



Published in final edited form as:

J Exp Psychol Anim Learn Cogn. 2015 January ; 41(1): 69–80. doi:10.1037/xan0000045.

Contextual control of instrumental actions and habits

Eric A. Thrailkill and Mark E. Bouton

University of Vermont

Abstract

After a relatively small amount of training, instrumental behavior is thought to be an action under the control of the motivational status of its goal or reinforcer. After more extended training, behavior can become habitual and insensitive to changes in reinforcer value. Recently, instrumental responding has been shown to weaken when tested outside of the training context. The present experiments compared the sensitivity of instrumental responding in rats to a context switch after training procedures that might differentially generate actions or habits. In Experiment 1, lever pressing was decremented in a new context after either short, medium, or long periods of training on either random-ratio or yoked random-interval reinforcement schedules. Experiment 2 found that more minimally-trained responding was also sensitive to a context switch. Moreover, Experiment 3 showed that when the goal-directed component of responding was removed by devaluing the reinforcer, the residual responding that remained was still sensitive to the change of context. Goal-directed responding, in contrast, transferred across contexts. Experiment 4 then found that after extensive training, a habit that was insensitive to reinforcer devaluation was still decremented by a context switch. Overall, the results suggest that a context switch primarily influences instrumental habit rather than action. In addition, even a response that has received relatively minimal training may have a habit component that is insensitive to reinforcer devaluation but sensitive to the effects of a context switch.

Keywords

action; habit; context; instrumental learning

Contemporary analyses of associative learning suggest that instrumental behavior can be controlled by separable action and habit processes (Dickinson, 1994; 2011; Dickinson & Balleine, 2002; Daw, Niv, & Dayan, 2005). Action is goal-directed in the sense that it is sensitive to the current motivational status of the reinforcer. In contrast, habit is not sensitive to the motivational status of the reinforcer and is thus performed mechanically and automatically, without a goal in mind. The distinction between goal-directed actions and habits has been found to have considerable generality across human and animal subjects (Balleine & O'Doherty, 2010). Investigators have made progress in understanding their neural underpinnings (Balleine, 2005; Balleine & Dickinson, 1998a; Balleine & O'Doherty, 2010; Coutureau & Killcross, 2003; Graybiel, 2008; Tanaka, Balleine, & O' Doherty, 2008; Yin & Knowlton, 2006), and the distinction has proven useful when applied to a range of

psychological problems including drug addiction (Everitt & Robbins, 2005; Hogarth, Balleine, Corbit, & Killcross, 2012), schizophrenia (Graybiel, 2008), and obsessive-compulsive disorder (Gillan et al., 2011; Robbins, Gillan, Smith, de Wit, & Ersche, 2012).

Action and habit processes can be separated empirically by changing the value of the reinforcer after the response has been learned. Such reinforcer revaluation is typically accomplished by conditioning a taste aversion to the reinforcer (Adams, 1982; Adams & Dickinson, 1981a) or through sensory-specific satiety (Balleine & Dickinson, 1998b). For example, an instrumental response might first be associated with a sucrose reinforcer. In a separate phase, consumption of sucrose might be paired with illness from a lithium chloride (LiCl) injection over several trials until the animal completely rejects the sucrose. In a final test, the response that previously produced sucrose is tested in extinction. Evidence for goal-directed action takes the form of a reduction in the response after reinforcer devaluation; response rate reflects the animal's knowledge (and memory) of the change in outcome value. In contrast, if the response were a habit, it would be unaffected by reinforcer devaluation, thus demonstrating its independence from the current value of the reinforcer. A suppression of responding, which suggests the role of action, has been demonstrated in many experiments (Adams, 1982; Adams & Dickinson, 1981a; Colwill & Rescorla, 1985a, b; Dickinson, Nicholas, & Adams, 1983). It is worth noting, however, that responses are seldom completely abolished after reinforcer devaluation, suggesting, perhaps, that some habit might be present along with the goal-directed action (see Colwill & Rescorla, 1990 for a critical discussion).

Recent studies from our laboratory have examined the role of the context in controlling instrumental behavior. In these studies, contexts have consisted of different training chambers located in separate rooms with distinctive visual, tactile, and olfactory cues. One consistent finding is that, in contrast to what is often observed with a Pavlovian conditioned response (e.g., Bouton & King, 1983; Bouton & Peck, 1989; Harris, Jones, Bailey, & Westbrook, 2000; see Rosas, Todd, & Bouton, 2013, for a recent review), an instrumental response trained in one context is weaker when it tested in a different context. For example, in studies of the renewal of instrumental responding after extinction, Bouton, Todd, Vurbic, and Winterbauer (2011) found that responding was weaker at the beginning of extinction training when extinction was conducted in a context different from the acquisition context. Todd (2013) reported similar findings when one response (e.g., lever pressing) was trained in Context A and a different response (e.g., chain pulling) was trained in Context B. The response trained in Context A was immediately weakened when extinction was conducted in Context B; the same was true for the response trained in Context B prior to extinction in Context A. This result suggests that the decrement in the response does not depend on different histories of reinforcement in the two contexts. In a related series of experiments, Bouton, Todd, and León (2014) found similar effects of a context switch on discriminated operant responses and on unsignaled responses that were trained with continuous reinforcement. They also found that lever pressing and chain pulling were weaker in a new context regardless of whether the rat was tested on the same lever or chain (moved between contexts), or copies of the lever or chain that were permanently affixed to each context. In sum, over a range of conditions, instrumental responding is weaker in a new environment (Context B) after being trained in another environment.

One interpretation of the results of Bouton et al. (2011, 2014) and Todd (2013) is that the context might enter into a direct association with the response. Such an association might be related to the stimulus-response (S-R) association that is commonly thought to underlie habit learning. This idea suggests the possibility that the context might be part of the stimulus that supports instrumental S-R habits.

Dual-process theories of instrumental behavior that distinguish between action and habit posit that a transition from action to habit control occurs over the course of training (Dickinson & Balleine, 2002; Dickinson, Balleine, Watt, Gonzales, & Boakes, 1995). Early in training, the response is thought to be goal-directed action, and reflects knowledge of a response-outcome association (R-O). Later in training, the response is thought to transition to habit and come under the control of a stimulus-response association (S-R; de Wit & Dickinson, 2009; Dickinson, 1994). The specific cause of the transition from R-O to S-R control is not entirely known. One idea is that the local correlation of responses and outcomes influences the speed of the transition from action to habit. This view is consistent with evidence that interval schedules of reinforcement generate more habit than do ratio schedules of reinforcement (Dickinson, Nicholas, & Adams, 1983).

The present experiments were designed to assess whether the context-switch effect on instrumental responding depends on the status of the response as an action or a habit. They did so in part by manipulating training procedures that were expected to differentially support action and habit learning. Experiment 1 compared the effects of a context switch on instrumental responding after either extended or less-extended training with either ratio or interval schedules of reinforcement. Experiment 2 then confirmed the effect of a context switch after even more limited training. Then, in order to identify the presence of actions and habits more precisely, Experiments 3 and 4 combined the context switch manipulation with a reinforcer devaluation treatment. In Experiment 3, responding after relatively minimal training on a ratio schedule of reinforcement was shown to be sensitive to reinforcer devaluation. However, the responding that still remained after reinforcer devaluation was found to be especially sensitive to the context switch. Experiment 4 then produced an extensively-trained response that was insensitive to reinforcer devaluation. Such a response was nonetheless sensitive to the effects of a context switch. The findings are thus consistent with the view that the context primarily controls the habit component, rather than the action component, of instrumental responding.

Experiment 1

Our previous studies demonstrating the effects of context change on instrumental responding have mainly studied responses trained with variable-interval (VI) schedules of reinforcement (Bouton et al., 2011, 2014; Todd, 2013). One purpose of Experiment 1 was to extend the analysis to the effects of a context switch on responding reinforced on a ratio schedule. Training with ratio or interval schedules is thought to differentially influence the development of actions and habits. Dickinson (e.g., 1985, 1994) has suggested that the differential strengths of response-outcome correlations arranged by these schedules contribute to the development of R-O and S-R control, respectively. A second variable studied was the amount of training. The amount or extent of training is widely thought to

influence the transition from R-O to S-R control (e.g., Adams, 1982; Adams & Dickinson, 1981b; Corbit, Chieng, & Balleine, 2014; Corbit, Nie, & Janak, 2013; Dickinson et al., 1983, 1995). In the present experiment, different groups were reinforced for responding on either a random-ratio (RR) or a yoked random-interval (yoked-RI) schedule in which pellet delivery was available at times in which reinforcers were earned in the RR group. (The advantage of the yoking procedure is that it produced an interval schedule that was approximately matched to the ratio schedule on reinforcement rate.) Three groups in each condition were further given different amounts of training. All training occurred in one context, Context A. The groups also received equal exposure to Context B (without the response manipulandum being present), where pellets were presented noncontingently at a comparable rate. Following training, each group was tested for lever pressing in Contexts A and B (order counterbalanced) under extinction conditions. The question was whether the schedule and/or the amount of training would yield instrumental responses that differed in their sensitivity to the effects of changing the context.

Method

Subjects—Thirty-six female rats (Charles River, St. Constance, Quebec, Canada), aged 75-90 days at the start of the experiment, were individually housed in suspended wire-mesh cages in a room maintained on a 16-8 light-dark cycle. Experimental sessions were conducted during the light portion of the cycle at 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 were allowed ad libitum access to water in their home cages. Weights were maintained by supplemental feeding when necessary at approximately 2 hrs post-session.

Apparatus—The apparatus consisted of two unique sets of four conditioning chambers (model ENV-088-VP; Med Associates, St. Albans, VT) located in separate rooms of the laboratory. Each chamber was housed in its own sound-attenuating chamber. All boxes measured 31.75 × 24.13 × 29.21 cm (length × width × height). The side walls and ceiling consisted of clear acrylic panels, and the front and rear walls were made of brushed aluminum. A recessed food cup was centered on the front wall approximately 2.5 cm above the level of the floor. In both sets of boxes, a retractable lever (model ENV-112CM, Med Associates) was positioned to the left of the food cup. The lever was 4.8 cm wide and was 6.3 cm above the grid floor. It protruded 2.0 cm from the front wall when extended. The chambers were illuminated by 7.5 W incandescent bulbs mounted to the ceiling of the sound-attenuation chamber. Ventilation fans provided background noise of 65 dBA.

The two sets of chambers were fully counterbalanced and had unique features that allowed them to serve as different contexts. In one set of boxes, the floor consisted of 0.5 cm diameter stainless steel floor grids spaced 1.6 cm apart (center-to-center) and mounted parallel to the front wall. The ceiling and a side wall had black horizontal stripes, 3.8 cm wide and 3.8 cm apart. A distinct odor was continuously presented by placing 1.5 ml of Anise Extract (McCormack, Hunt Valley, MD) in a dish immediately outside the chamber. In the other set of chambers, the floor consisted of alternating stainless steel grids with different diameters (0.5 and 1.3 cm), spaced 1.6 cm apart. The ceiling and the left side wall were covered with dark dots (2 cm in diameter). An odor was provided by 1.5 ml of Coconut

Extract (McCormack, Hunt Valley, MD) placed in a dish immediately outside the chamber. Reinforcement consisted of the delivery of a 45-mg food pellet into the food cup (MLab Rodent Tablets; TestDiet, Richmond, IN). The apparatus was controlled by computer equipment located in an adjacent room.

Procedure—Food restriction began one week prior to the beginning of training. During training, two sessions were conducted each day. Animals were handled each day and maintained at their target body weight with supplemental feeding when necessary.

Instrumental training: Rats were divided into three groups ($n = 12$). On Day 1, the group that received the most extensive training began its daily training sessions. The remaining groups experienced equivalent handling, but did not begin training until Days 9 and 13. The staggered starts allowed all groups to be tested on the same day, i.e., after equivalent amounts of handling and days on the food deprivation schedule.

Training always consisted of the following. On the first day, the rats were trained to eat from the food magazine in a session in which 60 pellets were delivered according to a random time (RT) 30 s schedule. The first session occurred in Context A; a similar session was then conducted in a box from the other set of chambers (Context B). Levers were retracted during these sessions. Lever-press training then began the next day. In Context A, pellets were delivered contingent on lever-pressing, whereas in Context B free pellets were delivered while the lever was absent according to the average rate earned in the preceding Context A session. The two types of sessions were double-alternated to control for time of day. The sessions were conducted at least 1.5 hrs apart.

Beginning the day following magazine training, there were two days of lever press training on a continuous reinforcement schedule (CRF). The sessions consisted of the insertion of the left lever and then its retraction after 60 reinforced responses. On the next day, a RR 5 was introduced for the RR group, and yoked-RI rats received pellets for the first response emitted following a reinforcer earned by the RR rat in another chamber. Following one RR 5 session, the RR schedule was increased to RR 15 for the remaining sessions. Groups received a total of 18, 6, and 2 training sessions on the terminal reinforcement schedules, or 1260, 540, and 300 total reinforced responses (including CRF and RR 5 sessions), respectively.

Testing: All rats were tested for lever pressing in both Contexts A and B on the day following the final training session. Test sessions were 10 min in duration; the lever was present, but presses had no scheduled consequences. Order of testing (Context A first, Context B first) was counterbalanced in each group. As usual, the sessions occurred at least 1.5 hours apart.

Data analysis: Response rates (responses per minute) during tests were analyzed in repeated measures analysis of variance. The rejection criterion was set at $p < .05$ for all statistical comparisons. Effect sizes are reported where appropriate. Confidence intervals for effect sizes were calculated according to methods suggested by Steiger (2004). We also analyzed

test responding in terms of its percentage of baseline responding. However, all critical differences remained unchanged, and this analysis is not reported further.

Results

Response rates for each group during training are presented in Figure 1. On the final session, the mean interreinforcer intervals for RR and yoked-RI groups were 8.89 s and 9.10, 10.49 and 10.67, and 19.09 and 19.96 for the 1260, 540, and 300 groups, respectively. Across groups, response rates were higher in the RR groups than yoked-RI groups. ANOVAs isolated the RR and RI groups given the same amounts of training. In the 1260 group, a Schedule (2) by Session (18) repeated-measures ANOVA found significant effects of Schedule, $F(1, 10) = 11.54$, $MSE = 2310.85$, $p = .007$, $\eta_p^2 = .34$, 95% CI [.06, .74], Session, $F(17, 170) = 19.82$, $MSE = 97.96$, $p < .001$, $\eta_p^2 = .66$, 95% CI [.55, .70], and a Schedule \times Session interaction, $F(17, 170) = 3.28$, $p < .001$, $\eta_p^2 = .25$, 95% CI [.07, .28]. In the 540 group, there were similar effects of Schedule, $F(1, 10) = 19.79$, $MSE = 467.48$, $p = .001$, $\eta_p^2 = .66$, 95% CI [.18, .81], Session, $F(5, 50) = 69.27$, $MSE = 31.96$, $p < .001$, $\eta_p^2 = .87$, 95% CI [.79, .90], and a Schedule \times Session interaction, $F(5, 50) = 11.83$, $p < .001$, $\eta_p^2 = .54$, 95% CI [.30, .64]. Finally, in the 300 group, there were significant effects of Schedule, $F(1, 10) = 14.90$, $MSE = 78.09$, $p = .003$, $\eta_p^2 = .59$, 95% CI [.11, .77], and Session, $F(1, 10) = 22.20$, $MSE = 47.17$, $p = .001$, $\eta_p^2 = .69$, 95% CI [.22, .82], but the interaction did not approach significance, $F(1, 10) < 1$.

We also compared response rates on the final session of training. There a Training Amount (3) by Schedule (2) ANOVA found significant effects of Training Amount, $F(2, 30) = 29.72$, $MSE = 111.46$, $p < .001$, $\eta_p^2 = .67$, 95% CI [.41, .77], Schedule, $F(1, 30) = 76.54$, $p < .001$, $\eta_p^2 = .72$, 95% CI [.51, .81], and a Training \times Schedule interaction $F(2, 30) = 5.87$, $p < .01$, $\eta_p^2 = .28$, 95% CI [.03, .47]. The ANOVA thus confirmed that responding increased with amount of training, and that RR maintained higher response rates.

Figure 2 presents the main results of the experiment, response rates during the test sessions in Contexts A and B for all the groups. As the figure suggests, there was a context switch effect evident in all groups. A Context (2) \times Schedule (2) \times Training Amount (3) repeated-measures ANOVA found significant main effects of Context, $F(1, 30) = 22.09$, $MSE = 124.92$, $p < .001$, $\eta_p^2 = .42$, 95% CI [.15, .60], Training amount, $F(2, 30) = 10.68$, $MSE = 109.74$, $p < .001$, $\eta_p^2 = .42$, 95% CI [.12, .58], and Schedule, $F(1, 30) = 14.14$, $p = .001$, $\eta_p^2 = .32$, 95% CI [.07, .52]. However, no interaction approached significance (largest $F = 1.39$).

Because actions as opposed to habits are especially likely after less extensive training, we also isolated and analyzed response rates in the 300 groups in a separate Context \times Schedule ANOVA. Once again, there was a significant effect of Context, $F(1, 10) = 9.89$, $MSE = 28.85$, $p = .01$, $\eta_p^2 = .49$, 95% CI [.04, .71], a significant effect of Schedule, $F(1, 10) = 7.30$,

$MSE = 51.29, p = .02, \eta_p^2 = .42, 95\% \text{ CI } [.00, .67]$, and no Context \times Schedule interaction ($F < 1$).

Discussion

Experiment 1 found a consistent decrement in instrumental responding when responding was tested in Context B. There was no evidence that the ratio or interval schedules or the different amounts of training had any impact on the clear effect of context change. It is worth noting that both the schedule and the amount of training variables had an impact on responding during the training phase. Despite this, there was no evidence that these factors interacted with the strength of the context switch effect. The results thus further extend the generality of our earlier context-switch findings with free-operant (Bouton et al., 2011) and discriminated-operant responding (Bouton et al., 2014).

If it is assumed that ratio vs. interval schedules and under- vs. over-training generate actions and habits, respectively, then the results of Experiment 1 could suggest that a context switch can weaken actions and habits equivalently. However, it is important to note that Experiment 1 provided no direct evidence that responding was either action or habit at the time of testing. Moreover, Dickinson et al. (1995) have shown that instrumental responding can become insensitive to reinforcer devaluation after fewer reinforced responses than were used here, suggesting that the least-trained rats in the present experiment could have received enough training for the action to have been converted into a habit. It is worth noting, however, that the amount of training required to convert an action into a habit has varied substantially across experiments. For example, Corbit et al. (2013) contrastingly reported no evidence of habit formation until rats had received four weeks of daily instrumental training. Therefore, although the results of Experiment 1 demonstrate that instrumental responding is weakened to a similar extent after a wide range of training amounts with ratio or interval schedules, it is not clear whether the lowest amount of training examined led to the development of action-based or habit-based responding.

Experiment 2

Experiment 2 was designed to provide a test of the possible effect of context after reducing the amount of instrumental training further. Adams (1982) provided early evidence for sensitivity to reinforcer devaluation (and thus action-based responding) after 100 reinforced responses. Dickinson et al. (1995) demonstrated that responding was sensitive to reinforcer devaluation after 120 reinforced responses. Experiment 2 therefore tested the effect of switching the context on an instrumental response that was paired with the reinforcer only 90 times.

Training consisted of three sessions in which rats learned to lever press in Context A as well as three alternating sessions of exposure to Context B. One group received noncontingent pellets during the exposure sessions in Context B (Group B+), as in Experiment 1, while a second group (Group B-) received equal exposure to Context B without noncontingent pellet deliveries. All rats were then tested in Contexts A and B (order counterbalanced). Our previous evidence suggests that robust and equivalent context-switch effects occur in animals exposed to the alternate context with or without free deliveries of the reinforcer

(Bouton et al., 2014). Experiment 2 was designed to determine whether equivalent experience with pellets in each context is likewise irrelevant to producing the context-switch effect using the present methods.

Method

Subjects and apparatus—The subjects were 16 female Wistar rats purchased from the same vendor as those in the Experiment 1 and maintained under the same conditions. The apparatus was also the same.

Procedure

Instrumental training: All rats received one magazine training session in each context, in which no levers were present and 30 free pellets were delivered into the food cup according to a RT 60 s schedule. Beginning the next day, daily lever-press training session began with the insertion of a lever and ended after 30 reinforced lever-presses. Each day, rats were given one session in Context A in the morning and one session in Context B in the afternoon. In Context A, a lever was inserted after a 2-min delay and rats were allowed to earn 30 pellets according to a CRF schedule, whereupon the lever was retracted. Following a 1.5 hour delay, rats received a session in Context B with no lever present. Rats in Group B- were placed in Context B for a period of time matched to the duration of the preceding session in Context A. Rats in Group B+ received 30 pellets delivered according to the average rate earned in the preceding Context A session. The two types of sessions were repeated on the next two days, with the CRF schedule replaced first with a RR 5 and then a RR 15 schedule.

Testing: Rats were then tested in each context on the day following the final training session. Test sessions were 10-min in duration; the lever was present, but presses had no scheduled consequences. Order of testing (Context A first, Context B first) was counterbalanced in each group. Sessions occurred at least 1.5 hours apart.

Results

Acquisition of lever pressing in Groups B- and B+ proceeded normally. Responding in the two groups was assessed by a Group (2) by Session (3) ANOVA. The groups' response rates did not differ reliably, $F(1, 14) = 3.37$, $MSE = 83.96$, $p = .09$. Responding increased over sessions, $F(2, 28) = 49.06$, $MSE = 28.19$, $p < .001$, $\eta_p^2 = .78$, 95% CI [.58, .85], but there was no interaction of group and session, $F(2, 28) < 1$. On the final day, response rates in Group B- ($M = 20.25$) and Group B+ ($M = 27.69$) did not differ, $F(1, 14) = 2.53$, $MSE = 87.38$, $p = .13$.

The results of the 10-min tests in Contexts A and B are presented in Figure 3. They suggest that the overall level of responding was higher in the group that received pellets in Context B during training. But more important, a context switch effect was plainly evident in both groups. These observations were confirmed in a Group (2) by Context (2) repeated measures ANOVA. There was a Group effect, $F(1, 14) = 11.61$, $MSE = 23.39$, $p = .004$, $\eta_p^2 = .45$, 95%

CI [.06, .67], as well as a Context effect, $F(1, 14) = 4.44$, $MSE = 10.28$, $p = .05$, $\eta_p^2 = .24$, 95% CI [.00, .52]. The interaction did not approach reliability, $F(1, 14) < 1$.

Discussion

Experiment 2 found that a context switch still weakened instrumental responding after quite limited training. Free pellets in Context B appeared to increase the overall level of responding in both contexts in Group B+. However, the presence or absence of pellets during training, which could have led to different Context-O associations in the two groups, did not change the decrement in responding produced by a switch to Context B. The results replicate and extend earlier work that found a context-switch effect in a group that received exposure to Context B during training without pellet deliveries in a discriminated operant procedure (Bouton et al., 2014, Experiment 2).

Dickinson et al. (1983, 1995; see also Adams, 1982) have extensively documented sensitivity to reinforcer devaluation, and thus action-based responding, after 100 reinforced responses. Prior research thus strongly suggests that the lever pressing studied in Experiment 2 was a goal-directed action. However, to provide more definitive information on whether the context effect depends on the response's status as an action or a habit, the next two experiments examined the context-switch effect following a reinforcer devaluation treatment.

Experiment 3

The main goal of Experiment 3 was to assess the context-switch effect on responding that could be more definitively characterized as an action. Three groups received training in Context A on an RR 15 schedule in which the response was paired with the food pellet 90 times (as in Experiment 2). Following training, the animals underwent a reinforcer devaluation experience. One group received trials in which the food pellets were devalued by pairing them with illness created by LiCl injections. Two control groups received either LiCl unpaired with the pellets or saline injections instead of LiCl. Lever pressing was then tested in extinction in Contexts A and B. If the response is an action, it should be at least partly suppressed in the rats that received the reinforcer paired with LiCl. The question, then, was whether the context switch would influence an instrumental response that could be characterized that way.

Method

Subjects and apparatus—The subjects were 24 female Wistar rats purchased from the same vendor as those in the Experiment 1 and maintained under the same conditions. The apparatus was also the same.

Procedure

Instrumental training: On the first day of training, all rats received magazine training consisting of 30 pellet deliveries on a RT 60 s schedule in each context. On the next three days (Days 2-4), rats were given a morning session in Context A and an afternoon session in Context B. In Context A, the lever was inserted after a 2-min delay and the rats were

allowed to earn 30 pellets, whereupon the lever was retracted. Following a three-hour delay, rats received a session in Context B with no lever present. Rats were placed in Context B for a period of time equal to the time last spent in Context A. Noncontingent pellets were not delivered in B in order to minimize the amount of exposure to the pellets (and thus latent inhibition to them) prior to the aversion conditioning treatment that was conducted in the next phase. (Recall that the results of Experiment 2, like previous results of Bouton et al. [2014], suggest that the context switch effect is not affected by the presence or absence of noncontingent pellets in Context B during training.) The A-to-B cycle was then repeated on the next two days so that the rats received 30 pellets on CRF, RR 5, and then RR 15 in the three sessions conducted in Context A.

Aversion conditioning: On Day 5, rats were matched on RR 15 response rates and then assigned to the Paired, Unpaired, or Saline groups. Aversion conditioning with the pellet reinforcer then proceeded over the next 12 days. The procedure involved pairing the pellets with LiCl equally often in each context. Aversion trials occurred every other day and were separated by a context exposure trial on the intervening days. On the first day of each two-day cycle, Paired rats received 50 noncontingent pellets in a particular context according to the average obtained interreinforcement interval in the final training session. They were then removed from the chamber and given an immediate intraperitoneal (i.p.) injection of 20 ml/kg LiCl (0.15 M) and put in the transport box prior to being returned to their home cages. Unpaired rats received the same exposure to the chamber and an immediate LiCl injection without receiving any pellets. On day 2 of each cycle, all rats were given an i.p. injection of isotonic saline (20 ml/kg). On this day, unpaired rats received 50 noncontingent pellets according to the average obtained interreinforcement interval in the final training session, while paired rats received exposure to the chamber for the same amount of time as in the preceding pellet session. The third group, Group Saline, received the same order of pellets and exposure-only sessions as the Paired group, but received saline injections after all sessions. There were 6 two-day conditioning cycles arranged so that the Paired group received three pellet-LiCl pairings in each context. Half of the rats in each group received three repetitions of sessions in the two contexts in the order of ABBA or BAAB. In order to maintain equivalent pellet exposure during aversion conditioning, Unpaired and Saline groups were only allowed to consume the average number of pellets eaten by the Paired group on each pellet trial.

Testing: Testing was conducted on the next three days. On the first day of testing, all rats received one 10-min lever press extinction test in each context in a manner similar to Experiment 1. The test order was counterbalanced in each group. On the next day, the rats were given a test of pellet consumption in each context (order counterbalanced) in order to assess the strength of the aversion to the pellets. In each of these tests, the rat was placed in the chamber with the lever removed and 20 food pellets were delivered on an RT 30 s schedule. Magazine entries and number of pellets consumed were recorded. Finally, on the last day, the rats were allowed to press the lever to earn food pellets according to the RR 15 schedule in Context A.

Results

Acquisition of lever pressing proceeded uneventfully. A Group (3) by Session (3) repeated-measures ANOVA found no differences between groups, $F(2, 21) < 1$. There was a significant increase in responding over sessions, $F(2, 42) = 57.65$, $MSE = 22.14$, $p < .001$, $\eta_p^2 = .73$, 95% CI [.56, .81], but the session effect did not interact with group, $F(4, 42) < 1$. A one-way ANOVA comparing response rates in the final session in the Paired ($M = 21.05$), Unpaired ($M = 20.38$), and Saline ($M = 20.15$) groups found no differences, $F(2, 21) < 1$.

Aversion conditioning in the chambers also proceeded smoothly. On the last cycle, rats in the Paired group consumed zero pellets, whereas Unpaired and Saline rats consumed all of them.

Figure 4 shows the crucial results from the test of instrumental responding in both contexts after the devaluation treatment. The figure suggests the presence of both reinforcer devaluation and context switch effects. A Group (Paired vs. Unpaired vs. Saline) by Context (A vs. B) ANOVA confirmed a main effect of group, $F(2, 21) = 5.19$, $MSE = 29.15$, $p = .01$, $\eta_p^2 = .33$, 95% CI [.02, .54], as well as a main effect of Context, $F(1, 21) = 5.10$, $MSE = 11.26$, $p = .03$, $\eta_p^2 = .19$, 95% CI [.00, .45]. The group and context factors did not interact statistically, $F(2, 21) < 1$. Planned LSD comparisons of each group found that the Paired group responded significantly less than both the Unpaired and Saline groups, $p = .005$ and $.03$, respectively, which did not reliably differ from each other, $p = .42$.

The consumption test conducted on the day following the context switch test confirmed that aversion conditioning in the Paired group was complete. No rat in the Paired group consumed a single pellet, whereas animals in the Unpaired and Saline groups consumed all 20 of them. Food cup entries for each group during the pellet consumption test are shown in Figure 5. A Group by Context ANOVA found a significant group effect, $F(2, 21) = 15.24$, $MSE = 41.32$, $p < .001$, $\eta_p^2 = .59$, 95% CI [.24, .73], and no context effector Context \times Group interaction, largest $F = 1.06$. Thus, magazine entry rates were higher in groups that had not acquired a taste aversion to the food pellets, and did not reliably differ in either context.

Results from the reacquisition test in which all groups earned pellets in Context A are shown in Figure 6. Unpaired and Saline groups showed reliable reacquisition of instrumental responding across 5-min blocks, whereas Paired rats showed very few responses. Group differences were confirmed in a Group (3) by Block (6) repeated-measures ANOVA, $F(2, 21) = 22.02$, $MSE = 465.29$, $p < .001$, $\eta_p^2 = .68$, 95% CI [.36, .78]. There was also a significant main effect of Block, $F(5, 105) = 22.03$, $MSE = 29.59$, $p < .001$, $\eta_p^2 = .51$, 95% CI [.36, .59], and Group \times Block interaction, $F(10, 105) = 7.34$, $p < 0.001$, $\eta_p^2 = .41$, 95% CI [.21, .48]. Planned LSD comparisons confirmed that Unpaired and Saline groups responded more than the Paired group, $p < .001$, and $.001$, and did not differ from one another, $p = .54$.

Discussion

The relatively small amount of instrumental training used in Experiment 3 (as well as Experiment 2) produced a response that had the classic hallmark of a goal-directed action: Its strength was significantly influenced by reinforcer devaluation. And equally important, responding in all of the groups that received the instrumental training was decremented by the switch from Context A to Context B. The results of this experiment therefore suggest that the context influences the strength of a response that meets the definition of a goal-directed action.

However, a closer look at the data suggests important qualifications to this conclusion. The strength of the devaluation effect, or size of the difference between the Paired and Unpaired groups, appeared equal when tested in the context in which the response had been trained (A) and in the different context (B). If we take the difference between the response level in the Paired and Unpaired groups as an indication of the strength of the R-O association, then the R-O association appears to have been unaffected by the context change. Instead, it was the residual responding that remained after the reinforcer was devalued that appeared to weaken with the context switch. It is unlikely that this residual responding was due to an incomplete aversion conditioning to the pellets: The Paired group consumed *zero* pellets in both contexts at the end of the revaluation phase and in the test that followed the extinction test of the instrumental response. Although the residual responding that remained after reinforcer devaluation may result from several factors (Colwill & Rescorla, 1985a, b), since it is a component of the response that is insensitive to the current status of the reinforcer, it can be interpreted as due to responding that is controlled by habit rather than action learning. According to this line of reasoning, the context switch in this experiment appears to have weakened habit, but not action.

Experiment 4

Experiment 4 was designed to replicate and extend the results of Experiment 3 while also testing the effects of the context switch on a response that could be entirely attributed to habit. Two groups of rats were trained to lever press on random-interval (RI) schedules for either 3 or 12 sessions. Following the completion of training, they underwent a reinforcer devaluation treatment. Half the animals in each group were given presentations of the reinforcer paired with LiCl, whereas the other half were given the same number of injections and food pellets in an unpaired arrangement. Following complete rejection of food pellets in the paired groups, all animals were tested for lever pressing in each context under extinction conditions, given an opportunity to consume pellets, and then an opportunity to reacquire lever pressing, as in Experiment 3. Prior research suggests that the lever-pressing response should become a habit after extensive training on a RI schedule (Dickinson et al., 1983, 1995), a possibility that was explicitly tested here by examining the effect of reinforcer devaluation. If the context influences habit, as suggested by Experiment 3, then the context switch should decrease an extensively-trained response that is not suppressed by reinforcer devaluation.

Method

Subjects and apparatus—The subjects were 32 female Wistar rats purchased from the same vendor as those in the previous experiments and maintained under the same conditions. The apparatus was also the same.

Procedure

Instrumental training: All rats received one 30-min magazine training session in each context, in which no levers were present and free pellets were delivered according to RT 60 s schedule. Beginning the next day, daily lever-press training sessions began with the insertion of a lever and ended after 30 reinforced lever-presses. Training was staggered as in Experiment 1 so that all groups finished training on the same day. The response was reinforced on a RI 2-s schedule on the first session and then increased to 15-s and 30-s on the second and third. Rats in Group 360 were trained for a total of 12 sessions, allowing for 360 reinforced lever presses, and rats in Group 90 were trained for a total of three sessions, allowing for 90 reinforced lever presses. Rats received two sessions each day, a lever-press session in Context A, and an equal-duration session without pellets or lever available in Context B.

Aversion conditioning: Rats were then given devaluation training in a manner similar to Experiment 2. A 20 ml/kg i.p. injection of .15 M LiCl was given on odd-numbered days, and injections of equivalent volume saline were given on even-numbered days. On odd-numbered days, rats in the Paired group were placed in a chamber with no lever present and given 50 food pellets according to a RT 30 s schedule. Rats in the Unpaired group were exposed to the chamber for an equivalent time without pellets. On even-numbered days, Paired rats were exposed to a chamber for the amount of time yoked to the previous pellet trial without pellet deliveries, and Unpaired rats were presented with the average number of pellets eaten by the Paired rats in the preceding LiCl trial. Across LiCl trials, Unpaired groups (360 and 90) were presented with the average pellets consumed in the preceding trial by the respective Paired group in order to minimize any differences in pellet exposure. Trials were conducted in each context in either ABBA or BAAB orders (counterbalanced); thus groups were balanced for order of LiCl presentation in each context. Aversion conditioning continued until the Paired rats consumed zero pellets. By the end of the phase, Group Paired had received four aversion conditioning trials in each context.

Testing: Rats were tested in each context for lever pressing, tested for pellet consumption, and allowed to reacquire responding according to RI 30 s following the three-day procedure used in Experiment 2.

Results

Lever pressing was readily acquired in all groups. The Paired and Unpaired groups within each training condition were compared in separate Group by Session repeated-measures ANOVAs. Groups that received 12 sessions of training showed increased responding over sessions, $F(11, 154) = 80.49$, $MSE = 15.48$, $p < .001$, $\eta_p^2 = .85$, 95% CI [.80, .87], with no group \times session interaction, $F(11, 154) = 1.04$, $p = .41$. The Group effect did not approach

significance, $F(1, 14) < 1$. Similarly, groups that received 3 sessions of lever-press training did not differ during the training phase, $F(1, 14) < 1$. Responding increased across sessions, $F(2, 28) = 84.55$, $MSE = 9.80$, $p < .001$, $\eta_p^2 = .86$, 95% CI [.72, .90], but session did not interact with group, $F(2, 28) < 1$. On the final session of acquisition, Groups 360-Paired, 360-Unpaired, 90-Paired, and 90-Unpaired had mean lever-press rates of 36, 0, 33.5, 20.7, and 19.9, respectively. A Training Amount (2) by Devaluation (2) ANOVA found higher response rates in the 360 groups, $F(1, 28) = 31.72$, $MSE = 55.84$, $p < .001$, $\eta_p^2 = .53$, 95% CI [.25, .68], but no difference between Paired and Unpaired groups or interaction, $F_s < 1$.

Devaluation training proceeded uneventfully. The decline in consumption was more rapid in the 90-Paired group than in 360-Paired, reflecting greater latent inhibition in the 360-paired group (Lubow & Moore, 1959). However, by the final devaluation trial, Paired rats in both the 90 and 360 groups consumed zero pellets in both contexts.

The critical test results are presented in Figure 7. The pattern suggests that extensive training produced a response that was not sensitive to reinforcer devaluation, but was still affected by the context switch. The data were analyzed with a Training Amount (90 vs. 360 pellets) by Devaluation (Paired vs. Unpaired) by Context (A vs. B) repeated-measures ANOVA. There was a main effect of the amount of training, $F(1, 28) = 10.71$, $MSE = 32.49$, $p = .003$, $\eta_p^2 = .28$, 95% CI [.04, .49], but not of devaluation, $F(1, 28) = 3.71$, $p = .06$. However, there was an interaction between amount of training and devaluation, $F(1, 28) = 5.28$, $p = .03$, $\eta_p^2 = .16$, 95% CI [.00, .38]. The main effect of context was reliable, $F(1, 28) = 6.45$, $MSE = 9.08$, $p = .02$, $\eta_p^2 = .19$, 95% CI [.01, .41]. Thus, the context switch once again decreased the amount of responding. Importantly, the effect of context did not interact with either of the other factors, $F_s < 1$. Follow-up ANOVAs isolating the groups within each level of training confirmed a highly significant effect of reinforcer devaluation training in the 90 group, $F(1, 14) = 25.86$, $MSE = 11.23$, $p < .001$, $\eta_p^2 = .65$, 95% CI [.25, .79], and no statistical evidence of an effect of reinforcer devaluation on response rates in the 360 group, $F(1, 14) < 1$.

The nonsignificant interaction between the Context and Training Amount factors suggests that, as in Experiment 1, the context effect may not have depended on the amount of instrumental training. However, we also subjected data from the groups at each level of Training Amount to separate Devaluation (Paired vs. Unpaired) by Context (A vs. B) ANOVAs. The 360 groups showed significantly decreased responding in Context B, $F(1, 14) = 6.16$, $MSE = 12.52$, $p = .03$, $\eta_p^2 = .31$, 95% CI [.00, .57], with no other reliable effects or interactions, $F_s < 1$. The 90 groups contrastingly demonstrated a significant effect of devaluation, $F(1, 14) = 25.86$, $MSE = 11.21$, $p < .001$, $\eta_p^2 = .65$, 95% CI [.25, .79], with no other effect or interaction, $F_s < 1$. The lack of a Context effect in the 90 groups may have been due to a lack of power. When we combined the data of the current 90 groups with the corresponding 90 groups from Experiment 3, an Experiment (3 vs. 4) by Devaluation (Paired vs. Unpaired) by Context (A vs. B) ANOVA found significant main effects of Context, $F(1, 28) = 5.45$, $MSE = 7.90$, $p = .03$, $\eta_p^2 = .16$, 95% CI [.00, .39], and Devaluation, $F(1, 28) = 25.66$, $MSE = 22.19$, $p < .001$, $\eta_p^2 = .48$, 95% CI [.19, .65]. There was no reliable

Experiment by Context interaction, $F(1, 28) = 1.72$. This result, along with the analogous nonsignificant Context by Training Amount interaction in the present experiment, continues to suggest that there is a reliable context switch effect after only 90 reinforced instrumental responses.

Results from the consumption test are presented in Figure 8. The mean number of pellets eaten are shown in the top panel. Rats in the Unpaired groups consumed every pellet in each context, whereas rats in the Paired groups consumed an average of less than one pellet in each context. The bottom panel of Figure 8 shows the number of entries into the food cup. Paired groups approached and entered the magazine at a much lower rate than Unpaired groups. This was confirmed in a Training group (2) by Devaluation (2) by Context (2) repeated-measures ANOVA, which revealed a Devaluation main effect, $F(1, 28) = 72.32$, $MSE = 20.75$, $p < .001$, $\eta_p^2 = .72$, 95% CI [.50, .81]. Food cup entry rates were not dependent on amount of training, $F(1, 28) < 1$, and did not interact with devaluation treatment, $F(1, 28) < 1$. Additionally, magazine entries during the consumption test did not depend on the testing context or any interaction involving context, $F_s < 1$.

Results from the reacquisition test in which all groups earned pellets in Context A are presented in Figure 9. Unpaired groups showed reliable reacquisition of responding across 5-min blocks, whereas Paired groups showed very little responding. Extensively trained animals reacquired responding to a higher rate relative to those who received relatively minimal training. A Training Amount (2) by Devaluation, (2) by Block (6) repeated-measures ANOVA confirmed main effects of Training Amount, $F(1, 28) = 4.78$, $MSE = 245.82$, $p = .04$, $\eta_p^2 = .15$, 95% CI [.00, .37], and Devaluation, $F(1, 28) = 74.03$, $p < .001$, $\eta_p^2 = .73$, 95% CI [.51, .82]. There was no interaction between training amount and devaluation, $F(1, 28) = 1.94$, $p = .17$. There was a significant main effect of Block, $F(5, 140) = 3.65$, $MSE = 11.20$, $p = .004$, $\eta_p^2 = .11$, 95% CI [.01, .19], and Devaluation \times Block interaction, $F(5, 140) = 31.71$, $p < .001$, $\eta_p^2 = .53$, 95% CI [.40, .60]. Training amount did not interact with Block, $F(5, 140) < 1$, and there was no three-way interaction, $F(5, 140) = 1.12$, $p = .35$.

Discussion

Devaluation of the reinforcer suppressed responding after relatively minimal, but not more extensive, training. The latter result suggests that habit, rather than goal-directed action, was controlling the response after extensive training. Nevertheless, the context switch still had an effect; as in Experiment 1, there was no evidence of an interaction between the context switch effect and the amount of training. The results with relatively minimal training on an RI schedule replicated those in Experiments 2 and 3 (which involved an RR schedule). Once again, the strength of the devaluation effect was consistent across contexts, and the context switch appeared to affect responding that remained after devaluation. Overall, the results are consistent with the hypothesis that the context controls the habit component, but not the action component, of instrumental responding.

General Discussion

The results of the present experiments continue to confirm that instrumental responding is weakened when it is tested outside the context in which it is trained (e.g., Bouton et al., 2011, 2014; Todd, 2013). This context switch effect was observed in each of the present experiments, and was demonstrated after a wide variation in the amount of training (90 – 1260 response-reinforcer pairings), and after reinforcement according to either interval or ratio reinforcement schedules. A switch from Context A to Context B had an apparently equivalent impact on responding whether Context B had been associated with the reinforcer or not (Experiment 2; see also Bouton et al., 2014). Thus, the effect of changing the context on instrumental behavior is general, and appears to occur whether or not there are differential associations with the reinforcer in the training and testing contexts

The present experiments further examined the role of context in action-based and habitual instrumental behavior. In Experiments 3 and 4, actions and habits were distinguished, as is customary, by their sensitivity to reinforcer devaluation. In Experiment 3, when a relatively minimal amount of training on a ratio schedule permitted the development of a goal-directed action (as confirmed by sensitivity to reinforcer devaluation), responding was weakened after a context switch. This context-specificity of the response was observed whether the reinforcer had been devalued or not. Two additional findings were noteworthy. First, the degree to which devaluation suppressed responding was highly similar in both the context where responding had been trained (Context A) and in the different context (Context B). Thus, the rats' knowledge of the R-O relation (cf., Dickinson, 1994) appeared to be similar in the two contexts. Second, although the rats reduced their instrumental responding after reinforcer devaluation, that responding was not completely abolished. This suggests that once the goal-directed R-O component maintaining behavior is removed, some other component remained. And interestingly, it was the other, residual component that was weaker in Context B than it was in Context A. Since the residual component was insensitive to devaluation, it can be seen as habit. The results thus suggest that the context switch primarily affected the habit component of the instrumental response.

In Experiment 4, extensively trained responding was not affected by reinforcer devaluation, and therefore met the customary definition of a habit (Dickinson, 1994). However, consistent with the foregoing interpretation of the results of Experiment 3, this habit was nonetheless weakened when it was switched from Context A to Context B. In contrast, relatively minimal training again led to responding that was partially sensitive to reinforcer devaluation, and thus partly an action. As in Experiment 3, sensitivity to reinforcer devaluation was the same in each context, and the context switch primarily affected residual responding. The fact that there was no interaction between the effects of context, amount of training, and reinforcer devaluation further suggests that the context switch effect was due to the weakening of a component of responding that was independent of the other factors. Extensive training led to responding that was entirely insensitive to reinforcer devaluation. Yet, a component supporting insensitive responding was removed when the response was tested in Context B.

The notion that instrumental responding can be at once partly an action and partly a habit is consistent with previous research (Dickinson et al., 1995). Our results suggest that the

removal of the response from the acquisition context causes a loss of support of the habit. Thus, at least a part of the S-R association that supports habit is identifiably Context-R. There was no evidence of contextual sensitivity of the R-O action component of instrumental behavior. Experiments 3 and 4 show that after relatively few reinforced responses, responding was decreased by reinforcer devaluation equally in Contexts A and B. Once the R-O component had been removed, the context switch effect was still present. Experiment 4 also showed that once control by the goal-directed action component was removed by overtraining, the response remained sensitive to a context switch.

Previous studies of habitual control have universally relied on null results to identify habitual responding. That is, instrumental habits have been inferred when responding is not affected by aversion conditioning of the reinforcer (Adams, 1982), reinforcer-specific satiety (Balleine & Dickinson, 1998b), or other disruptive manipulations such as contingency reversal (Dickinson, Squire, Varga, & Smith, 1998). Yet, null results are notoriously difficult to interpret, and the lack of an effect of reinforcer devaluation on a response can potentially have multiple causes (e.g., see Rescorla, 1982; Colwill & Rescorla, 1985b, 1990, for discussion). One implication of the current results is that if the habit component of instrumental behavior is context-specific, then the context sensitivity of an instrumental response, a positive result, may provide a potential index of that component.

It should be noted that the context switch never abolished the instrumental response. As one especially important example, the extensively trained response in Experiment 4, which was identified as a habit because of its insensitivity to reinforcer devaluation, was still strong, though decremented, when it was tested in Context B. Although it is likely that there is some generalization between the present contexts, the incomplete drop in performance as a result of context change might also suggest that the context is not the only stimulus that controls a habitual response. Although several previous writers have suggested a role for context in evoking habit (e.g., de Wit & Dickinson, 2009), others have suggested that preceding *responses* in the behavior chain or sequence provide a stimulus for later responses (e.g., Dezfouli & Balleine, 2012, 2013). Thus, a switch to Context B would have removed contextual support for lever pressing, but not support from earlier responses in the chain that were still initiated. For instance, the sight of the lever might be sufficient to elicit approach to the lever regardless of the context. The approach response might produce proprioceptive feedback that then elicits some habitual lever pressing. The significant, but not complete, decrement in responding of the extensively-trained but devalued group when tested in Context B might be consistent with such a possibility.

Dickinson et al. (1995) developed a set of qualitative predictions for the development of habitual control. According to their descriptive model, "Habit strength" develops from zero as a linear function of the amount of training. Contrary to this description, our results suggest that a habit component might already exist after relatively minimal instrumental training. Moreover, it seems preserved, perhaps unchanged, after extensive training. (The fact that it is unchanged over training is suggested by the fact that the context effect did not interact with the amount of training in either Experiments 1 or 4.) It is possible that the growth of habit that occurs with extensive practice might occur with other, non-contextual, controllers of habit. As noted above, one potential contributor to habit strength is the

development of response sequences of different durations and initiation rates with different amounts of training and schedules of reinforcement (Brackney, Cheung, Neisewander, & Sanabria, 2011; Dezfouli & Balleine, 2012; Reed, 2011). Response sequence formation and other potential non-contextual factors warrant further investigation.

Finally, although the fact that the context played a role regardless of reinforcer devaluation treatments suggests that the context controlled responding independently of the R-O association. The present results are consistent with the idea that actions supported by R-O associations depend less on environmental stimuli than habits supported by S-R associations (Dickinson, 1994; de Wit & Dickinson, 2009). However, it is possible that contexts might also operate under other conditions by signaling that a specific response will produce a specific outcome in a specific context. Indeed, it appears that such *hierarchical* control is possible when the training procedure allows the context to signal or disambiguate response-outcome associations (Trask & Bouton, 2014; see also Colwill & Rescorla, 1990). In such cases, the context may enter into specific context-(R – O) associations that allow animals to display selective knowledge of which response will lead to which outcome.

Acknowledgments

This research was supported by Grant RO1 DA033123 from the National Institute on Drug Abuse to MEB. We thank Sydney Trask for comments on the manuscript.

References

- Adams CD. Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Quarterly Journal of Experimental Psychology*. 1982; 34B:77–98.
- Adams CD, Dickinson A. Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology*. 1981a; 33B:109–121.
- Adams, CD.; Dickinson, A. Actions and Habits: Variations in associative representations during instrumental learning. In: Spear, NE.; Miller, RR., editors. *Information Processing in Animals: Memory Mechanisms*. Hillsdale, N. J.: Erlbaum; 1981b.
- Balleine BW. Neural bases of food-seeking: Affect, arousal and reward in corticostriatal limbic circuits. *Physiology and Behavior*. 2005; 86:717–730. [PubMed: 16257019]
- Balleine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*. 1998a; 37:407–419. [PubMed: 9704982]
- Balleine BW, Dickinson A. The role of incentive learning in instrumental outcome revaluation by sensory-specific satiety. *Animal Learning and Behavior*. 1998b; 26:46–59.
- Balleine BW, O'Doherty JP. Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology Reviews*. 2010; 35:48–69. [PubMed: 19776734]
- Bouton ME, King DA. Contextual control of the extinction of conditioned fear: Tests for the associative value of the context. *Journal of Experimental Psychology: Animal Behavior Processes*. 1983; 9:248–265. [PubMed: 6886630]
- Bouton ME, Peck CA. Context effects on conditioning, extinction, and reinstatement in an appetitive conditioning preparation. *Animal Learning & Behavior*. 1989; 17:188–198.
- Bouton ME, Todd TP, León SP. Contextual control of discriminated operant behavior. *Journal of Experimental Psychology: Animal Learning and Cognition*. 2014; 40:92–105. [PubMed: 24000907]
- Bouton ME, Todd TP, Vurbic D, Winterbauer NE. Renewal after the extinction of free operant behavior. *Learning and Behavior*. 2011; 39:57–67. [PubMed: 21279496]

- Brackney RJ, Cheung THC, Neisewander JL, Sanabria F. The isolation of motivational, motoric, and schedule effects on operant performance: A modeling approach. *Journal of the Experimental Analysis of Behavior*. 2011; 96:17–38. [PubMed: 21765544]
- Colwill, RM. Associative representations of instrumental contingencies. In: Medin, DL., editor. *The Psychology of Learning and Motivation*. Vol. 31. San Diego, California: Academic Press; 1994. p. 1-72.
- Colwill RM, Rescorla RA. Postconditioning devaluation of a reinforcer affects instrumental responding. *Journal of Experimental Psychology: Animal Behavior Processes*. 1985a; 11:120–132.
- Colwill RM, Rescorla RA. Instrumental responding remains sensitive to reinforcer devaluation after extensive training. *Journal of Experimental Psychology: Animal Behavior Processes*. 1985b; 11:520–536.
- Colwill RM, Rescorla RA. Effect of reinforcer devaluation on discriminative control of instrumental behavior. *Journal of Experimental Psychology: Animal Behavior Processes*. 1990; 16:40–47. [PubMed: 2303793]
- Corbit LH, Nie H, Janak PH. Habitual alcohol seeking: Time course and the contribution of subregions of the dorsal striatum. *Biological Psychiatry*. 2013; 72:389–395. [PubMed: 22440617]
- Coutureau E, Killcross S. Inactivation of the infralimbic cortex reinstates goal-directed responding in overtrained rats. *Behavioural Brain Research*. 2003; 146:167–174. [PubMed: 14643469]
- Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience*. 2005; 8:1704–1711.
- de Wit S, Dickinson A. Associative theories of goal-directed behavior: a case for animal-human translational models. *Psychological Research*. 2009; 73:463–476. [PubMed: 19350272]
- Dezfouli A, Balleine BW. Habits, action sequences and reinforcement learning. *European Journal of Neuroscience*. 2012; 35:1036–1051. [PubMed: 22487034]
- Dezfouli A, Balleine BW. Actions, action sequences and habits: Evidence that goal-directed and habitual control are hierarchically organized. *PLoS Computational Biology*. 2013; 9:e1003364.10.1371/journal.pcbi.1003364 [PubMed: 24339762]
- Dickinson A. Actions and habits: The development of behavioral autonomy. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*. 1985; 308:67–78.
- Dickinson, A. Instrumental conditioning. In: Mackintosh, NJ., editor. *Animal Learning and Cognition Handbook of Perception and Cognition Series*. 2nd. San Diego, CA: Academic Press; 1994. p. 45-79.
- Dickinson, A. Goal-directed behavior and future planning in animals. In: Menzel, R.; Fischer, J., editors. *Animal Thinking: Contemporary issues in Comparative Cognition*. Cambridge, MA: MIT Press; 2011. p. 79-91.
- Dickinson A. Associative learning and animal cognition. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*. 2012; 367:2733–2742.
- Dickinson, A.; Balleine, BW. The role of learning in the operation of motivational systems. In: Pashler, H.; Gallistel, R., editors. *Steven' Handbook of Experimental Psychology Vol 3: Learning, Motivation, and Emotion*. New York, NY: Wiley; 2002. p. 497-534.
- Dickinson A, Balleine B, Watt A, Gonzalez F, Boakes RA. Motivational control after extended instrumental training. *Animal Learning and Behavior*. 1995; 23:197–206.
- Dickinson A, Nicholas DJ, Adams CD. The effect of instrumental training contingency on susceptibility to reinforcer devaluation. *Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology*. 1983; 35:35–51.
- Dickinson A, Squire S, Varga Z, Smith JW. Omission learning after instrumental pretraining. *Quarterly Journal of Experimental Psychology*. 1998; 51B:271–286.
- Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nature neuroscience*. 2005; 8:1481–1489.
- Gillan CM, Pappmeyer M, Morein-Zamir S, Sahakian BJ, Fineberg NA, Robbins TW, de Wit S. Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *American Journal of Psychiatry*. 2011; 168:718–726. [PubMed: 21572165]
- Graybiel AM. Habits, actions, and the evaluative brain. *Annual Review of Neuroscience*. 2008; 31:359–387.

- Harris JA, Jones ML, Bailey GK, Westbrook RF. Contextual control over conditioned responding in an extinction paradigm. *Journal of Experimental Psychology: Animal Behavior Processes*. 2000; 26:174–185. [PubMed: 10782432]
- Hogarth L, Balleine BW, Corbit LH, Killcross S. Associative learning mechanisms underpinning the transition from recreational drug use to addiction. *Annals of the New York Academy of Sciences*. 2013; 1282:12–24. [PubMed: 23126270]
- Killcross S, Coutureau E. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex*. 2003; 13:400–408. [PubMed: 12631569]
- Lubow RE, Moore AU. Latent inhibition: The effect of non-reinforced exposure to the conditional stimulus. *Journal of Comparative and Physiological Psychology*. 1959; 52:416–419.
- Reed P. An experimental analysis of steady-state response rate components on variable ratio and variable interval schedules of reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*. 2011; 37:1–9. [PubMed: 20718555]
- Rescorla, RA. Comments on a technique for assessing associative learning. In: Commons, ML.; Herrnstein, RJ.; Wagner, AR., editors. *Quantitative Analysis of Behavior: Acquisition*. Vol. 3. Cambridge, MA: Ballinger; 1982. p. 41-61.
- Robbins TW, Gillan CM, Smith DG, de Wit S, Ersche KD. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends in Cognitive Sciences*. 2012; 16:81–91. [PubMed: 22155014]
- Rosas JM, Todd TP, Bouton ME. Context change and associative learning. *WIREs Cognitive Science*. 2013; 4:237–244. [PubMed: 23772263]
- 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]
- Tanaka SC, Balleine BW, O'Doherty JP. Calculating consequences: Brain systems that encode the causal effects of actions. *Journal of Neuroscience*. 2008; 28:6750–6755. [PubMed: 18579749]
- Todd TP. Mechanisms of renewal after the extinction of instrumental behavior. *Journal of Experimental Psychology: Animal Behavior Processes*. 2013; 39:193–207. [PubMed: 23627796]
- Trask S, Bouton ME. Contextual control of operant behavior: Evidence for hierarchical associations in instrumental learning. *Learning and Behavior*. 2014; 42:281–288. [PubMed: 24961749]
- Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*. 2006; 7:464–476.

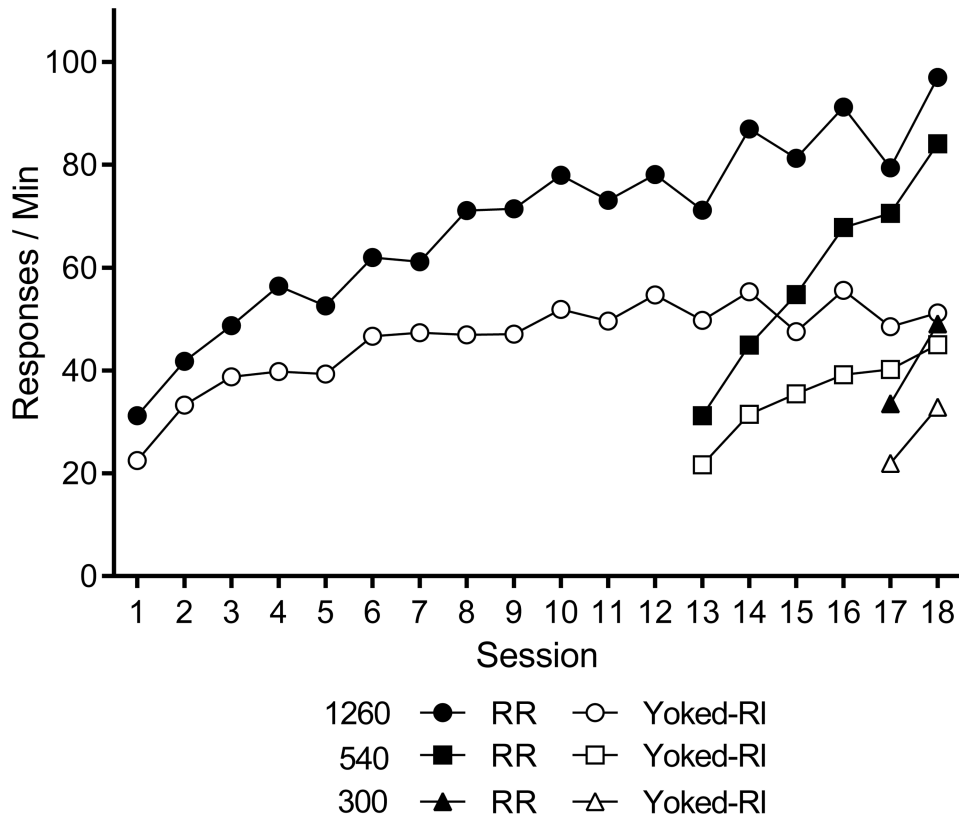


Figure 1. Results of Experiment 1. Mean response rates of the groups during acquisition in Context A. RR is Random Ratio, RI is Random interval, 1260, 540, 300 are total number of reinforced lever presses in training (including CRF and RR 5 sessions).

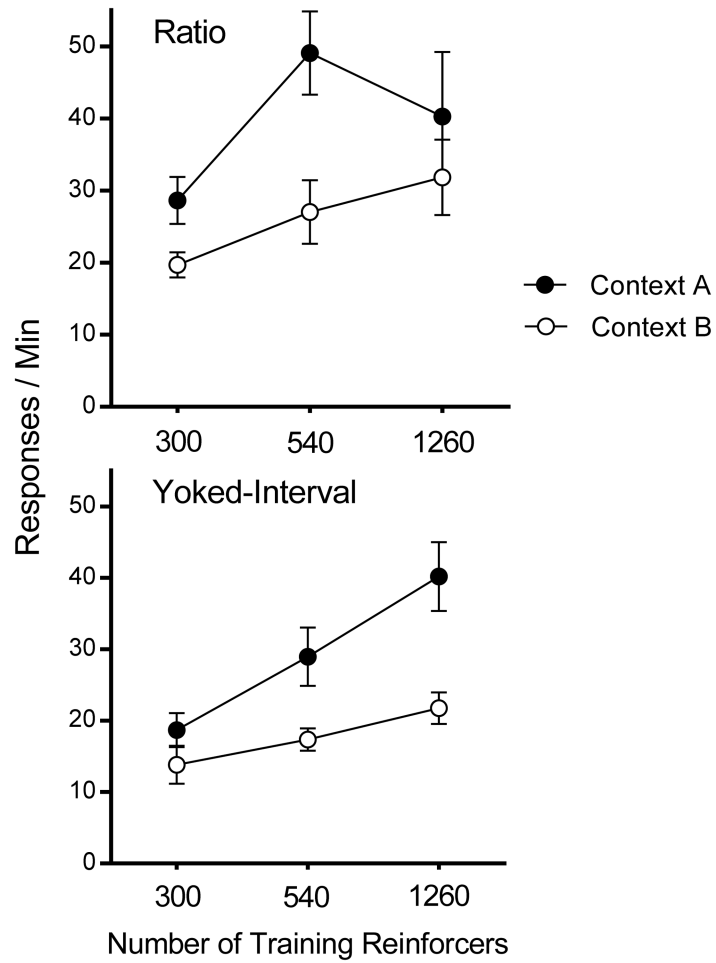


Figure 2.

Results of Experiment 1. Mean response rates during the extinction test in Contexts A and B plotted as a function of total number of reinforced responses in training. Error bars represent *SEM*. Error bars are included for comparison of between-group differences, and are not relevant for interpreting within-subject differences (i.e., the context effect).

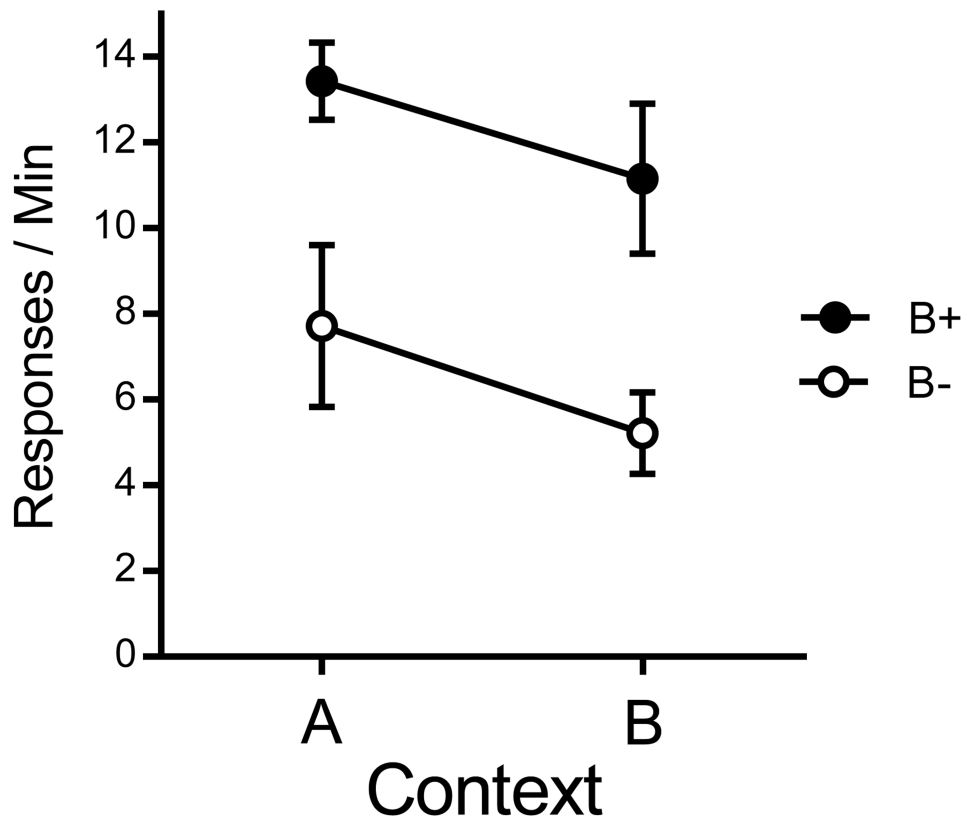


Figure 3. Results of Experiment 2. Mean response rates for each group in Contexts A and B. Error bars represent *SEM*. Error bars are included for comparison of between-group differences, and are not relevant for interpreting within-subject differences (i.e., the context effect).

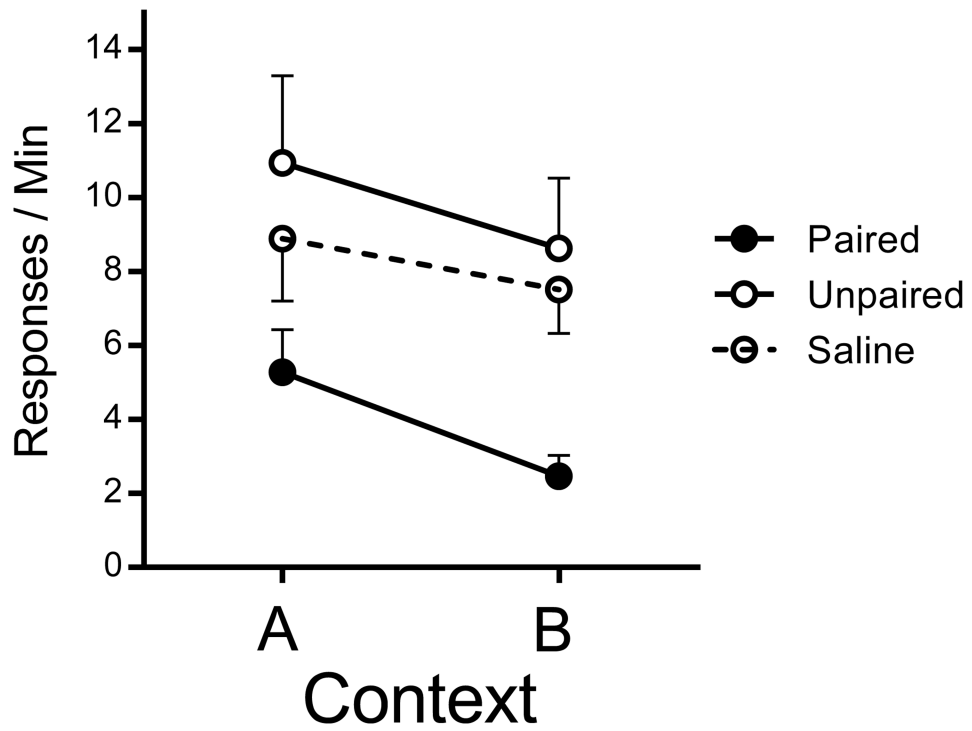


Figure 4. Results of Experiment 3. Mean response rates for each group in Contexts A and B. Error bars represent *SEM*. Error bars are included for comparison of between-group differences, and are not relevant for interpreting within-subject differences (i.e., the context effect).

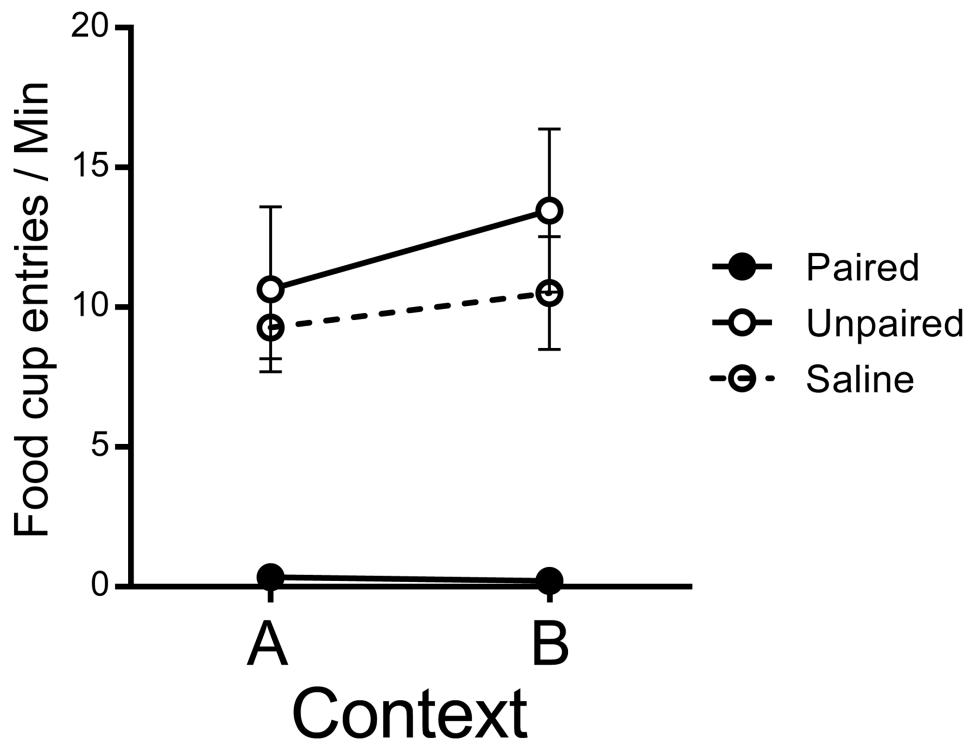


Figure 5. Results from Experiment 3. Mean magazine entry rates during the pellet consumption test in Contexts A and B. Error bars represent *SEM*. Error bars are included for comparison of between-group differences, and are not relevant for interpreting within-subject differences (i.e., the context effect).

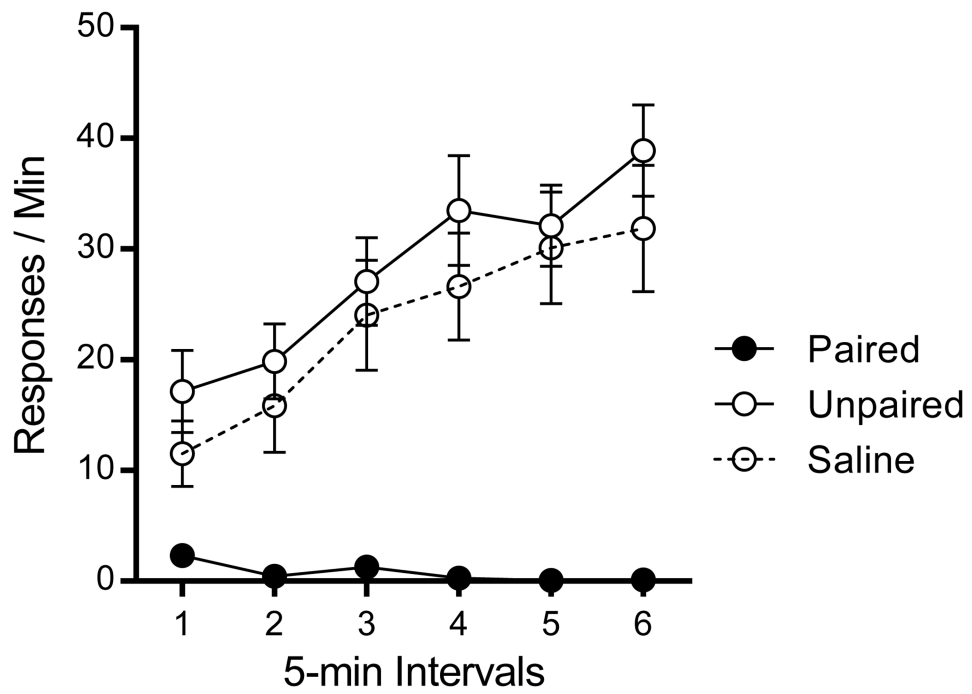


Figure 6. Results from Experiment 3. Mean response rates of the groups plotted as a function of 5-min bin in the reacquisition test session. Error bars represent *SEM*.

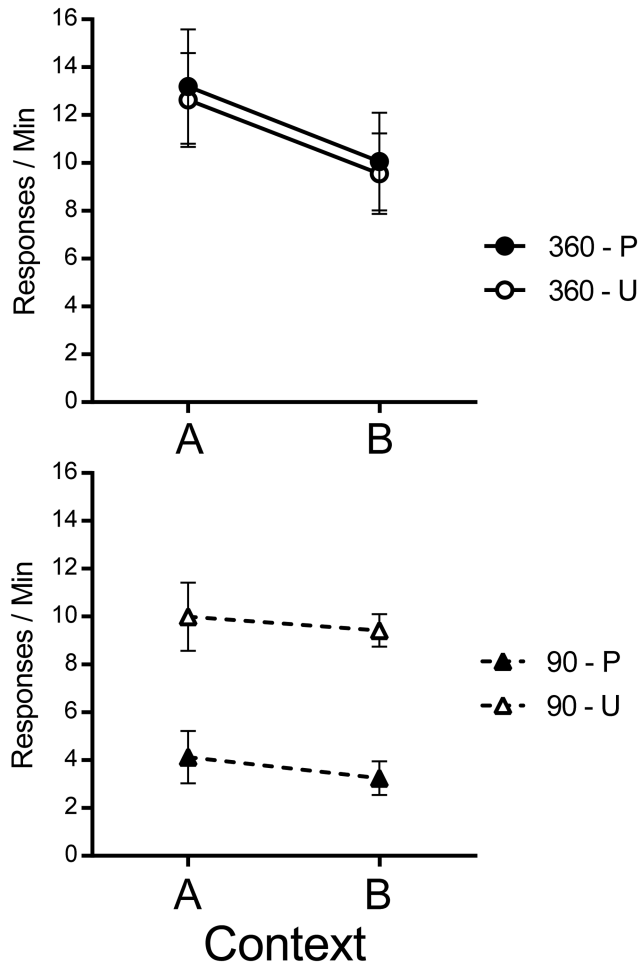


Figure 7.

Results of Experiment 4. Mean response rates for each group in Contexts A and B during the testing sessions. P is Paired, U is Unpaired, 360 and 90 refer to total number of reinforced lever presses in training. Error bars represent *SEM*. Error bars are included for comparison of between-group differences, and are not relevant for interpreting within-subject differences (i.e., the context effect).

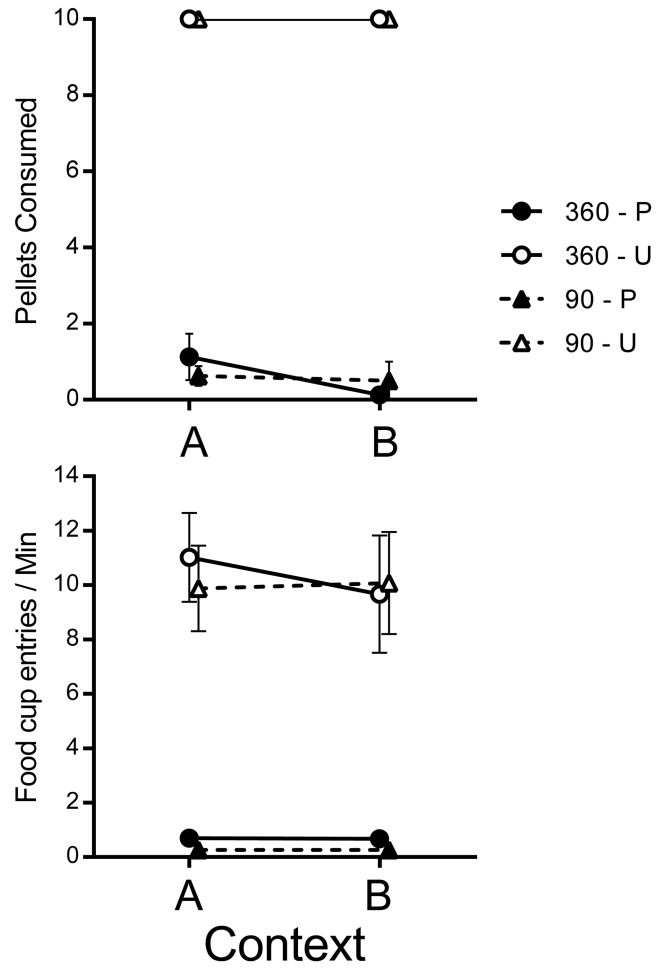


Figure 8. Results from Experiment 4. Mean number of pellets consumed (top panel) and mean magazine entry rates (Bottom) for the groups in the consumption test. P is Paired, U is Unpaired, 360 and 90 refer to total number of reinforced lever presses in training. Error bars represent *SEM*. Error bars are included for comparison of between-group differences, and are not relevant for interpreting within-subject differences (i.e., the context effect).

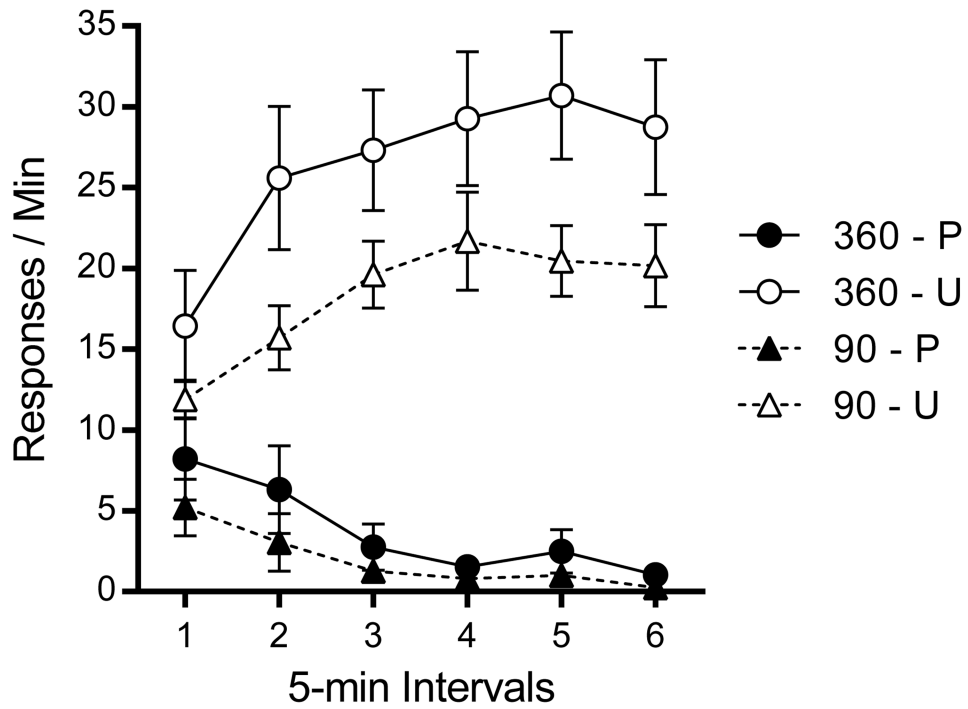


Figure 9.

Results from Experiment 4. Mean response rates for each group plotted as a function of 5-min bin in the reacquisition test. P is Paired, U is Unpaired, 360 and 90 refer to total number of reinforced lever presses in training. Error bars represent *SEM*.