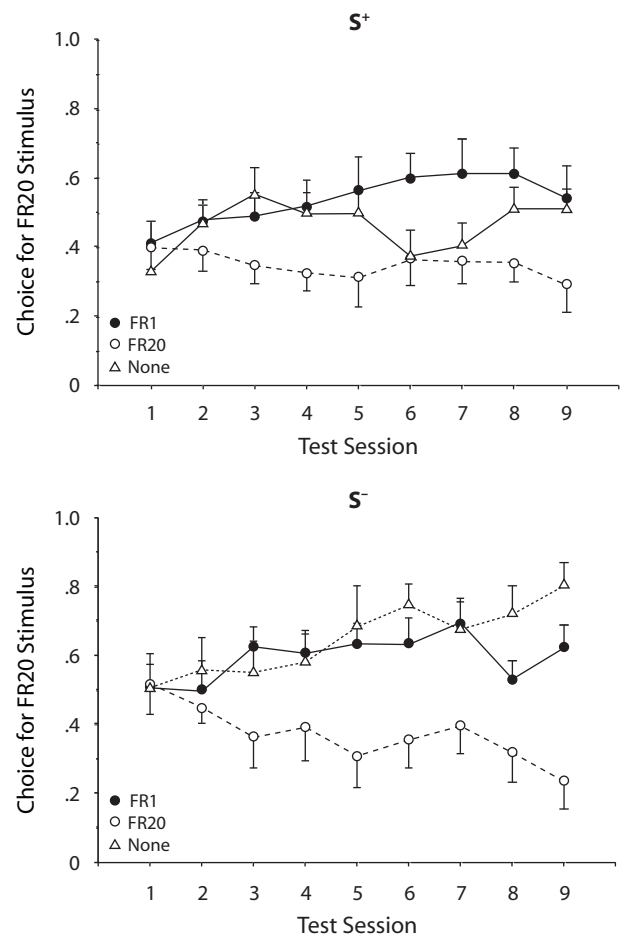


Figure 5. Mean proportions of choices for the FR20 stimuli by test session in Experiment 2. Results for trials preceded by FR1, FR20, or no response requirement are marked as noted in the legend. Error bars indicate the standard errors of the means. Upper panel: Results from S⁺ trials. Lower panel: Results from S⁻ trials.

significant interaction confirms that performance on test trials depends on the amount of training. Neither the main effect of test number nor any of the other interactions were significant (all p s > .20).

Measures of the final preference levels were obtained for each pigeon by averaging across the last three test sessions. The resulting preferences for the FR20 stimulus are shown in the upper panel of Figure 6. There was a significant preference for S_{FR20}^- over S_{FR1}^- for trials preceded by FR1 [$M = 73.7\%$; $t(11) = 3.64, p < .01$], and that for trials preceded by no requirement approached significance [$M = 61.5\%$; $t(11) = 2.19, p = .05$]. The preference was significant in the opposite direction (i.e., for S_{FR1}^- over S_{FR20}^-) on trials preceded by FR20 [$M = 31.7\%$; $t(11) = -2.33, p < .05$]. For S^+ trials, the only preference that deviated significantly from chance was for S_{FR1}^+ over S_{FR20}^+ on trials preceded by FR20 [$M = 33.4\%$; $t(11) = -2.94, p < .05$].

Results for the subgroup of pigeons given extended training are shown in the bottom panel of Figure 6. A repeated measures ANOVA with initiating event and S^+ / S^- as factors found a significant effect of initiating event

[none, FR1, or FR20; $F(2,6) = 25.91, p < .01$], but neither the main effect of S^+ / S^- nor the interaction was significant. For each type of test trial, the average preference deviated from indifference in the same direction as the data for the full group in the upper panel of Figure 6, except that preferences were generally more extreme. With the exception of S^- trials preceded by FR20, for which the results in the upper and lower panels of Figure 6 were the same ($M = 32\%$), the preference for the FR20 stimulus (or for the FR1 stimulus) in the lower panel was more extreme than in the upper panel. This provides additional evidence that the test preferences become more extreme with increased amounts of training.

Overall, these results are consistent with those from the simultaneous condition in Experiment 1 and confirm that probe preference develops gradually with increased training. Thus, they provide some support for Zentall and Singer's (2007) claim that extensive training is required before test preferences that deviate significantly from chance can be observed. However, as in Experiment 1, we found no evidence of a general work ethic effect—that is, a consistent preference for the stimulus preceded by FR20—but instead found that choice on test trials depended on the type of initiating event. In Experiment 2, we again found that the only significant preferences were away from the stimulus associated with the baseline discrimination whose initiating event was most similar to that on the test trial.

GENERAL DISCUSSION

The major goals of the present research were to determine whether the work ethic effect reported by Clement et al. (2000) depended on whether subjects were baseline trained on simultaneous or successive discriminations (Experiment 1) and to investigate the effects of amount of training on the preference for the high-effort stimulus (Experiment 2). Neither Experiment 1 nor 2 replicated the work ethic effect reported by Clement et al.: Our results did not show a consistent preference for the high-effort stimuli. Instead, we found that for both S^+ and S^- test trials, preference for the FR20 stimulus depended on the initiating event that preceded the choice stimuli. Whereas pigeons generally preferred the FR20 stimulus on trials preceded by FR1 or no response requirement, they responded more to the FR1 stimulus on trials preceded by FR20. In other words, pigeons chose the stimulus whose initiating event in baseline was most dissimilar from that which preceded the test trial. Experiment 2 showed that preferences on test trials became more extreme as amount of training increased. Moreover, in both experiments, preferences deviated more from chance on S^- than on S^+ trials. Whereas the results of Experiment 2 are consistent with Zentall and Singer's (2007) suggestion that amount of training is a relevant variable, they provide no support for the claim that the work ethic effect is more likely to be obtained when pigeons have had extensive training. We found that test preferences diverged more from indifference as the cumulative exposure to baseline and test sessions increased, but the direction of preference depended on the initiating event, just as in Experiment 1.

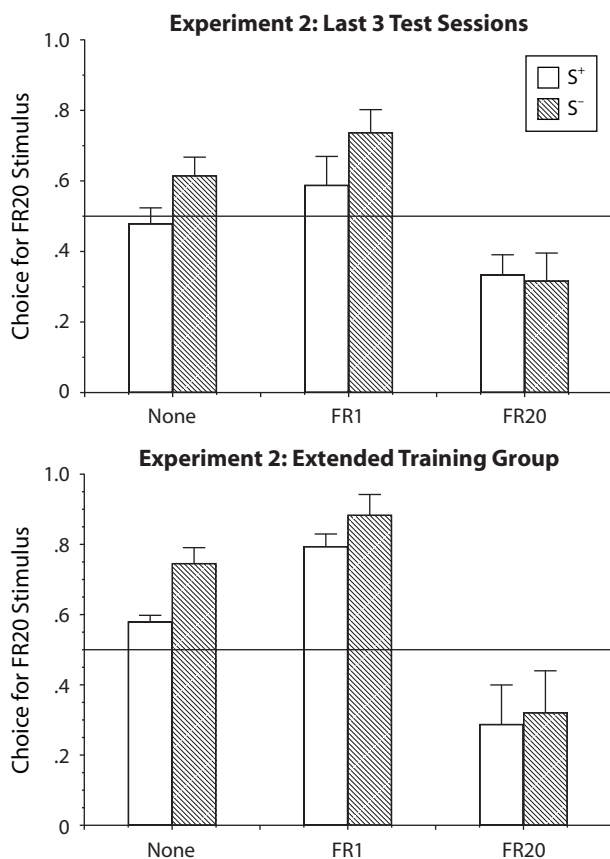


Figure 6. Mean proportions of choices for the FR20 stimuli during test in Experiment 2, when there was no initial stimulus (None) or the test trial required 1 peck (FR1) or 20 pecks (FR20) to be initiated. Unfilled bars show the results from the S^+ test trials; filled bars show the results from the S^- test trials. Error bars indicate the standard errors of the means for each condition. The top panel shows the data from the full group of 12 pigeons averaged across the final three test sessions (7–9); the bottom panel shows the data from the subgroup of pigeons that received extended training.

Our failure to replicate the work ethic effect is consistent with Vasconcelos et al. (2007a). In addition to failing to obtain an overall preference for the FR20 stimuli, their data showed a tendency for the S_{FR1}^- to be preferred on test trials preceded by FR20, and for the S_{FR20}^- to be preferred on test trials preceded by FR1 or by no initial stimulus. Thus, there was some indication in Vasconcelos et al.'s (2007a) data that preferences depended on the initiating event, in a manner similar to that found in the present study. However, their results generally failed to reach statistical significance.

There are several possible reasons why we may have obtained stronger results than Vasconcelos et al. (2007a) in terms of magnitude of preference on test trials. One likely factor is that our pigeons received more training overall—an average of 37.2 baseline sessions in Experiment 1 and up to 98 in Experiment 2, rather than 22–23 sessions in their study. The results of Experiment 2 confirmed that test preferences became more extreme (in terms of deviation from indifference) with increased training, so the greater number of baseline sessions our pigeons received would be expected to generate stronger preferences than those observed by Vasconcelos et al. (2007a).

Another factor that may have influenced the results is the reinforcement contingencies on test sessions. In the present experiments as well as those of Vasconcelos et al. (2007a) and Clement et al. (2000), a single response was necessary to terminate a probe trial, and reinforcement was delivered with 50% probability for both choices. The advantage of random, nondifferential reinforcement is that responding can be maintained indefinitely over multiple test sessions, as compared with probe trials that always end in nonreinforcement (cf. Grace & Savastano, 2000). There is no reason to expect that such reinforcement would bias responding one way or the other, so any preferences found can be interpreted as baseline transfer. Even if a feedback relation were present, such that probe preferences were strengthened across multiple test sessions by the nondifferential reinforcement contingencies, the initial preference would have to reflect the effects of baseline training. Thus, the probe procedure used in research on the work ethic effect should represent a valid assay of baseline transfer.

One procedural difference is that outcomes (presence or absence of reinforcement) for responses on test trials occurred immediately in our experiments, but after a 6-sec delay in Vasconcelos et al. (2007a). Thus, any feedback relation or strengthening effect of reinforcement on test trials, if present, would have been greater in the present study. Although this procedural difference was unintended,¹ it may have contributed to the stronger test preferences observed here. However, it is important to note that our use of immediate outcomes on test trials cannot explain the direction of the observed preferences, which must represent transfer from baseline.

There have now been at least two independent failures to replicate the work ethic effect in pigeons. The results of the present study and Vasconcelos et al. (2007a) were generally consistent and were based on larger sample sizes (pooled across experiments, 22 pigeons in the former and

20 in the latter) than those of Clement et al. (2000; 8 pigeons), and our Experiment 2 provided evidence against Zentall and Singer's (2007) rejoinder that the work ethic effect can be reliably obtained with sufficient training. Thus, it seems increasingly likely that Clement et al.'s original results may have derived from a Type I error.

If the value of a stimulus in a simultaneous discrimination is not enhanced by following a relatively effortful response requirement, the apparent discrepancy noted in the introduction between the work ethic effect in pigeons and the findings of Grace and Savastano (2000) is resolved. However, there is evidence that after training on multiple-chain schedules, a preference does emerge for the terminal-link stimulus that follows a longer initial link (O'Daly, Angulo, Gipson, & Fantino, 2006; O'Daly, Meyer, & Fantino, 2005). Exactly why transfer tests of relative value should differ after training on concurrent or multiple chains is unclear, and an answer must await future research.

The major empirical finding from our study is that the pigeons generally preferred the stimulus on test trials whose initiating event in baseline was most dissimilar from the event that preceded the test trial. How is this result to be understood? In discussing possible outcomes for their experiment, Clement et al. (2000) noted that if the center-key response requirement served an occasion-setting function on baseline trials (Ross & Holland, 1981), pigeons might prefer S_{FR1}^+ on test trials that began with FR1 and S_{FR20}^+ on test trials that began with FR20. Although this prediction is the opposite of what we obtained, our data do suggest that the response requirement provided part of the context in which the stimuli occurred (Balsam, 1985).

A possible explanation for the results on test trials involving the S^- stimuli was suggested by Vasconcelos et al. (2007a). They noted that instead of learning to peck the S^+ in baseline, pigeons may have learned not to peck the S^- . That is, after completing the center-key response requirement, the pigeons may have looked first to avoid the S^- and then responded to the S^+ . If pigeons learned not to respond to the S_{FR1}^- after the FR1 center-key requirement, and not to respond to the S_{FR20}^- after the FR20, they should have shown a preference for the S_{FR20}^- on trials that began with FR1 or no response requirement (because presumably no response requirement is more similar to FR1 than to FR20), and for the S_{FR1}^- on trials that began with FR20. However, this explanation fails to account for the S^+ results, which were generally in the same direction (if less extreme) than the S^- results.

Why would the pigeons have preferred the S^+ stimulus from the baseline discrimination whose initiating event was most dissimilar from that for the test trial? One possible explanation for this counterintuitive finding is that a process similar to choice by exclusion (Clement & Zentall, 2003) may have influenced responding on S^+ test trials. In Clement and Zentall's (2003) study, pigeons were trained on a matching-to-sample task and later exposed to a novel sample-comparison choice. Clement and Zentall (2003) found that pigeons were likely to reject a comparison stimulus that was already associated with a different sample, choosing instead an alternative stimulus that did not have such an association. To apply this result to the

present study, two assumptions are necessary: (1) that the center-key response requirement might have activated a particular S^+ association and (2) that the S^- constituted part of the context for S^+ responding on baseline trials. For example, consider an S^+ test trial preceded by an FR1 response requirement. Completion of the FR1 should activate the association for S^+_{FR1} , but the absence of the S^-_{FR1} (and presence of S^+_{FR20}) signals a change from the baseline context for S^+_{FR1} responding. To the extent that this change in context can be considered a novel sample, choice by exclusion should increase the likelihood of responding to the S^+_{FR20} . A parallel argument would apply to explaining choice of S^+_{FR1} on FR20 test trials. Thus, choice of the S^+ whose baseline initiating event was most dissimilar from the one that preceded the test trials might have resulted from choice by exclusion. However, why this mechanism should have overridden the presumably strong tendency to respond to the S^+ whose association was activated by the response requirement is unclear.

Whether choice by exclusion or some other process can account for the results on S^+ trials, our findings suggest that the conventional interpretation that transfer tests provide an unambiguous assessment of stimulus value may be mistaken. According to this view, stimuli acquire value during baseline training, and when stimuli from different baseline contexts are presented together in transfer tests, response allocation reflects differences in relative value. In other words, this interpretation assumes that value is a valid construct that can be measured outside of the context in which it is acquired. Our results provide a strong challenge to this view by showing that the relative “values” of S^+_{FR20} and S^+_{FR1} (and of S^-_{FR20} and S^-_{FR1}) depend on the event that initiated the transfer trial. This means either that factors other than value influence responding on transfer tests or that value may have limited validity as a theoretical construct for understanding the present results. Regardless of which of these possibilities (or both) is correct, a clear implication is that researchers need to interpret the results of transfer tests with caution.

A conservative methodology would be always to include different types of transfer tests involving the same stimuli in individual experiments, effectively providing converging operations to measure stimulus value (Garner, Hake, & Eriksen, 1956; Grace, 2001). If results from different transfer tests involving the same stimuli are in agreement, we can be more confident about conclusions regarding the relative values of the stimuli in question. However, when results of different transfer tests are in conflict—as was obtained in the present experiments—we must question the assumptions described above. A task for future research will be to identify more precisely the variables that control responding under different transfer conditions, and thereby hopefully clarify the scope and status of value as an explanatory construct for the understanding of learned behavior.

AUTHOR NOTE

This research was presented at the 4th ABA International Conference (Sydney, August 2007). The first author was supported by a PhD fellowship from the Portuguese Foundation for Science and Technology (FCT).

Correspondence and reprint requests may be addressed to J. Arantes or R. C. Grace, University of Canterbury, Department of Psychology, Private Bag 4800, Christchurch, New Zealand (e-mail: joana.arantes@canterbury.ac.nz, randolph.grace@canterbury.ac.nz).

REFERENCES

- BALSAM, P. D. (1985). The functions of context in learning and performance. In P. D. Balsam & A. Tomie (Eds.), *Context and learning* (pp. 1-22). Hillsdale, NJ: Erlbaum.
- CLEMENT, T. S., FELTUS, J. R., KAISER, D. H., & ZENTALL, T. R. (2000). “Work ethic” in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review*, *7*, 100-106.
- CLEMENT, T. S., & ZENTALL, T. R. (2002). Second-order contrast based on the expectation of effort and reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*, 64-74.
- CLEMENT, T. S., & ZENTALL, T. R. (2003). Choice based on exclusion in pigeons. *Psychonomic Bulletin & Review*, *10*, 959-964.
- DIGIAN, K. A., FRIEDRICH, A. M., & ZENTALL, T. R. (2004). Discriminative stimuli that follow a delay have added value for pigeons. *Psychonomic Bulletin & Review*, *11*, 889-895.
- FANTINO, E. (1969). Choice and rate of reinforcement. *Journal of the Experimental Analysis of Behavior*, *12*, 723-730.
- FANTINO, E. (2001). Context: A central concept. *Behavioural Processes*, *54*, 95-110.
- FANTINO, E., PRESTON, R. A., & DUNN, R. (1993). Delay reduction: Current status. *Journal of the Experimental Analysis of Behavior*, *60*, 159-169.
- FANTINO, E., & ROMANOWICH, P. (2007). The effect of conditioned reinforcement rate on choice: A review. *Journal of the Experimental Analysis of Behavior*, *87*, 409-421.
- FESTINGER, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- FLAHERTY, C. F. (1996). *Incentive relativity*. Cambridge: Cambridge University Press.
- FRIEDRICH, A. M., CLEMENT, T. S., & ZENTALL, T. R. (2005). Discriminative stimuli that follow the absence of reinforcement are preferred by pigeons over those that follow reinforcement. *Learning & Behavior*, *33*, 337-342.
- FRIEDRICH, A. M., & ZENTALL, T. R. (2004). Pigeons shift their preference toward locations of food that take more effort to obtain. *Behavioural Processes*, *67*, 405-415.
- GARNER, W. R., HAKE, H. W., & ERIKSEN, C. W. (1956). Operationism and the concept of perception. *Psychological Review*, *63*, 149-159.
- GRACE, R. C. (1994). A contextual model of concurrent-chains choice. *Journal of the Experimental Analysis of Behavior*, *61*, 113-129.
- GRACE, R. C. (2001). On the failure of operationism. *Theory & Psychology*, *11*, 5-33.
- GRACE, R. C. (2002). The value hypothesis and acquisition of preference in concurrent chains. *Animal Learning & Behavior*, *30*, 21-33.
- GRACE, R. C., & SAVASTANO, H. I. (1997). Transfer tests of stimulus value in concurrent chains. *Journal of the Experimental Analysis of Behavior*, *68*, 93-115.
- GRACE, R. C., & SAVASTANO, H. I. (2000). Temporal context and conditioned reinforcement value. *Journal of Experimental Psychology: General*, *129*, 427-443.
- HERRNSTEIN, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, *4*, 267-272.
- HERRNSTEIN, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, *13*, 243-266.
- MAZUR, J. E. (2001). Hyperbolic value addition and general models of animal choice. *Psychological Review*, *108*, 96-112.
- O'DALY, M., ANGULO, S., GIPSON, C., & FANTINO, E. (2006). Influence of temporal context on value in the multiple-chains and successive-encounters procedures. *Journal of the Experimental Analysis of Behavior*, *85*, 309-328.
- O'DALY, M., MEYER, S., & FANTINO, E. (2005). Value of conditioned reinforcers as a function of temporal context. *Learning & Motivation*, *36*, 42-59.
- REYNOLDS, G. S. (1961). Behavioral contrast. *Journal of the Experimental Analysis of Behavior*, *4*, 57-71.

- ROSS, R. T., & HOLLAND, P. C. (1981). Conditioning of simultaneous and serial feature-positive discriminations. *Animal Learning & Behavior*, **9**, 293-303.
- SINGER, R. A., BERRY, L. M., & ZENTALL, T. R. (2007). Preference for a stimulus that follows a relatively aversive event: Contrast or delay reduction? *Journal of the Experimental Analysis of Behavior*, **87**, 275-285.
- VASCONCELOS, M., URCUIOLI, P. J., & LIONELLO-DE NOLF, K. M. (2007a). Failure to replicate the "work ethic" effect in pigeons. *Journal of the Experimental Analysis of Behavior*, **87**, 383-399.
- VASCONCELOS, M., URCUIOLI, P. J., & LIONELLO-DE NOLF, K. M. (2007b). When is a failure to replicate not a Type II error? *Journal of the Experimental Analysis of Behavior*, **87**, 405-407.
- VON FERSEN, L., WYNNE, C. D. L., DELIUS, J. D., & STADDON, J. E. R. (1991). Transitive inference formation in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **17**, 334-341.
- WILLIAMS, B. A. (1983). Another look at contrast in multiple schedules. *Journal of the Experimental Analysis of Behavior*, **39**, 345-384.
- WILLIAMS, B. A. (1991). Behavioral contrast and reinforcement value. *Animal Learning & Behavior*, **19**, 337-344.
- WILLIAMS, B. A. (1992). Inverse relations between preference and contrast. *Journal of the Experimental Analysis of Behavior*, **58**, 303-312.
- WILLIAMS, B. A. (2002). Behavioral contrast redux. *Animal Learning & Behavior*, **30**, 1-20.
- WILLIAMS, B. A., & McDEVITT, M. A. (2001). Competing sources of stimulus value in anticipatory contrast. *Animal Learning & Behavior*, **29**, 302-310.
- ZENTALL, T. R. (2005). A within-trial contrast effect and its implications for several social psychological phenomena. *International Journal of Comparative Psychology*, **18**, 273-297.
- ZENTALL, T. R., & SHERBURNE, L. M. (1994). Transfer of value from S+ to S- in a simultaneous discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, **20**, 176-183.
- ZENTALL, T. R., & SINGER, R. A. (2007). Within-trial contrast: When is a failure to replicate not a Type I error? *Journal of the Experimental Analysis of Behavior*, **87**, 401-404.

NOTE

1. Although we intended the simultaneous condition of Experiment 1 to be a direct replication of Clement et al. (2000), their description of the probe procedure did not specify whether outcomes (presence or absence of reinforcement) followed responses immediately or after a delay.

(Manuscript received July 7, 2007;
revision accepted for publication August 30, 2007.)