Effects of outcome devaluation on instrumental behaviors in a discriminated heterogeneous chain

Eric A. Thrailkill and **Mark E. Bouton**

University of Vermont

Abstract

Operant behavior often takes place in a sequence, or chain, of linked responses that lead to a reinforcer. We have recently studied rats performing a discriminated heterogeneous behavior chain that involves the presentation of a discriminative stimulus (e.g., a panel light) to set the occasion for a procurement behavior (e.g., a lever press) that leads to a second stimulus (e.g., a second panel light) that indicates that a consumption response (e.g., a chain pull) will be reinforced. The present study assessed the role played by a representation of the reinforcer in controlling the performance of the responses in this chain. After acquisition of the chain, rats received a reinforcer devaluation treatment in the form of repeated paired, or unpaired, presentations of the food-pellet outcome and lithium-chloride illness. Once paired rats came to reject the pellets, half the animals in each group were tested on procurement, and the other half were tested on consumption. Neither response was affected by the outcome devaluation treatment, although entries into the food cup were suppressed. Combined with other results, the findings suggest that the "goal" for goal-directed procurement responding in a discriminated heterogeneous chain may be the consumption response rather than the primary reinforcer.

Keywords

Behavior often takes place in a chain of linked responses. Such a chain minimally requires completing two distinct, but serially dependent, behaviors in order to attain a reinforcer. For instance, completing a procurement behavior (e.g., purchasing junk food at a mini-mart) must occur before gaining access to an opportunity to perform a consumption response (e.g., actually opening the package and eating the snack). In many cases, behavior chains are discriminated: Each response takes place in the presence of its own distinct discriminative stimulus. We and others have argued that understanding behavior chains has implications for how to both reduce unwanted behaviors (e.g., smoking, drug taking, and overeating; Ostlund& Balleine, 2009; Thrailkill & Bouton, 2015a, 2016a) and promote wanted behaviors (e.g., searching behavior in police dogs; Thrailkill, Porritt, Kacelnik, & Bouton, 2016) that occur in chains.

Address correspondence to: Eric A. Thrailkill, Department of Psychological Science, University of Vermont, Burlington, VT 05405, eathrail@uvm.edu, mbouton@uvm.edu.

Recently, we used extinction to reveal several features of the associative structure learned in a discriminated heterogeneous chain (Thrailkill & Bouton, 2015a, 2016b). In such a chain, a procurement response (e.g., lever press) in the presence of a procurement discriminative stimulus (SD; e.g., a panel light) leads to a consumption SD (e.g., panel light) that sets the occasion for a consumption response (e.g., chain pull) that earns a reinforcer. The results suggest that the responses in the chain are inter-associated: Extinction of the procurement response weakens consumption responding, and extinction of the consumption response weakens procurement responding (Thrailkill & Bouton, 2015a, 2016b). Each effect depends critically on the animals learning to inhibit the response in extinction; that is, extinction of the SDs without the opportunity to make the extinguished response had no effect on the other response (see also Bouton, Trask, & Carranza-Jasso, 2016). Importantly, the effects were specific to the response association learned in the chain: When two separate chains were trained, extinction of one response (either a procurement or a consumption response) weakened only its associated response, and not the corresponding response in the other chain.

In the present study, we addressed the role of the animal's representation of the outcome in the discriminated chain. We did so by using a reinforcer devaluation technique. Adams (1982) showed that rats could adjust their instrumental behavior in response to a change in the value of its outcome. Briefly, after acquiring a lever-press response for food pellets, rats learned a taste aversion to the food pellets (devaluation) and came to reject the food pellets. Next, rats were allowed to emit the lever press response. Importantly, no outcomes were delivered in this test (extinction); therefore, any effect of devaluation on responding reflects the integration of response-outcome (R-O) learning with the change in the value of the out come across phases. Although instrumental responding was suppressed by the devaluation treatment, in subsequent experiments, Adams (1982) found that responding was not depressed if the rats had more extensive instrumental training. Many writers have subsequently argued that extensive training leads to control by a stimulus-response (S-R) association, under which the presence of the lever elicits the lever press response without regard to the value of its outcome (Dickinson, 1985).

In studies of instrumental chains, extinction of the consumption response produces a result that is analogous to this effect of devaluing the reinforcer. That is, procurement responding is weakened after consumption responding is extinguished (Olmstead, Lafond, Everitt, & Dickinson, 2001; Thrailkill & Bouton, 2016b; Zapata, Minney, & Shippenberg, 2010). However, recent results suggest that this effect depends critically on the opportunity to make and inhibit the consumption response (Thrailkill & Bouton, 2016b). Importantly, extinction exposure to the consumption SD alone (without the opportunity to make the consumption response) had no effect on the procurement response. Thus, devaluation of a possible conditioned reinforcer had no impact on the procurement response. This pattern led us to the hypothesis that the consumption response, rather than a reinforcer per se, may be the valued goal for procurement responding in an instrumental chain.

Other studies have tested the effects of reinforcer devaluation on chained instrumental behaviors. Balleine, Garner, Gonzales, and Dickinson (1995; Balleine, Paredes-Olay, & Dickinson, 2005) assessed the effect of outcome devaluation on rats' responding in a chain

procedure that did not involve control of the chained responses by different SDs. The results showed that consumption responses (which occur temporally proximal to the reinforcer) were immediately suppressed following outcome devaluation (satiation or taste aversion), but procurement responding was not. Animals only reduced their procurement responding in a test if they previously had an opportunity to learn they no longer valued the outcome (incentive learning; Balleine, 1992, see also Corbit & Balleine, 2003). Balleine et al. (1995) noted the parallel between the control of chained instrumental responses and findings in Pavlovian second-order conditioning, where the CR elicited by a proximal CS (CS2) is sensitive to a change in the outcome representation (habituation to the US) but responding to a more distal CS (CS1) is not (e.g., Rescorla, 1977).

To date, the effects of outcome devaluation have not been assessed in a *discriminated* heterogeneous chain. In the present experiment, rats acquired a chain consisting of the presentation of a procurement SD (S1) that signaled that a procurement response (R1; e.g., lever press) could lead to the presentation of a consumption SD (S2) that set the occasion for a consumption response (R2; e.g., a chain pull) to be reinforced. Rats then received taste-aversion conditioning consisting of either paired or unpaired presentations of the food-pellet reinforcer and an injection of lithium chloride (LiCl). Once the paired animals stopped consuming the pellets, half the rats in each group received a test with either S1 or S2 in extinction. In general, if the representation of the primary reinforcer controls performance of the discriminated instrumental chain, then rats should reduce their responding after devaluation. But if the consumption response, rather than the reinforcer, is the goal of the procurement response, there should be no devaluation effect. It was also possible that devaluation could weaken R2, but not R1 (Balleine et al., 1995, 2005; Corbit & Balleine, 2003). Following the test, to further confirm the effectiveness of outcome devaluation, all rats received a second test in which they could freely consume the food pellets. Finally, a reacquisition session compared the ability of the reinforcer to support performance of the chain in Paired and Unpaired groups.

Method

Subjects

Thirty-two naïve female Wistar rats (Charles River, St. Constance, Canada), aged 75–90 days at the start of the experiment, were housed in suspended wire-mesh cages in a room with a 16:8 light-dark cycle. Experimental sessions were conducted during the light portion of the cycle at approximately the same time each day. Rats were food deprived and maintained at 80% of their free-feeding weights for the duration of the experiment. Rats had unlimited access to water in their home cages.

Apparatus

The apparatus was the same as that described in previous studies of instrumental chains (Thrailkill & Bouton, 2015a, Experiment 1; Thrailkill & Bouton, 2016b, Experiment 1). Each of eight operant chambers was housed in its own sound attenuation chamber. Each operant chamber had a recessed floor-level food cup positioned in the center of the front wall; a response lever was mounted to the left of the food cup and a chain suspended from

the ceiling (which activated a microswitch when pulled) was positioned to the right. An infrared beam in the food cup detected food-cup entries. Two 28 V panel lights were positioned on the front wall near the lever and the chain. Reinforcement consisted of the delivery of a 45 mg food pellet into the food cup. The apparatus was controlled by computer equipment in an adjacent room

Procedure

Food restriction began one week prior to the beginning of training. During training, one session was conducted each day, 7 days a week. Animals were handled daily and maintained at their target weight with supplemental feeding at approximately 2 hr postsession when necessary.

Acquisition—On the first two days of training, all rats received sessions of magazine training consisting of 30 free pellets delivered on a Variable Time (VT) 60-s schedule. Training of the consumption response (R2) then began on Day 3. Only the R2 manipulandum (chain or lever, counterbalanced) was present. A response on R2 was reinforced according to fixed ratio (FR) 1 until 20 pellets were earned, then an additional 30 pellets could be earned according to FR 1 during presentations of the consumption SD (S2); responses in the absence of the SD were no longer reinforced. Completing the FR 1 requirement turned S2 off, delivered a pellet, and initiated a variable 45-s intertrial interval (ITI). S2 was always the panel light near the R2 manipulandum. If a response was not made during S2, it terminated after 60s and a new ITI was initiated. On Day 4, there were 30 presentations of S2 with the FR 1 requirement. Beginning on Day 5, the procurement manipulandum (R1) was introduced to the chamber, and rats now received 30 presentations of S1 (the panel light near R1). A response on R1 turned S1 off and turned on S2, which set the occasion for reinforcement of R2 with a food pellet. There was a variable 45-s ITI. On Days 5 and 6, the response requirement for each link was increased to random ratio (RR) 2. The requirement was further increased to RR 4 for Days 7–11. During this period, the maximal duration in each S was gradually reduced from 60s to 20s. At the end of the acquisition phase, the rats had earned a total of 290 reinforcers and had 7 sessions of training with the full chain.

Aversion Conditioning—On Day 12, rats were matched on response rates and then assigned to the paired or unpaired group. Aversion conditioning with the pellet reinforcer proceeded over the next 12 days. Conditioning trials took place in the conditioning chambers with response manipulanda removed. Time in the conditioning chambers in each cycle depended on the mean time required to deliver all the pellets. On the first day of each two-day cycle, half the rats (Paired) received 50 free pellet deliveries in the chamber according to a VT 45-s schedule, and all rats then received an intraperitoneal (ip) injection of 20 ml/kg (0.15 M) LiCl immediately following the session and placed back in their home cages. On the second day, the remaining half of the rats (Unpaired) received 50 free pellets according to VT 45 s, and all rats then received an injection of isotonic saline (20 ml/kg, ip). There were six 2-day conditioning cycles. In order to maintain equivalent pellet exposure during aversion conditioning, the Unpaired group received the mean number of pellets consumed by the Paired group on each trial.

Testing—Testing was then conducted on the next 3 days. On the first and crucial day of the test, all rats received one 16-trial session (approximately 20 min in duration) with both response manipulanda in place. Half the rats in the Paired and Unpaired groups received trials in which S1 was presented alone (Groups S1 Paired and S1 Unpaired), and the remaining half received S2 (Groups S2 Paired and S2 Unpaired). Responses turned off the S according to RR 4, but otherwise had no consequences (Thrailkill & Bouton, 2015a). On the second day, the rats were given a test with the pellets delivered noncontingently to assess (again) the strength of the aversion to the pellets. The rats were placed in the chamber with the manipulanda removed and 10 food pellets were delivered on VT 45-s. Food cup entries and number of pellets consumed were recorded. Finally, on the last day, the rats received a reacquisition test in which they could again perform the usual S1R1–S2R2 chain to earn food pellets. There were 30 trials in the reacquisition session. If the rat did not meet the R1 response requirement (RR 4), S1 went off (and S2 was presented) after 20 s. An S2 ended without a pellet after 20 s and initiated the next ITI if the rat did complete the RR 4.

Data Analysis—Response rates (responses per min) were evaluated with analyses of variance (ANOVAs) using a rejection criterion of $p < .05$. Effect sizes are reported where appropriate. Confidence intervals (CIs) for effect sizes were calculated according to the method suggested by Steiger (2004). When support for the null hypotheses was relevant for interpreting the results, we calculated Bayes factors (BF) using the scaled Jeffrey-Zellner-Siow prior following the method suggested by Rouder, Speckman, Dongchu, and Morey (2009). Due to small sample and effect size in the tests, the scaled-information prior (r) was set to 0.5 in calculating BF.

Results

Acquisition

Acquisition of the chain proceeded without incident and the results were similar to previously published work (Thrailkill & Bouton, 2015a, 2016b). Rats learned to perform the appropriate response in each SD over sessions. In the final session, mean procurement (R1) response rates during the pre-S1, S1, and S2 periods were 8.1, 22.4, and 1.9. Mean consumption (R2) response rates during the same periods were 4.2, 6.8, and 43.2 A Devaluation (Paired, Unpaired) by to-be-tested Stimulus (S1, S2) by Response (R1, R2) ANOVA comparing response rates during the pre-S1 period found greater R1 than R2 responding, $F(1, 28) = 6.80$, $MSE = 36.61$, $p = .01$, $\eta_p = .20$, and no other effects, largest $F = 1.66$. A similar analysis of responding during S1 also found greater R1 than R2 responding, $F(1, 28) = 55.74$, $MSE = 69.66$, $p < .001$, $\eta_p = .67$, and no other effects, largest $F(1, 28) = 1.61$, $MSE = 73.88$. Finally, an analysis of responding during S2 found greater responding on R2 than R1, $F(1, 28) = 378.86$, $MSE = 72.18$, $p < .001$, $\eta_p = .93$, and no other effects, largest $F(1, 28) = 1.39$, $MSE = 81.71$.

Aversion Conditioning

In order to roughly match pellet exposure between Paired and Unpaired groups, only the mean number of pellets consumed by the Paired group on the preceding cycle was presented to each group in the next cycle. Thus, the number of pellets presented decreased across

cycles of aversion conditioning (50, 50, 45, 41, 13, and 3). The Unpaired groups consumed all the pellets offered in each cycle. The Paired groups consumed all the pellets in the first two trials, and then a decreasing proportion of the pellets presented over the four remaining trials. The mean proportions of pellets consumed on those trials were .89, .67, .46, and .00, and .91, .73, .42, and .00, in the groups tested with R1 and R2, respectively. The two Paired groups did not differ in the proportion of pellets consumed. This observation was confirmed by a to-be-tested Stimulus (S1, S2) by Cycle (5) ANOVA, which found only a significant effect of Cycle, $F(4, 56) = 81.86$, $MSE = 0.03$, $p < .001$, $\eta_p = .85$, other $Fs < 1$.

Testing

The results of testing are shown in Figure 1, which separately presents the groups tested with S1 (left column) and S2 (right column). The layout presents responding occasioned by the tested SD in the top row, responding on the manipulandum not occasioned by the SD in the middle row, and food-cup entries for each group in the bottom row.

Response occasioned by the tested stimulus—Figure 1a shows R1 response rates in the S1 test during the 30-s pre S1, S1, and post-S1 periods. Responding on R1 was low during the pre-S1 period, increased during S1, and then decreased during post-S1. There was no effect of the devaluation treatment. A Devaluation (Paired, Unpaired) by Stimulus period (pre, S1, post) by Block (4) ANOVA found effects of Stimulus period, $F(2, 28) = 56.49$, $MSE = 45.11$, $p < .001$, $\eta_p = .80$, Block, $F(3, 42) = 10.41$, $MSE = 41.13$, $p < .001$, $\eta_p = .43$, and a SD period by Block interaction, $F(6, 84) = 5.53$, $MSE = 19.93$, $p < .001$, $\eta_p = .28$. However, no effects involving Devaluation approached significance, largest $F(2, 28) = 1.39$, BF = 1.47.

Figure 1b shows R2 response rates in the S2 test during the pre S2, S2, and post S2 periods. Responding on R2 was low during pre S2, increased during S2, and then decreased during the post S2 period. A Devaluation (Paired, Unpaired) by Stimulus period (pre, S2, post) by Block (4) ANOVA found effects of Stimulus period, $F(2, 28) = 56.49$, $MSE = 78.44$, $p < .001$, $\eta_p = .80$, Block, $F(3, 42) = 39.12$, $MSE = 43.28$, $p < .001$, $\eta_p = .74$, and a Stimulus period by Block interaction, $F(6, 84) = 16.30$, $MSE = 29.65$, $p < .001$, $\eta_p = .54$. There was also a three-way interaction, $F(6, 84) = 3.65$, $p = .003$, $\eta_p = .21$, and a marginal Devaluation by Block interaction, $F(3, 42) = 2.63$, $p = .06$, $\eta_p = .16$. Follow up Devaluation by Block ANOVAs compared responding during each stimulus period. For pre-S2, there was only an effect of Block, $F(3, 42) = 19.06$, $MSE = 2.93$, $p < .001$, $\eta_p = .58$; other $Fs < 1$. For S2, R2 appeared lower in Paired group in the first block, but then crossed over in the second block and remained higher than that in the Unpaired group for the remainder of the test. A Devaluation by Block interaction supported this observation, $F(3, 42)\ 25.28 = 3.50$, $p = .02$, $\eta_p = .20$, along with an effect of Block, $F(3, 42) = 25.28$, $MSE = 92.80$, $p < .001$, $\eta_p = .64$, and no effect of Devaluation, $F < 1$. Follow up comparisons assessed the interaction. R2 responding in the Paired group was significantly *higher* than the Unpaired group in Block 2, $F(1, 14) = 5.92$, $MSE = 100.50$, $p = .03$, $\eta_p = .30$, but there were no group differences in the other blocks, largest $F = 1.01$. The results provide no evidence that the pairing treatment depressed either R1 or R2 responding.

Response not occasioned by the tested stimulus—Test responding on the other, non-occasioned, response is presented in Figures 2c and 2d. Generally, responses on the alternative manipulandum were low during each stimulus period, and were also not affected by devaluation. For the groups tested with S1, R2 responding was assessed during pre-S1, S1, and post S1 periods. A Devaluation by Stimulus period by Block ANOVA found effects of Stimulus period, $F(2, 28) = 12.85$, $MSE = 3.72$, $p < .001$, $\eta_p = .48$, Block, $F(3, 42) = 11.27$, $MSE = 13.27$, $p < .001$, $\eta_p = .45$, and a Stimulus period by Block interaction, $F(6, 84) = 4.35$, $MSE = 3.08$, $p = .001$, $\eta_p = .24$. But there were no effects involving Devaluation, largest $F(2, 28) = 2.21$. For the groups tested with S2, R1 responding assessed during pre-S2, S2, and post-S2 periods. A Devaluation by Stimulus period by Block ANOVA found effects of Stimulus period, $F(2, 28) = 3.39$, $MSE = 12.79$, $p < .05$, $\eta_p = .19$, and Block, $F(3, 42) = 4.44$, $MSE = 36.58$, $p < .01$, $\eta_p = .24$, but no other effects or interactions $Fs < 1$.

Food cup entries during testing—Food cup entries in the test are presented in Figures 1e and 1f. In contrast to lever pressing and chain pulling, food cup entries were strongly suppressed in the Paired groups.

For the groups tested with S1 (Figure 1e), food-cup entry rate was first assessed across each period, and then within pre-S1, S1, and post S1 periods separately. A Devaluation by Stimulus period by Block ANOVA found a reliable Stimulus period effect, $F(2, 28) = 6.12$, $MSE = 20.54$, $p = .006$, $\eta_p = .30$, and no other effects or interactions, largest $F(1, 14) = 2.94$, $MSE = 187.85$. For the pre-S1 period, there was a Devaluation by Block interaction, $F(3, 42) = 3.12$, $MSE = 29.20$, $p = .04$, $\eta_p = .18$, and no other effects, largest $F(1, 14) = 2.81$, $MSE = 111.39$. The Devaluation effect was significant in the first, $F(1, 14) = 11.78$, $MSE = 8.27$, $p = .004$, $\eta_p = .46$, 95% CI [.07, .67], second, $F = 5.03$, $MSE = 22.96$, $p = .04$, $\eta_p = .26$, 95% CI [.00, .54], and fourth blocks, $F = 4.96$, $MSE = 71.90$, $p = .04$, $\eta_p = .26$, 95% CI [.00, .54], but not in the third, $F < 1$. In S1, the Devaluation by Block ANOVA found no effects, largest $F(1, 14) = 1.56$, $MSE = 33.47$, and there was no effect of Devaluation in the individual blocks of S1 presentations, largest $F = 1.71$, $MSE = 20.23$. For the post-S1 period, there were no significant effects or interactions, largest $F(1, 14) = 2.97$, $MSE = 84.08$. The devaluation effect was significant in the first, $F(1, 14) = 6.58$, $MSE = 12.99$, $p = .02$, $\eta_p = .32$, 95% CI [.00, .58], and second, $F = 10.01$, $MSE = 19.51$, $p < .01$, $\eta_p = .42$, 95% CI [.04, .65], but not the third and fourth blocks, $Fs < 1$.

For groups tested with S2, the same analysis found an effect of Devaluation, $F(1, 14) = 6.23$, $MSE = 211.80$, $p = .03$, $\eta_p = .31$, 95% CI [.00, .57], a SD by Block interaction, $F(6, 84) = 2.99$, $MSE = 8.09$, $p = .01$, $\eta_p = .18$, and no other effects, largest $F(3, 42) = 1.86$, $MSE = 40.98$. For the pre-S2 period, there were effects of Devaluation, $F(1, 14) = 8.48$, $MSE = 40.11$, $p = .01$, $\eta_p = .38$, 95% CI [.02, .62], and Block, $F(3, 42) = 6.15$, $MSE = 14.18$, $p = .001$, $\eta_p = .31$, that did not interact, $F = 2.14$. For the S2 period, there was only a Devaluation effect, $F(1, 14) = 4.48$, $MSE = 118.09$, $p = .05$, $\eta_p = .24$, 95% CI [.00, .52], other $Fs < 1$. For the post-S2 period, there was an effect of Devaluation, $F(1, 14) = 5.27$, $MSE = 87.74$, $p = .04$, $\eta_p = .27$, 95% CI [.00, .55], that did not involve Block, largest $F(3, 42) = 1.30$, $MSE = 21.05$. Overall, devaluation had its strongest effect on food cup entries during periods that had predicted food during acquisition; that is, during and after S2 for the S2 groups, and after S1 for the S1 groups.

Food pellet test—In the subsequent food pellet test, the Unpaired groups ate all of the 10 pellets presented, and the S1-Paired and S2-Paired groups respectively ate means of 0.4 and 0.6 pellets, which did not differ, $F < 1$.

Reacquisition test—The results of the reacquisition test, when the groups could again make R1 and R2 to receive the pellet reinforcer, are shown in Figure 2. Both the paired and unpaired groups completed trials at the start of the test. This meant that both groups earned some food pellets, but not that they necessarily ate them. The number of pellets earned and actually eaten are presented in Figure 2a. A Devaluation by Test stimulus by Pellet status (Earned, Eaten) ANOVA revealed effects of Devaluation, $F(1, 28) = 113.88$, $MSE = 62.52$, $p < .001$, $\eta_p = .80$, Pellet status, $F(1, 28) = 54.59$, $MSE = 7.61$, $p < .001$, $\eta_p = .66$, and a Devaluation by Pellet status interaction, $F = 54.59$, $p < .001$, $\eta_p = .66$. A follow-up ANOVA compared the number of earned and eaten pellets in the Devalued groups and found a significant difference between the number of pellets earned and eaten, $F(1, 14) = 54.56$, $MSE = 15.21$, $p < .001$, $\eta_p = .80$. No other effects reached significance, largest $F(1, 14) = 1.22$, $MSE = 24.55$.

R2 and R1 responding are shown in Figures 2b and 2c. Interestingly, both responses occurred at a high and unsuppressed rate in the devalued group at the start of testing. We calculated elevation scores to describe R1 and R2 responding occasioned by the corresponding SD by subtracting response rate during the 30 s immediately before S1 from the response rate during each SD (cf. Thrailkill & Bouton, 2015a, 2016b). There were no differences in responding during the pre-S1 to period complicate our analysis of elevation scores, and we omit the analysis for brevity. Concerning R2 responding (Figure 2b), a Devaluation by Test stimulus by 5-trial Block ANOVA found an effect of Devaluation, $F(1, 28) = 39.30$, $MSE = 1240.86$, $p < .001$, $\eta_p = .58$, and a Devaluation by Block interaction, $F(5, 140) = 13.18$, $MSE = 130.66$, $p < .001$, $\eta_p = .32$. There were no other effects, largest $F(5, 140) = 1.80$. Planned ANOVAs compared responding in the first and second 5-trial blocks of the test and found no differences in the first block, $F < 1$, smallest BF = 1.77; the devaluation effect became significant in the second block, $F(1, 28) = 22.46$, $MSE = 288.97$, $p < .001$, $\eta_p = .45$. The fact that devaluation had no effect during the early blocks supports the other test data in suggesting that the devaluation treatment initially had no effect on R2 responding.

Concerning R1 responding, a Devaluation by Test stimulus by 5-trial Block ANOVA found an effect of Devaluation, $F(1, 28) = 22.62$, $MSE = 481.24$, $p < .001$, $\eta_p = .45$, and a Devaluation by Block interaction, $F(5, 140) = 3.76$, $MSE = 111.86$, $p = .003$, $\eta_p = .12$. There were no other effects, largest $F(5, 140) = 1.66$. Again, planned ANOVAs compared responding in the first and second blocks of reacquisition and found no differences in the first 5-trial block, $Fs < 1$, and the devaluation effect became significant in the second block, $F(1, 28) = 7.30$, $MSE = 210.48$, $p = .01$, $\eta_p = .12$. This was also true when we compared the number of completed chains (i.e., trials in which the RR contingencies were met in S1 and S2) in the first and second 5-trial blocks. Devaluation had no effect in the first block, $Fs < 1$, smallest BF = 1.82, but did in the second block, $F(1, 28) = 10.56$, $MSE = 2.84$, $p = .003$, $\eta_p = .27$. Once again, the absence of a devaluation effect early in re-training supports the other

results in suggesting little effect of the devaluation treatment on R1—until, perhaps, the animals began tasting the food pellets.

Discussion

Rats acquired the discriminated heterogeneous chain and learned to make the appropriate response in S1 and S2 over a relatively brief period of training. Outcome devaluation then led to complete rejection of the food pellets in the Paired groups, an effect that was still strong in the food pellet test that followed the extinction tests of operant responding. However, when tested for procurement (R1) or consumption (R2) responding, there was no evidence that outcome devaluation depressed either behavior. In contrast, food cup entries were suppressed. Strikingly, a reacquisition test also revealed that animals initially performed the entire instrumental chain even when it produced the devalued reinforcer, which was itself rejected during the reacquisition test. Thus, devaluing the outcome representation had no effect on the responses tested (1.) individually or (2.) within the chain, i.e., under the conditions of acquisition. The overall pattern of results suggests that a representation of the reinforcing outcome plays a very weak role, if any, in motivating R1 and R2 responding in the present heterogeneous instrumental chain.

One possible explanation is that the instrumental training might have been extensive enough to convert the procurement and consumption responses into habits (Adams, 1982; Dickinson, Balleine, Watt, Gonzales, & Boakes, 1995; Thrailkill & Bouton, 2015b). We know that extended training does diminish the effect of extinguishing consumption on procurement responding in a related procedure (Zapata et al., 2010). However, the present experiment involved considerably less training than that study and found no evidence of a reinforcer devaluation effect. More important, the present experiment also involved less training than our own study with the same training methods in which procurement responding was weakened by extinction of consumption responding (Thrailkill & Bouton, 2016b, Experiment 1). In the earlier experiment, rats received a total of 420 reinforced consumption responses (360 trials with the full chain) during chain training. In the present study, rats received only 290 reinforced consumption responses (and 210 trials with the chain). The difference suggests that the present insensitivity of procurement to reinforcer devaluation was not merely due to the creation of a habit. Instead, the "goal" representation for procurement may be the consumption response (Thrailkill & Bouton, 2016b) rather than the reinforcer representation tested here.

The idea that R2 is the goal (and thus perhaps the reinforcer) for R1 is broadly consistent with reinforcement theories that have emphasized the relationship between operant and contingent responses (e.g., Premack, 1965; Timberlake & Allison, 1974). However, to our knowledge those theories have not directly addressed behavior chains, and more important, they do not suppose that the animal learns a *representation* of the contingent behavior in a way that would allow a devaluation effect of the type we have observed (Thrailkill & Bouton, 2016b). That is, they would not predict the immediate reduction in R1 when it is tested after separate extinction of R2. Nonetheless, the importance of the relationship between the two responses is perhaps reminiscent of those theories.

The results may appear to contrast with those reported by Balleine et al. (1995). In their experiment, which used a non-discriminated procedure that lacked the present SDs, rats decreased their R2 responding (but not R1 responding) immediately after being sated on the food reinforcer. Further testing revealed that rats could reduce R1 responding (but not R2 responding), only after posttraining re-exposure to the food reinforcer in a sated state (i.e., incentive learning; Balleine, 1992). It is notable that evidence of a devaluation effect in the present experiment was found in food cup entries, which were clearly reduced in the Paired groups and especially apparent following the offset of a consumption SD. The food cup data are consistent with Balleine et al.'s report of devaluation effects on later responses in the chain (Balleine, 1992; Balleine et al., 1995, 2005; Corbit & Balleine, 2003). This aspect of the present results is consistent with Balleine et al.'s (1995) idea that the closer the behavior is to the end of a chain the more likely it is to be sensitive to outcome devaluation.

In the present experiment, rats were required to make each response in its own SD, and reinforcement depended on discriminative control of each response; the rats made few responses in the absence of its stimulus (see also Thrailkill & Bouton, 2015a, 2016b). Acquisition of stimulus control in a discriminated chain involves learning to use the SDs to partition R1, R2, and O. Such partitioning may explain the difference in control by the outcome value in the present procedure and that observed in chains that do not involve separate SDs for R1 and R2 (e.g., Balleine et al., 1995; Ostlund, Winterbauer, & Balleine, 2009).

In summary, the results suggest that rats are not motivated by a representation of the reinforcer when they perform R1 and R2 in a discriminated heterogeneous chain. The fact that procurement may instead depend on a representation of the next *response* (Thrailkill & Bouton, 2016b) may be analogous to the cocaine user who is motivated to engage in procurement behaviors (e.g., finding and buying cocaine) while anticipating consumption activity (e.g., doling out the cocaine and actually using it). Consumption behavior may function as the "goal" for procurement. The present work may further support an important role for the association between responses in controlling performance of a discriminated heterogeneous chain (Thrailkill & Bouton, 2015a, 2016b, Thrailkill, Trott, Zerr, & Bouton, 2016).

Acknowledgments

References

Adams CD. Variations in the sensitivity of instrumental responding to reinforcer devaluation. Quarterly Journal of Experimental Psychology. 1982; 34B:77–98.

Balleine BW. The role of incentive learning in instrumental performance following shifts in primary motivation. Journal of Experimental Psychology: Animal Behavior Processes. 1992; 18:236–250. [PubMed: 1619392]

Balleine BW, Garner C, Gonzalez F, Dickinson A. Motivational control of heterogeneous instrumental chains. Journal of Experimental Psychology: Animal Behavior Processes. 1995; 21:203–217.

Balleine BW, Paredes-Olay C, Dickinson A. Effects of Outcome Devaluation on the Performance of a Heterogeneous Instrumental Chain. International Journal of Comparative Psychology. 2005; 18:257–272.

Bouton ME, Trask S, Carranza-Jasso R. Learning to inhibit the response during instrumental (operant) extinction. Journal of Experimental Psychology: Animal Learning and Cognition. 2016; 42:246–258. [PubMed: 27379715]

Corbit LH, Balleine BW. Instrumental and Pavlovian incentive processes have dissociable effects on components of a heterogeneous instrumental chain. Journal of Experimental Psychology: Animal Behavior Processes. 2003; 29:99–106. [PubMed: 12735274]

Dickinson A. Actions and habits: the development of behavioural autonomy. Philosophical Transactions of the Royal Society of London. 1985; 308B:67–78.

Dickinson A, Balleine B, Watt A, Gonzalez F, Boakes RA. Motivational control after extended instrumental training. Animal Learning & Behavior. 1995; 23:197–206.

Olmstead MC, Lafond MV, Everitt BJ, Dickinson A. Cocaine seeking by rats is a goal-directed action. Behavioral Neuroscience. 2001; 115:394–402. [PubMed: 11345964]

Ostlund SB, Balleine BW. On habits and addiction: An associative analysis of compulsive drug seeking. Drug Discovery Today: Disease Models. 2009; 5:235–245.

Ostlund SB, Winterbauer NE, Balleine BW. Evidence of action sequence chunking in goal-directed instrumental conditioning and its dependence on the dorsomedial prefrontal cortex. Journal of Neuroscience. 2009; 29:8280–8287. [PubMed: 19553467]

Premack, D. Reinforcement theory. In: Levine, D., editor. Nebraska Symposium on Motivation. Lincoln, NE: University of Nebraska Press; 1965. p. 123-180.

Rescorla, RA. Pavlovian second-order conditioning: Some implications for instrumental behaviour. In: Davis, H.; Hurwitz, HMB., editors. Operant-Pavlovian Interactions. Hillsdale, NJ: Erlbaum; 1977. p. 133-164.

Rouder JN, Speckman PL, Dongchu S, Morey RD. Bayesian *t* tests for accepting and rejecting the null hypothesis. Psychonomic Bulletin & Review. 2009; 16:225–237. [PubMed: 19293088]

Steiger JH. Beyond the *F* test: Effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis. Psychological Methods. 2004; 9:164–182. [PubMed: 15137887]

Thrailkill EA, Bouton ME. Extinction of chained instrumental behaviors: Effects of procurement extinction on consumption responding. Journal of Experimental Psychology: Animal Learning and Cognition. 2015a; 41:232–246. [PubMed: 25915751]

Thrailkill EA, Bouton ME. Contextual control of instrumental actions and habits. Journal of Experimental Psychology: Animal Learning and Cognition. 2015b; 41:69–80. [PubMed: 25706547]

Thrailkill EA, Bouton ME. Extinction and the associative structure of heterogeneous instrumental chains. Neurobiology of Learning and Memory. 2016a; 133:61–68. [PubMed: 27296700]

Thrailkill EA, Bouton ME. Extinction of chained instrumental behaviors: Effects of consumption extinction on procurement responding. Learning & Behavior. 2016b; 44:85–96. [PubMed: 26276367]

Thrailkill EA, Porritt F, Kacelnik A, Bouton ME. Increaseing the persistence of a heterogeneous behavior chain: Studies of extinction in a rat model of search behavior in working dogs. Behavioural Processes. 2016; 129:44–53. [PubMed: 27306694]

Thrailkill EA, Trott JM, Zerr CL, Bouton ME. Contextual control of chained instrumental behaviors. Journal of Experimental Psychology: Animal Learning and Cognition. 2016 in press.

Timberlake W, Allison J. Response deprivation: An empirical approach to instrumental performance. Psychological Review. 1974; 81:146–164.

Zapata A, Minney VL, Shippenberg TS. Shift from goal-directed to habitual cocaine seeking after prolonged experience in rats. Journal of Neuroscience. 2010; 30:15457–15463. [PubMed: 21084602]
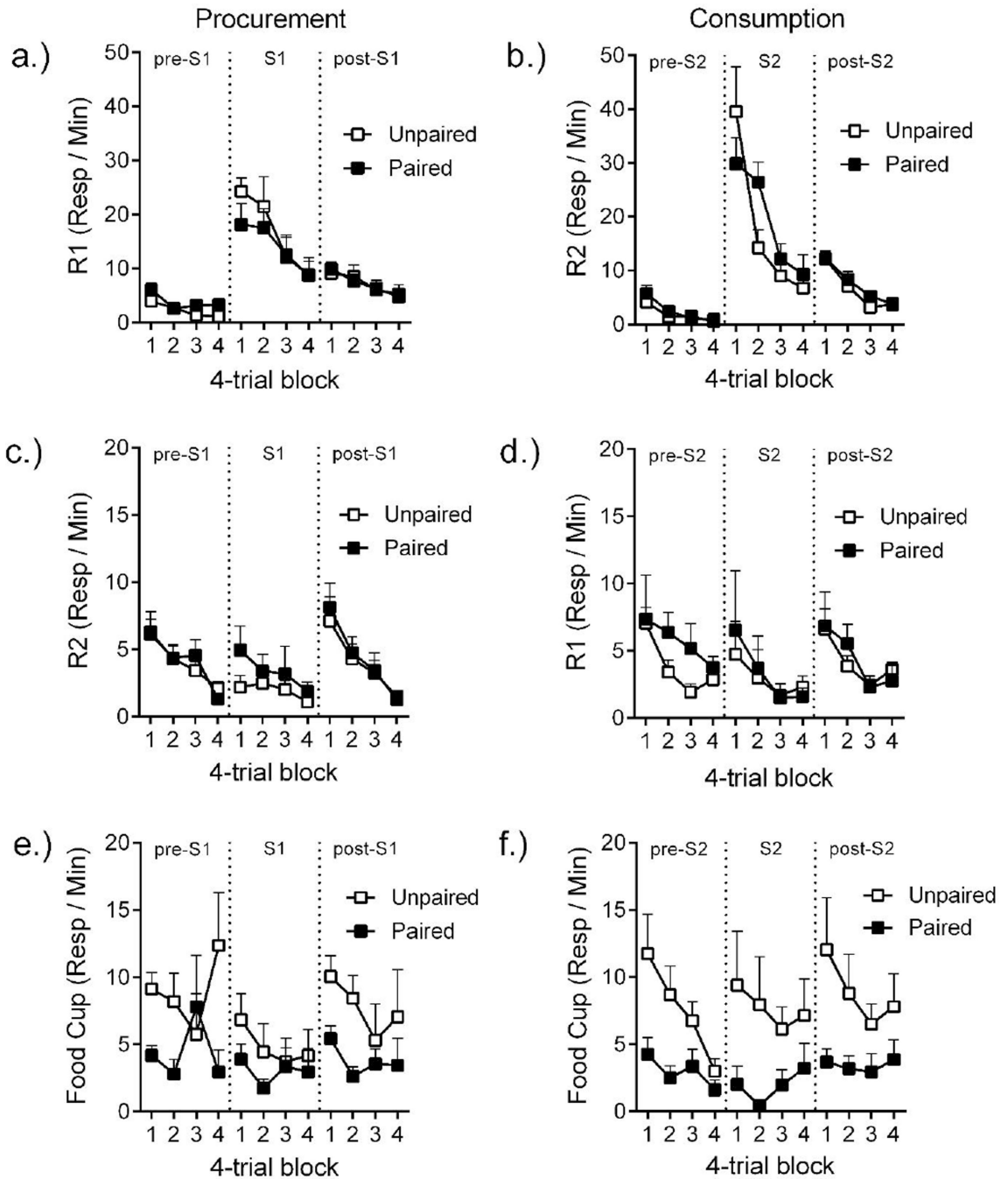
Figure 1.

Test responding. a.) Procurement (R1) and b.) consumption (R2) response rates in each 4-trial block for groups tested with procurement (S1) and consumption (S2) stimuli, respectively. c.) Consumption (R2) response rates in groups tested with S1 and d.) procurement (R1) response rates in groups tested with S2 in each 4-trial block. e.) and f.) Food cup entry rates in each 4-trial block for groups tested with S1 and S2, respectively. Error bars are the standard error of the mean and only appropriate for between group comparisons.
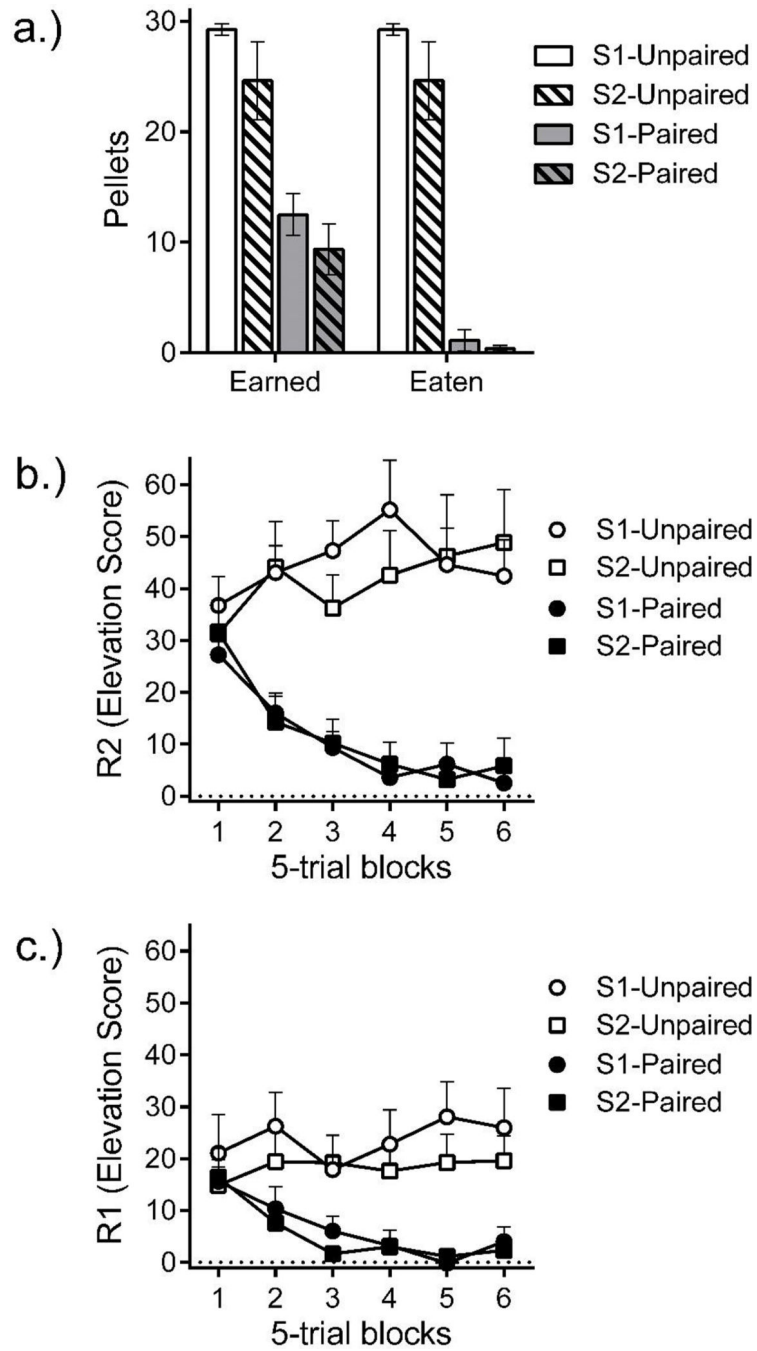
Figure 2.
Reacquisition test. a.) Number of pellets earned and number eaten in the chain reacquisition session. R2 (b.) and R1 (c.) response rate elevation scores across 5-trial blocks of the reacquisition session. Error bars are the standard error of the mean and only appropriate for between group comparisons.