Renewal after the punishment of free operant behavior

Mark E. Bouton and **Scott T. Schepers**
University of Vermont

Abstract

Three experiments examined the role of context in punishment learning. In Experiment 1, rats were trained to lever press for food in Context A and then punished for responding in Context B (by presenting response-contingent footshock). Punishment led to complete suppression of the response. However, when responding was tested (in extinction) in Contexts A and B, a strong renewal of responding occurred in Context A. In Experiment 2, renewal also occurred when initial reinforcement occurred in Context A, punishment occurred in Context B, and testing occurred in a new context (Context C). In both experiments, behavioral suppression and renewal were not observed in groups that received noncontingent (yoked) footshocks in Context B. In Experiment 3, two responses (lever press and chain pull) were separately reinforced in Contexts A and B and then punished in the opposite context. Although the procedure equated the contexts on their association with reinforcement and punishment, renewal of each response was observed when it was tested in its non-punished context. The contexts also influenced response choice. Overall, the results suggest that punishment is specific to the context in which it is learned, and establish that its context-specificity does not depend on a simple association between the context and shock. Like extinction, punishment may involve learning to inhibit a specific response in a specific context. Implications for theories of punishment and for understanding the cessation of problematic operant behavior (e.g., drug abuse) are discussed.

Keywords

Context; punishment; operant conditioning; renewal; relapse

Recent research suggests that extinction of instrumental behavior is specific to the context in which it is learned. For example, if operant lever pressing is reinforced in one context and then extinguished (allowed to occur without reinforcement) in a second context, extinguished responding will return if the response is tested again in the original context (e.g., Bouton, Todd, Vurbic, & Winterbauer, 2011; Nakajima, Tanaka, Urushihara, & Imada, 2000). Such ABA renewal, where conditioning, extinction, and testing occur in Contexts A, B, and A (respectively) has also been demonstrated in drug self-administration (e.g., Bossert, Liu, Lu, & Shaham, 2004; Crombag & Shaham, 2002) and in signaled shuttle box avoidance (Nakajima, 2014). In addition, the ABC and AAB forms of renewal have been reported (e.g., Bouton et al., 2011; Nakajima, 2014; Todd, 2013; Todd, Vurbic, & Bouton, 2014; Todd, Winterbauer, & Bouton, 2012; but see Bossert et al., 2004; Crombag &

Correspondence: Mark E. Bouton, Department of Psychology, University of Vermont, Burlington, VT 05405-0134, 802 656-4164, Mark.Bouton@uvm.edu.

Shaham, 2002). In these cases, renewal occurs in a context (C or B) that is different from the original conditioning context (A); they therefore suggest that mere removal from the context of extinction can be sufficient to produce a recovery of responding. The context of extinction thus appears to inhibit the response in some way. Recent research suggests that the inhibition might involve a direct inhibitory association between the context and the response (e.g., Todd, 2013; Todd, Vurbic, & Bouton, 2014; cf. Rescorla, 1993, 1997). In instrumental extinction, the animal may thus learn not to make a particular response in a particular context (see Bouton & Todd, 2014, for further review).

The context-dependence of extinction after instrumental learning is consistent with what is known about extinction after Pavlovian conditioning, where a conditioned stimulus (CS) is presented without the unconditioned stimulus (US) after the two have been associated in an earlier phase. The ABA, ABC, and AAB forms of renewal have been repeatedly demonstrated in several Pavlovian conditioning preparations (see Bouton & Woods, 2008; Vurbic & Bouton, 2014, for reviews). In the Pavlovian domain, extinction has been argued to be a representative form of retroactive interference or inhibition that occurs when learning in a second phase interferes with or inhibits performance acquired in the earlier phase (Bouton, 1993; Miller & Escobar, 2002). As one illustration, response recovery effects such as renewal, spontaneous recovery, and reinstatement have also been demonstrated after counterconditioning, where a CS is associated with one US (e.g., footshock or food pellet) and subsequently paired with a different US (e.g., food pellet or footshock; Bouton & Peck, 1992; Brooks, Hale, Nelson, & Bouton, 1995; Peck & Bouton, 1990). In any of a number of retroactive interference paradigms, performance based on what is learned in the first phase can return with various manipulations of context or time (e.g., Bouton, 1993). There is some evidence that this is true in instrumental learning, too. For example, if a negative contingency between the response and reinforcer is introduced after instrumental training, it produces a context-specific suppression of instrumental behavior. That is, ABA renewal has been demonstrated after behavior is suppressed by omission training (Nakajima, Urushihara, & Masaki, 2002; see also Bouton & Schepers, 2014; Kearns & Weiss, 2007).

The present experiments were directed at whether what is known about instrumental extinction similarly applies to another example of retroactive interference in instrumental learning. In *punishment*, an instrumental behavior that has been positively reinforced is subsequently suppressed by making an aversive event or stimulus contingent on it. Punishment is worth understanding in its own right (e.g., Azrin & Holz, 1966; Church, 1963; Solomon, 1964). But in addition, it is interesting from a translational perspective because humans who quit taking drugs, for example, may do so because they learn to appreciate the negative consequences of drug taking (e.g., Marchant, Khuc, Pickens, Bonci, & Shaham, 2013; Panlilio, Thorndike, & Schindler, 2003). Punishment also suppresses behavior despite the fact that the reinforcer remains available. These features of punishment may make it a better model than extinction of the cessation of human drug taking (e.g., Marchant et al., 2013). However, consistent with the idea that extinction and punishment nevertheless share common principles, punished behavior is known to recover if time elapses after punishment (e.g., Estes, 1944; see also Krasnova et al., 2014), as it does in spontaneous recovery after extinction (e.g., Rescorla, 2004). In addition, drug self-

administration behaviors that have been suppressed by punishment may be reinstated if the drug reinforcer is presented noncontingently after the behavior is suppressed (e.g., Panlilio et al., 2003).

There is also evidence that the suppressive effects of punishment, like those of extinction, may be context-specific. In an experiment reported by Marchant et al. (2013), rats were initially reinforced for lever pressing in Context A with alcohol. Then, during sessions conducted in a second context (Context B), footshock was presented contingent on responding, which was otherwise still reinforced with alcohol. After lever pressing was suppressed by punishment, the rats were returned to the original context (Context A) and lever pressing was tested in extinction. Responding was renewed on the return to Context A. However, the behavioral mechanism underlying contextual control was not established. One possibility is that the rats simply associated the footshock with Context B, and that lever pressing was suppressed there by contextual fear conditioning (e.g., Bouton & Swartzentruber, 1986). To control for this possibility, Marchant et al. gave another group a fixed number of noncontingent shocks during each session in Context B; the number of shocks (10) was chosen "because this was the approximate mean number of shocks/session given to the rats in group Punished during the first three sessions" (Marchant et al., 2013, p. 258). The procedure thus did not match the control and experimental groups on the number or temporal distribution of the shocks, and, because some rats in the punished group also received shock intensities that increased over sessions, the groups were not matched on shock intensity either. Experiments in a more recent report (Marchant et al., 2014) involved only rats that were reinforced in A and punished in B; renewal was replicated on return to A, but the experiments continue to leave the underlying behavioral mechanism(s) unknown.

The present experiments therefore examined the contextual control of punishment in more detail. Throughout, care was taken to isolate effects that were specific to true punishment learning created by response-contingent footshock. The results confirm that the effects of punishment can be context-specific, have implications for theories of punishment, and suggest that, analogous to extinction, animals subjected to punishment may learn not to make a specific response in a specific context.

Experiment 1

The first experiment examined ABA renewal after punishment. Over several sessions, two groups of rats were first reinforced with food pellets for lever pressing on a variable interval (VI) 30-s schedule of reinforcement in Context A. Once responding had stabilized, the groups were switched to Context B. Over the next several sessions, lever pressing in B continued to be reinforced on the VI 30-s schedule. However, brief footshocks (0.5-s, 0.6-mA) were also introduced. For one group (Group Punished), the shocks were delivered on a VI 90-s schedule. That is, a shock was presented for the first response made 90 s (on average) since the last shock had been earned. In contrast, each rat in a control group (Group Yoked) received a noncontingent shock at the same point in time that an animal in the punished group earned one. The yoking procedure equated the groups on the number and temporal distribution of shocks delivered during the phase. The groups therefore had

comparable opportunities to associate the context with shock, but only Group Punished received a true punishment contingency.

In a final test, the rats were allowed to lever press during brief sessions conducted in Contexts A and B, with order counterbalanced (e.g., Bouton et al., 2011). Responding was tested in extinction; that is, neither food pellets nor shocks were presented. If ABA renewal occurs after punishment, we would expect a recovery of responding in Context A for Group Punished. To the extent that the context controls responding that has been suppressed by true punishment learning enabled by response-contingent shock, we should not expect the same result in Group Yoked.

Method

Subjects—The subjects were 32 naïve female Wistar rats (*n*s = 16) purchased from Charles River Laboratories (St. Constance, Quebec). They were between 75 and 90 days old at the start of the experiment and were individually housed in suspended wire mesh cages in a room maintained on a 16:8-h light:dark cycle. The rats were food-deprived to 80% of their initial body weights by providing a controlled daily feeding throughout the experiment.

Apparatus—Conditioning proceeded in two sets of four standard conditioning boxes (Med-Associates, St. Albans, VT; model: ENV-008-VP) that were housed in different rooms of the laboratory. Boxes from both sets measured 31.75 × 24.13 × 29.21 cm (l × w × h) and were housed in sound-attenuation chambers. The front and back walls were aluminum; the sidewalls and ceiling were clear acrylic plastic. There was a 5.08 × 5.08 cm recessed food cup centered in the front wall near floor-level. 4.8-cm stainless steel retractable operant levers were located to the left and to the right of the food cup, 6.2 cm above the floor. The levers protruded 1.9 cm into the box when extended. Ventilation fans provided background noise of 60 dB(A), and illumination was provided by two 7.5-W incandescent bulbs mounted on the ceiling of the sound-attenuation chamber.

Each of the two sets of four boxes contained features that allowed them to be used as different contexts. In one set of boxes, the floor consisted of 0.48-cm diameter stainless steel grids spaced 3.81 cm and mounted parallel to the front wall. The ceiling and left sidewall had black horizontal stripes (3.81 cm wide). A distinct odor emanated continuously from a 5-ml dish containing a 2% anise solution (McCormick, Sparks, MD) outside of each box near the front wall. In the other set of boxes, the floor consisted of alternating stainless steel grids with different diameters (0.48 and 1.27 cm), spaced 1.59 cm. The ceiling and left sidewall were covered with dark dots (1.9 cm in diameter). A dish containing 5 ml of a 4% coconut solution (McCormick, Sparks, MD) was placed outside each box to provide odor.

Reinforcers were a 45-mg grain-based rodent food pellet (TestDiet, Richmond, IN, USA). Shocks could be delivered to box floors by Med Associates Aversive Stimulator Modules (model ENV-414). The equipment was controlled by a computer located in an adjacent room.

Procedure—All experimental sessions were 30 min in duration unless otherwise noted.

Magazine Training: On the first day, all rats were assigned to a box within each of the two sets of boxes representing Contexts A or B (counterbalanced). They then received two sessions in which pellets were delivered freely according to a Random Time (RT) 30-s schedule. The first session occurred in Context A and the second occurred in Context B approximately 2 hrs later. No response manipulanda were available during these sessions.

Acquisition: On each of the next 6 days, all rats received a single daily session in Context A, in which lever presses resulted in a pellet delivery every 30 s on average (a VI 30-s reinforcement schedule). Sessions began with insertion of the right-hand lever following a 2-min delay. No special response shaping was necessary. Sessions ended with the retraction of the lever after 30 min.

Punishment: On each of the next 4 days, all rats received one session in Context B where, after the usual 2-min delay, the right lever was inserted in the box and, as during acquisition sessions, presses were reinforced on the VI 30-s reinforcement schedule. However, for Group Punished, lever presses now also delivered a 0.6-mA 0.5-s shock on a VI 90-s schedule. The shock schedule was constrained by an interval being randomly selected without replacement from a pre-defined set of intervals (60 s, 75 s, 90 s, 105 s, or 120 s). Animals in Group Yoked also received 0.6-mA 0.5-s shocks. However, these shocks were not response contingent, but were instead delivered when a shock was earned by a linked master rat in the punished group. Sessions ended after 30 min, when the lever was retracted.

Renewal Test: On the final day, all rats received two 10-min test sessions. Following the usual 2-min delay, the right lever was inserted in the box and responses were counted but did not result in pellet or shock deliveries. Half the rats were first tested in Context A and half in Context B. Approximately 60 min later, each animal was given a second 10-min test session in the alternate context.

Data analyses: The results were evaluated with analyses of variance (ANOVAs) and follow-up t tests using a rejection criterion of $p < .05$. For ANOVAs with more than one factor, we report partial eta squared as our measure of effect size; for comparisons between two means, we report eta squared. For either measure of effect size, we computed 95% confidence intervals (CIs) using procedures described by Steiger (2004).

Results

The results of the acquisition, punishment, and test phases are shown in Figure 1. As suggested by the left panel of the figure, the two groups acquired lever responding and increased their rate of responding similarly over the 6 sessions of acquisition in Context A. This was confirmed by a 2 (Group) x 6 (Session) ANOVA, which found a significant main effect of Session, $F(5, 150) = 57.83$, $MSE = 18.91$, $p < .001$, $\eta_p^2 = .66$, 95% CI [.56, .72], but no Group effect or Group x Session interaction, Fs < 1.

The center panel of Figure 1 shows that responding initially declined in both groups when punishment was first introduced in Context B. However, Group Punished made fewer responses than Group Yoked in each session. A 2 (Group) x 4 (Session) ANOVA indicated a significant main effect of Session, $F(3, 90) = 9.89$, $MSE = 12.69$, $p < .001$, $\eta_p^2 = .25$, 95%

CI [.09, .37], a main effect of Group, $F(3, 90) = 91.63$, $MSE = 111.13$, $p < .001$, $\eta_p^2 = .75$, 95% CI [.66, .80], and a Group x Session interaction, $F(3, 90) = 25.04$, $MSE = 12.69$, $p < .001$, $\eta_p^2 = .45$, 95% CI [.29, .56]. Subsequent t-tests conducted to decompose the interaction confirmed that Group Punished made fewer presses than Group Yoked during all sessions, ts(30) > 4.28, $p < .001$; $\eta^2 = .38$, 95% CI [.11, .57]. Both groups received an average total of 18.5 shocks over the punishment sessions.

Renewal Test: The right panel of Figure 1 shows that during the renewal tests Group Punished, but not Group Yoked, exhibited a renewal of responding when tested in the acquisition context (Context A) compared to the punished context (Context B). A 2 (Group) x 2 (Context) ANOVA revealed a main effect of Context, $F(1, 30) = 31.74$, $MSE = 22.94$, $p < .001$, $\eta_p^2 = .51$, 95% CI [.24, .67], which confirmed differential responding between test contexts. The main effect of Group, $F(1, 30) = 41.24$, $MSE = 43.83$, $p < .001$, $\eta_p^2 = .58$, 95% CI [.32, .71], and the crucial Group x Context interaction, $F(1, 30) = 38.67$, $MSE = 22.94$, $p < .001$, $\eta_p^2 = .56$, 95% CI [.30, .70], were also reliable and indicated that the effect of context depended on group. Importantly, follow up comparisons confirmed that Group Punished made more responses in Context A than in Context B, $t(15) = 8.08$, $p < .001$; $\eta^2 = .69$, 95% CI [.46, .78]. Sixteen out of 16 rats in that group (100%) made more lever presses in Context A than in Context B. In contrast, Group Yoked exhibited no change in responding between Contexts A and B, $t(15) = 0.43$, $p = .67$.

Discussion

Rats in Group Punished showed a complete suppression of lever pressing by the end of the punishment phase in Context B. The fact that Group Yoked showed far less suppression indicates that suppression of behavior in the punished group was a true punishment effect. It is worth noting that the complete suppression of responding in Group Punished meant that Group Yoked was receiving few shocks (but could continue to earn pellets) by the end of the phase.

Consistent with the idea that punishment results in a context-specific suppression of responding, the punished group demonstrated a strong recovery of instrumental responding when it was tested in the original context (Context A). In fact, response recovery reached a level that was indistinguishable from that in the control group. The effects of punishment are thus strongly context-specific. The results amply support Estes's (1944) claim that "...the strength of a response at the end of a period of punishment is not a reliable index of its true state" (p. 16). In contemporary terms, punishment suppressed, but did not erase, the original behavior.

Experiment 2

The ABA renewal design employed in Experiment 1 (see also Marchant et al., 2013, 2014) does not distinguish between two processes that could separately contribute to the recovery of responding in Context A. One possibility, of course, is that punishment learning is context-specific. But in addition, the context in which responding had originally been reinforced might separately excite or activate responding during renewal testing there. Such

a possibility is consistent with results indicating that operant behavior can be strongly controlled by the context in which it is learned (e.g., Bouton et al., 2011; Bouton, Todd, & León, 2014; Thrailkill & Bouton, 2015). To separate the role of the two possibilities, it is necessary to ask whether punished behavior can be renewed when responding is tested in a *new* context that has never been associated with reinforcement or punishment of the response. The purpose of Experiment 2 was therefore to test whether ABC renewal occurs in the punishment situation.

The design of Experiment 2 was similar to that of Experiment 1. Two groups first received lever-press training in Context A followed by response-contingent or yoked shocks delivered while responding in Context B. After the second phase was complete, responding was tested in Context B and Context *C*, a new context in which operant responding had never been reinforced. If mere removal from the context of punishment is sufficient to cause a recovery of punished behavior, we should observe higher responding in Context C than Context B in Group Punished. Based on the results of Experiment 1, we expected no such difference in Group Yoked. In fact, since the test session in Context C was the first time that responding was tested in that context, it was possible that Group Yoked would contrastingly show *less* responding in Context C than in Context B (e.g., Bouton et al., 2011, 2014).

Method

Subjects—The subjects were 24 naïve female Wistar rats (*n*s = 12) purchased from the same vendor as those in the previous experiment. They were of similar age and maintained under the same conditions.

Apparatus—Three new sets of four conditioning boxes housed in separate rooms of the laboratory served as three contexts (counterbalanced). Each box was housed in its own sound-attenuation chamber. All boxes were of the same design (Med Associates model: ENV-008-VP). They measured $30.5 \times 24.1 \times 21.0$ cm ($l \times w \times h$). A recessed 5.1 cm × 5.1 cm food cup was centered in the front wall approximately 2.5 cm above the level of the floor. Twenty-eight-V panel lights (2.5 cm in diameter) were attached to the wall 10.8 cm above the floor and 6.4 cm both to the left and right of the food cup. The boxes were illuminated by one 7.5-W incandescent bulb mounted to the ceiling of the sound-attenuation chamber. Ventilation fans provided background noise of 65 dBA. The sidewalls and ceiling were made of clear acrylic plastic, while the front and rear walls were made of brushed aluminum.

In one set of boxes, the floor was made of stainless steel grids (0.48 cm diameter) staggered such that odd- and even-numbered grids were mounted in two separate planes, one 0.5 cm above the other. Retractable levers (Med Associates model: ENV-112CM) were positioned to the left and right of the food cup. These levers were 4.8 cm long and positioned 6.2 cm above the grid floor. The left lever protruded 1.9 cm into the box when extended (the right lever remained retracted over the course of this experiment). This set of boxes had no distinctive visual cues. Olfactory cues were continuously presented by placing a dish containing 5 ml of Rite Aid lemon cleaner (Rite Aid corporation, Harrisburg, PA) outside of each box near the front wall. The second set of boxes also contained features to provide a

distinct context. The grids of the floor were mounted on the same plane and were spaced 1.6 cm apart (center-to-center). The left sidewall and ceiling had black diagonal stripes, 3.8 cm wide and 3.8 cm apart. A distinct odor was created by placing 5 ml of Pine-Sol (Clorox Co., Oakland, CA) in a dish outside the box. The third set of boxes had the following noteworthy features. The grids of the floor were mounted on the same plane and were spaced 1.6 cm apart (center-to-center). A retractable lever, identical to those in the Lemon- and Pine-scented boxes was positioned on the left side of the food cup. A distinct odor was continuously presented by placing 5 ml of vinegar (Heinz, Pittsburgh, PA) in a dish outside the box. The three sets of boxes were counterbalanced as Contexts A, B, and C.

Reinforcers and shocks were the same as those in Experiment 1, and a computer located in an adjacent room again controlled the equipment.

Procedure

Magazine Training: On the first day, all rats were assigned to a box within each of the three sets of boxes representing distinct contexts A, B or C (counterbalanced). They then received three 30-min magazine training sessions (one in each context), separated by approximately 60 min, in which pellets were delivered freely on average every 30 s (VT 30-s). The first session was in Context A, followed by Context B and then Context C. Response manipulanda were not available during these sessions.

Acquisition: Lever press training then proceeded as described in Experiment 1, except that in this experiment animals responded on a lever to the left of the food cup. Both groups received these sessions in a counterbalanced box designated as Context A.

Punishment: This phase also followed the procedure described in Experiment 1 except that rats responded on a lever to the left of the food cup. As in Experiment 1, all punishment sessions occurred in Context B. Animals that received acquisition sessions in a lemon-scented box now received punishment in a pine-scented box. Animals that received acquisition in the pine-scented boxes received punishment sessions in a vinegar-scented box. Rats assigned to a vinegar-scented box during acquisition sessions received punishment in a lemon-scented box. Four rats (two from Group Punished and two from Group Yoked) were removed from the experiment and all analyses because a technical error prevented two animals in Group Punished from receiving shocks during the first and second sessions. Although this ultimately affected the completeness of counterbalancing, an ANOVA on the crucial renewal test data (see below) indicated no effects or interactions involving the box type factor, Fs(1, 14)   1.99.

Renewal Test: On the final day, all rats received two 10-min test sessions, where after the usual 2-min delay, the left lever was extended and responses were counted but no pellets or shocks were delivered. Half of the rats were first tested in the context that they previously received punishment (Context B) and half were tested first in the relatively novel context that had only been experienced during magazine training (Context C). Approximately 60 min after the first test session, each animal was tested in a second 10-min test in the other context.

Results

The results of the acquisition, punishment, and test phases are shown in Figure 2. As suggested by the left panel, both groups acquired lever pressing similarly in Context A over six sessions. This was confirmed by a significant main effect of Session, $F(5, 90) = 51.57$, $MSE = 10.66$, $p < .001$, $\eta_p^2 = .74$, 95% CI [.63, .79]. Neither the main effect of Group nor the Group x Session interaction were significant, $Fs < 1$.

The center panel of Figure 2 suggests that the introduction of shocks in Context B again resulted in an initial decline in lever pressing in both groups. However, Group Punished responded less over all sessions. In Contrast, lever pressing in Group Yoked increased across Sessions 2, 3, and 4. The main effect of Group, $F(1, 18) = 102.90$, $MSE = 83.05$, $p < .001$, $\eta_p^2 = .85$, 95% CI [.66, .91], and the Group x Session interaction, $F(3, 54) = 25.44$, $MSE = 10.14$, $p = .001$, $\eta_p^2 = .59$, 95% CI [.38, .68], were both reliable. The main effect of Session was not significant, $F(3, 54) = 2.02$, $MSE = 10.14$, $p = .12$. Follow up comparisons decomposing the interaction confirmed that Group Punished made fewer responses than Group Yoked during all four sessions, $ts(18)$ 3.38, p   .01, $\eta^2 = .39$, 95% CI [.06, .61]. The groups received an average total of 13.4 shocks during the phase.

Renewal Test: As suggested in the right panel of Figure 2, Group Punished exhibited a renewal of responding when tested in the relatively novel context (Context C). In contrast, Group Yoked actually made fewer lever presses in the novel context (Context C) than in the context in which non-contingent shocks had been received (Context B). The results of a 2 (Group) x 2 (Context) ANOVA revealed a significant Group x Context interaction, $F(1, 18) = 15.36$, $MSE = 12.98$, $p = .001$, $\eta_p^2 = .46$, 95% CI [.11, .66], as well as a main effect of Group $F(1, 18) = 18.39$, $MSE = 62.89$, $p < .001$, $\eta_p^2 = .50$, 95% CI [.15, .69]. The main effect of Context was not significant, $F < 1$. Follow up comparisons confirmed that Group Punished responded significantly more in Context C than in Context B, $t(9) = 3.20$, $p = .01$, $\eta^2 = .53$, 95% CI [.04, .74], demonstrating ABC renewal after punishment. Nine out of 10 rats in this group made more responses in Context C than Context B. In contrast, Group Yoked responded more in Context B than in the new Context C, $t(9) = 2.36$, $p < .05$, $\eta^2 = 30.22$, $\eta^2 = .38$, 95% CI [.00, .65]. Eight out of 10 rats in that group made *fewer* responses in Context C than Context B.

Discussion

The results of Experiment 2 demonstrated ABC renewal after punishment. They indicate that a switch out of the punishment context is sufficient to allow punished behavior to recover.

As in analogous studies of extinction (e.g., Bouton et al., 2011; Todd, Winterbauer, & Bouton, 2012), the size of the ABC renewal effect was not numerically large, and it appeared to be weaker than the ABA effect studied in Experiment 1 (although any cross-experiment comparison should be made with caution). However, it is worth reiterating that 90% of the rats in Group Punished lever pressed more in Context C than Context B during the test. Moreover, recovery of responding with the switch from Context B to Context C had to work against the fact that instrumental lever pressing is weaker when it is tested in a new

context (see also Bouton et al., 2011; Bouton et al., 2014; Todd, 2013), a finding that was replicated in the present yoked group. In the case of instrumental extinction, the strength of ABC renewal can be increased by providing more extended Phase 1 training (e.g., 12 rather than the present 6 sessions) or by conducting Phase 1 in more than one context (Todd et al., 2012). Thus, the ABC renewal effect observed here is important and convincing.

Experiment 3

The fact that ABC renewal was observed in Experiment 2 suggests that the inhibition of behavior that occurs in punishment is controlled at least partly by learning about the punishment context. The fact that no such effect was observed in the yoked group (see also Experiment 1) suggests that the punishment context did not merely suppress responding through its direct association with shock; the punished and yoked groups were equated on that. However, it is worth noting that, because responding was less suppressed in the yoked groups, the yoked groups also earned more reinforcers during the punishment phase than did the punished groups. Although that does not invalidate the conclusion that response suppression caused by response-contingent punishment is context-specific, one purpose of Experiment 3 was to use a method in which the number of reinforcers as well as the number of shocks could both be controlled. It employed a within-subject design similar to one used by Todd (2013) in an investigation of extinction (see Table 1). All rats were first reinforced for performing two instrumental responses (pressing a lever and pulling a chain). One response (R1, counterbalanced over lever and chain) was reinforced in Context A during sessions that were intermixed with sessions in which the other response (R2) was reinforced in Context B. During the punishment phase that followed, the two responses were switched and punished in the opposite context. That is, R1 was punished in Context B and R2 was punished in Context A. In final tests, the responses were tested in each context. The question was whether R1 and R2 would renew in their unpunished contexts. Since the two contexts were both associated with the reinforcement and punishment of an instrumental behavior, the idea that context controls punishment through simple contextual fear conditioning would not predict a renewal effect.

There were two renewal tests. In the first, R1 and R2 were simultaneously made available during test sessions in Contexts A and B. In the second, R1 and R2 were tested individually in separate sessions, in a manner more like that of Experiments 1 and 2. In either case, if the context controls suppression of a specific behavior, then the rats would make more R1 than R2 responses in Context A and more R2 than R1 responses in Context B. However, the test in which both responses were available simultaneously allowed us to ask whether contextual control of punishment learning can affect response *choice*.

Method

Subjects and Apparatus—The subjects were 16 naïve female Wistar rats of the same age and from the same vendor as those in the previous experiments. They were also maintained under the same conditions. The apparatus was the same as that used in Experiment 1, except that a response chain (Med Associates model ENV-111C) could now be suspended from a microswitch mounted on top (outside) of the ceiling. When inserted, the chain hung 1.9 cm from the front wall, 3.1 cm to the right of the food cup, and 6.2 cm

above the floor. The alternate response was provided by the lever positioned on the left of the food cup.

Procedure

Magazine Training: Magazine training proceeded as described in Experiment 1.

Acquisition: On each of the next 6 days, all rats were given two daily response-training sessions, one in Context A and the other in Context B. The sessions were separated by approximately 120 min. Context order was double alternated so that contexts were experienced equally often as the first and second session of the day. In Context A, only R1 (lever or chain, counterbalanced) was available. It was reinforced with pellets on a VI 30-s schedule immediately upon placement in the conditioning box. In Context B, only R2 (chain or lever, counterbalanced) was available and was also reinforced on a VI 30-s schedule.

Punishment: On each of the next 2 days, the rats received two sessions in which responding was now punished. On each day there was one session in each context, separated by approximately 2 hr. In Context A, R2 was now reinforced on the VI 30-s schedule, but in addition, 0.6-mA 0.5-s shocks were delivered on a VI 90-s schedule as in Experiment 1. In Context B, R1 was similarly reinforced on the VI 30-s schedule and punished on the VI 90-s schedule. During the two sessions on Day 1, the punishment schedule began after the rats had worked on the VI 30-s schedule for 10 min. During the sessions on Day 2, the reinforcement and punishment schedules were both in effect immediately upon placement in the box. All sessions ended at the end of 30 min.

Simultaneous Response Renewal Test: On the final day, each rat received two 10-min test sessions, one in each context (A and B), with the order counterbalanced. Test sessions were separated by approximately 60 min. In each context, both the lever and chain were simultaneously available for the first time. Responses were counted but no shocks or pellets were delivered at any time.

Single Response Renewal Test: Approximately 2 hr after completion of the simultaneous response renewal test, each rat received two additional 10-min test sessions (separated by approximately 60 mins) in which each response was tested alone in each context. In each session, only one response (lever or chain) was available and counted but neither pellets nor shocks were delivered. Half the rats were first tested with a response in the context in which it had been punished followed by a test of the same response in its non-punished context. The other half were tested in the reverse order. The final two sessions were conducted in the same punished-non-punished (or non-punished-punished) order with the alternate response. Therefore, for each rat, both R1 and R2 (counterbalanced) were separately tested in the context in which it had been punished and in the context in which it had only been reinforced.

Results

The results of the acquisition and punishment phases are presented in Figure 3. Acquisition of R1 and R2 proceeded uneventfully. A 2 (Response) x 6 (Session) ANOVA found a main

effect of Session, $F(5, 75) = 96.73$, $MSE = 13.83$, $p < .001$, $\eta_p^2 = 87$, 95% CI [.80, .90], but no Response x Session interaction or main effect of Response, Fs < 1. When each response was then punished in its alternate context, responding unsurprisingly declined. A 2 (Response) x 2 (Session) ANOVA identified a main effect of Session, $F(1, 15) = 133.18$, $MSE = 19.69$, $p < .001$, $\eta_p^2 = .90$, 95% CI [.73, .94], as well as a Response x Session interaction, $F(1, 15) = 16.08$, $MSE = 17.69$, $p = .001$, $\eta_p^2 = .52$, 95% CI [.12, .71]. The main effect of Response was not significant, $F(1, 15) < 1$, $MSE = 69.34$, $p = .47$. We have no explanation of the Response x Session interaction, although it is worth noting that all rats first received R2 punishment in Context A followed by R1 punishment in Context B. Importantly, there were no differences in R1 and R2 responding during the final day of punishment, $t(15) = 1.19$, $p = .25$. Over the phase, the rats received 11.6 and 48.4 shocks and food pellets in Context A and 11.8 and 51.3 shocks and food pellets in Context B.

Simultaneous Response Renewal Test: The results of the first renewal test, when both responses were available simultaneously in each context, are presented at left in Figure 4. Each response was relatively suppressed in its punishment context. A 2 (Response) x 2 (Context) ANOVA identified a significant Response x Context interaction $F(1, 15) = 20.18$, $MSE = 29.84$, $p < .001$, $\eta_p^2 = .57$, 95% CI [.18, .75]. The main effects of Response and Context were not significant, Fs < 1. Rats responded more on R1 in its renewal context (Context A) than in its punished context (Context B), $t(15) = 2.88$, $p = .01$, $\eta^2 = .36$, 95% CI [.02, .60], and similarly responded more on R2 in its renewal context (Context B) than in its punished context (Context A), $t(15) = 4.12$, $p = .001$, $\eta^2 = .53$, 95% CI [.14, .71].

Single Response Renewal Test: As indicated by the right panel of Figure 4, a similar pattern was observed when each response was then tested separately in each context. That is, here the rats also responded less in sessions when the available response was in its punished context. This was confirmed by a 2 (Response) x 2 (Context) ANOVA, which found a significant Response x Context interaction, $F(1, 15) = 12.70$, $MSE = 23.29$, $p < .01$, $\eta_p^2 = .46$, 95% CI [.08, .67]. Main effects of Response and Context were not significant, Fs < 1. Follow-up comparisons confirmed that animals responded more on R1 in its renewal context (Context A) than in its punished context (Context B), $t(15) = 3.06$, $p < .01$, $\eta^2 = .38$, 95% CI [.03, .62]. Similarly, the rats responded more on R2 in its renewal context (Context B) than in its punished context (Context A), $t(15) = 2.45$, $p < .05$, $\eta^2 = .29$, 95% CI [.00, .55].

Discussion

The results indicate that renewal can occur after punishment when each of the contexts is associated with shock. Instead of learning to inhibit responding because the punishment context is associated with shock, the animal learns to inhibit a particular response in a particular context. This was evident when R1 and R2 were tested either simultaneously or separately. The former result indicates that the contextual control of punishment can influence the organism's choice to perform R1 over R2 (in Context A) or R2 over R1 (in Context B).

The results continue to challenge the idea that the context controls punishment as a simple Pavlovian signal for shock. Although such control can almost certainly occur (because

shock-associated contexts can suppress operant behavior, e.g., Bouton & Swartzentruber, 1986), it is not necessary to produce the present punishment effects. The results of the experiment instead suggest that punishment leads the animal to learn to suppress performance of a specific response in a specific context. Two mechanisms that might capture this feature of the contextual control of punishment will be considered in the General Discussion.

General Discussion

The results of the present experiments indicate that the learning that occurs during punishment can be highly specific to the context in which it occurs. In Experiment 1, a switch out of the punishment context and return to the original training context produced a complete recovery of punished lever pressing. Rats given the same number and temporal distribution of noncontingent shocks showed no such effect. In Experiment 2, a switch out of the punishment context and into a new context also produced a renewal of behavior, indicating that a context switch is itself sufficient to renew punished behavior. Once again, no such effect was observed in a group given the same number and distribution of shocks in a noncontingent manner. It is worth noting that the ABC renewal in the punished group occurred despite the fact that operant responding in our laboratory is generally attenuated by a context switch (e.g., Bouton et al., 2011; Bouton et al., 2014; Todd, 2013). In Experiment 3, two responses that were separately reinforced in different contexts and then punished in the opposite context were renewed when they were tested in their original contexts. Those results demonstrated a high degree of specificity in what the animal learned during punishment. The results of the present experiments establish that the context-specificity of punishment does not depend on the animal merely learning to associate the punishment context with shock and that the renewal of punished behavior does not depend on excitation or activation of behavior in the original context, as in the ABA design (cf. Marchant et al., 2013, 2014).

Experiments 1 and 2 used a yoked control procedure to assess the importance of the response-shock contingency. Although that procedure equates response-contingent and noncontingent groups perfectly on the number and temporal distribution of the shocks they receive, Church (1964; see also 1963, 1989) noted a possible bias in the yoked-control design if there are individual differences in response-suppressing effects that the shock elicits. For example, consider the possibility that shock presentations can elicit an emotional response that suppresses operant behavior. As noted by Church (1989), "if an experimental subject required a few more shocks than its yoked control before it became suppressed, the experimental subject would make only a few more responses than its yoked control...; if a control subject required a few more shocks than its yoked experimental before it became suppressed, the control subject would continue to respond indefinitely" (p. 405). It is not possible to talk one's way out of the yoked-control artifact, although it is worth noting that the differences between responding in the punished and yoked groups observed in Experiments 1 and 2 were substantial. More important, we observed punishment/renewal effects that were quite specific to R1 and R2 in Experiment 3. Since shocks that were contingent on R1 were not contingent on R2 (and vice versa), the results of Experiment 3 isolate a role for response contingency without incurring a possible yoked-control artifact.

There can thus be little question of the importance of the response-shock contingency in creating the punishment effect that was shown to be context-specific there.

Several theories of punishment have been proposed over the years. The present demonstrations of renewal after punishment are clearly inconsistent with the idea that punishment causes unlearning (e.g., Thorndike, 1913), and the specific renewal and suppression effects on R1 and R2 in Experiment 3 also appear to rule out the possibility that punishment is due to the conditioning of an emotional response to the context (e.g., Estes, 1969) or to the learning of incompatible responses that compete with the target response (Guthrie, 1935; Solomon, 1964). The role of a conditioned emotional response is challenged by the importance of the response contingency observed in all three of the present experiments; and an incompatible behavior (e.g., freezing) would presumably compete equally with lever pressing or chain pulling, making it impossible to observe the response-specific renewal and suppression effects observed in Experiment 3. Instead, the results may be more compatible with the idea that the organism learns to associate the response with the punishing outcome (e.g., Bolles, Holtz, Dunn, & Hill, 1980; see also Bolles, 1972; Mackintosh, 1974). From this perspective, the punishment context might function as an occasion-setter (e.g., Holland, 1992) that signals that the response now leads to shock. There is evidence that animals can learn such hierarchical context-response-outcome relationships (Trask & Bouton, 2014). But one challenge for this view is that occasion setting is known to transfer across similarly-trained targets (e.g., Holland, 1992). Therefore, learning that Context A signaled an R2-shock relation (for example) would be expected to transfer and influence a similarly-trained target response (i.e., R1) in Experiment 3. To account for the response-specific effects in that experiment, one would need to assume that transfer of occasion-setting across target responses was incomplete (see Todd, 2013; Todd et al., 2014). An alternative possibility is that the punishment of responding might somehow encourage the animal to learn to inhibit the specific response in the punishment context, an idea that has been formalized as a direct inhibitory context-response association (e.g., Bouton & Todd, 2014; Rescorla, 1996; Todd, 2013; Todd et al., 2014).

It is worth noting that the occasion-setting and response inhibition accounts of punishment are directly analogous to recent accounts of instrumental extinction (e.g., Bouton & Todd, 2014; Todd, 2013; Todd et al., 2014). Indeed, one point of the present results is that they begin to establish strong parallels between punishment and extinction. Either phenomenon can be characterized as a retroactive interference effect in which performance is highly dependent on the context in which testing occurs. As is true of Pavlovian extinction, instrumental extinction may thus be regarded as representing a general instrumental retroactive interference process (e.g., Bouton, 1993). Of course, the fact that punishment uniquely involves presentation of an aversive event means that there may be nontrivial differences between punishment and extinction. For example, Panlilio, Thorndike, and Schindler (2005) reported that injection of a benzodiazepine tranquilizer (lorazepam) caused a recovery of punished but not extinguished instrumental responding. Although punishment and extinction will inevitably differ in their details, the present results encourage consideration of their similarities. Neither form of behavioral suppression necessarily entails unlearning, and performance after each is sensitive to the effects of context and time.

Finally, the present results have possible clinical implications. If punishment is used as a means of suppressing unwanted behavior in the clinic (e.g., Lerman & Vorndran, 2002), it should be borne in mind that the present demonstration of its context-specificity suggests that behaviors that are punished in one clinical setting might not be as suppressed in other contexts. Second, although punishment may be considered a more realistic model than extinction of the cessation of human drug taking (e.g., Marchant et al., 2013; Panlilio et al., 2003), a common set of behavioral principles may still apply. Thus, when drug-taking, over-eating, or gambling are suppressed by knowledge of the behavior's aversive consequences, lapse and relapse in the form of the renewal effect may be ready to occur in new contexts.

Acknowledgments

References

Azrin, NH.; Holz, WC. Punishment. In: Honig, WK., editor. Operant behavior: Areas of research and application. New York, NY: Appleton-Century-Crofts; 1966. p. 380-447.

Bolles RC. Reinforcement, expectancy, and learning. Psychological Review. 1972; 79:394–409.

Bolles RC, Holtz R, Dunn T, Hill W. Comparisons of stimulus learning and response learning in a punishment situation. Learning and Motivation. 1980; 11:78–96.

Bossert JM, Liu SY, Lu L, Shaham Y. A role of ventral tegmental area glutamate in contextual cue-induced relapse to heroin seeking. The Journal of Neuroscience. 2004; 24:10726–10730. [PubMed: 15564590]

Bouton ME. Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. Psychological Bulletin. 1993; 114:80–99. [PubMed: 8346330]

Bouton ME, Peck CA. Spontaneous recovery in cross-motivational transfer (counterconditioning). Animal Learning & Behavior. 1992; 20:313–321.

Bouton ME, Schepers ST. Resurgence of instrumental behavior after an abstinence contingency. Learning & Behavior. 2014; 42:131–143. [PubMed: 24366673]

Bouton ME, Swartzentruber D. Analysis of the associative and occasion-setting properties of contexts participating in a Pavlovian discrimination. Journal of Experimental Psychology: Animal Behavior Processes. 1986; 12:333–350.

Bouton ME, Todd TP. A fundamental role for context in instrumental learning and extinction. Behavioural Processes. 2014; 104:13–19. [PubMed: 24576702]

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 & Behavior. 2011; 39:57–67. [PubMed: 21279496]

Bouton, ME.; Woods, AM. Extinction: Behavioral mechanisms and their implications. In: Byrne, JH.; Sweatt, D.; Menzel, R.; Eichenbaum, H.; Roediger, H., editors. Learning and memory: A comprehensive reference. Vol. 1. Oxford, UK: Elsevier; 2008. p. 151-171.

Brooks DC, Hale B, Nelson JB, Bouton ME. Reinstatement after counterconditioning. Animal Learning & Behavior. 1995; 23:383–390.

Church RM. The varied effects of punishment on behavior. Psychological Review. 1963; 70:369–402. [PubMed: 14049776]

Church RM. Systematic effect of random error in the yoked control design. Psychological Bulletin. 1964; 62:122–131. [PubMed: 14199645]

Church, RM. The yoked control design. In: Archer, T.; Nisson, L., editors. Aversion, avoidance, and anxiety: Perspectives on aversively motivated behavior. Hillsdale, NJ: Erlbaum; 1989. p. 403-415.

Crombag HS, Shaham Y. Renewal of drug seeking by contextual cues after prolonged extinction in rats. Behavioral Neuroscience. 2002; 116:169–173. [PubMed: 11895178]

Estes WK. An experimental study of punishment. Psychological Monographs: General and Applied. 1944; 57:i–40.

Estes, WK. Outline of a theory of punishment. In: Campbell, BA.; Church, RM., editors. Punishment and aversive behavior. New York: Appleton-Century-Crofts; 1969. p. 57-82.

Guthrie, ER. The psychology of learning. New York: Harper; 1935.

Holland, PC. Occasion setting in Pavlovian conditioning. In: Medin, DL., editor. The psychology of learning and motivation. Vol. 28. New York: Academic Press; 1992. p. 69-125.

Kearns DN, Weiss SJ. Contextual renewal of cocaine seeking in rats and its attenuation by the conditioned effects of an alternative reinforcer. Drug and Alcohol Dependence. 2007; 90:193–202. [PubMed: 17451890]

Krasnova IN, Marchant NJ, Ladenheim B, McCoy MT, Panlilio LV, Bossert JM, Shaham Y, Cadet JL. Incubation of methamphetamine and palatable food craving after punishment-induced abstinence. Neuropsychopharmacology. 2014; 39:2008–2016. [PubMed: 24584329]

Lerman DC, Vorndran CM. On the status of knowledge for using punishment: Implications for treating behavior disorders. Journal of Applied Behavior Analysis. 2002; 35:431–464. [PubMed: 12555918]

Mackintosh, NJ. The psychology of animal learning. San Diego: Academic Press; 1974.

Marchant NJ, Khuc TN, Pickens CL, Bonci A, Shaham Y. Context-induced relapse to alcohol seeking after punishment in a rat model. Biological Psychiatry. 2013; 73:256–262. [PubMed: 22883434]

Marchant NJ, Rabei R, Kaganovsky K, Caprioli D, Bossert JM, Bonci A, Shaham Y. A critical role of lateral hypothalamus in context-induced relapse to alcohol seeking after punishment-imposed abstinence. The Journal of Neuroscience. 2014; 34:7447–7457. [PubMed: 24872550]

Miller RR, Escobar M. Associative interference between cues and between outcomes presented together and presented apart: An integration. Behavioural Processes. 2002; 57:163–185. [PubMed: 11947996]

Nakajima S. Renewal of signaled shuttle box avoidance in rats. Learning and Motivation. 2014; 46:27–43.

Nakajima S, Tanaka S, Urushihara K, Imada H. Renewal of extinguished lever-press responses upon return to the training context. Learning and Motivation. 2000; 31:416–431.

Nakajima S, Urushihara K, Masaki T. Renewal of operant performance formerly eliminated by omission or noncontingency training upon return to the acquisition context. Learning and Motivation. 2002; 33:510–525.

Panlilio LV, Thorndike EB, Schindler CW. Reinstatement of punishment-suppressed opioid self-administration in rats: An alternative model of relapse to drug abuse. Psychopharmacology. 2003; 168:229–235. [PubMed: 12845420]

Panlilio LV, Thorndike EB, Schindler CW. Lorazepam reinstates punishment-suppressed remifentanil self-administration in rats. Psychopharmacology. 2005; 179:374–382. [PubMed: 15821953]

Peck CA, Bouton ME. Context and performance in aversive-to-appetitive and appetitive-to-aversive transfer. Learning and Motivation. 1990; 21:1–31.

Rescorla RA. Inhibitory associations between S and R in extinction. Animal Learning & Behavior. 1993; 21:327–336.

Rescorla RA. Response inhibition in extinction. The Quarterly Journal of Experimental Psychology: Section B. 1997; 50:238–252.

Rescorla RA. Spontaneous recovery. Learning & Memory. 2004; 11:501–509. [PubMed: 15466300]

Solomon RL. Punishment. American Psychologist. 1964; 19:239–253.

Thorndike, EL. The psychology of learning. Vol. 2. Teachers College, Columbia University; New York: 1913.

Thrailkill EA, Bouton ME. Contextual control of instrumental actions and habits. Journal of Experimental Psychology: Animal Learning and Cognition. 2015

Todd TP. Mechanisms of renewal after the extinction of instrumental behavior. Journal of Experimental Psychology: Animal Behavior Processes. 2013; 39:193–207. [PubMed: 23627796]

Todd TP, Vurbic D, Bouton ME. Mechanisms of renewal after the extinction of discriminated operant behavior. Journal of Experimental Psychology: Animal Learning and Cognition. 2014; 40:355–368. [PubMed: 25545982]

Todd TP, Winterbauer NE, Bouton ME. Effects of the amount of acquisition and contextual generalization on the renewal of instrumental behavior after extinction. Learning & Behavior. 2012; 40:145–157. [PubMed: 22002545]

Trask S, Bouton ME. Contextual control of operant behavior: Evidence for hierarchical associations in instrumental learning. Learning & Behavior. 2014; 42:281–288. [PubMed: 24961749]

Vurbic, D.; Bouton, ME. A contemporary behavioral perspective on extinction. In: McSweeney, FK.; Murphy, ES., editors. The Wiley-Blackwell handbook of operant and classical conditioning. Chichester, UK: John Wiley & Sons, Ltd; 2014. p. 53-76.
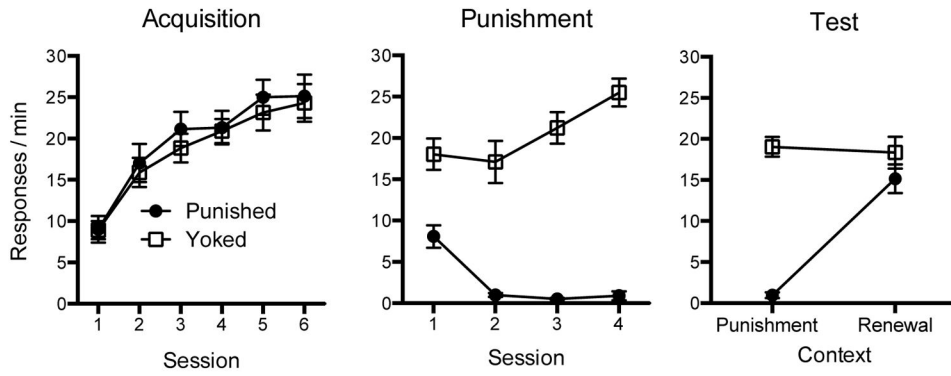
Figure 1.
Results of Experiment 1. Mean responses per minute during each 30-min acquisition session in Context A (left) and sessions in which punishment was introduced in Context B (middle). The right panel shows the mean responses per minute during each 10-min test session in the punishment (Context B) and renewal (Context A) contexts. Error bars represent the standard error of the mean (*SEM*). Error bars are included to help interpret between-group differences, and are not relevant for interpreting within-subject differences (i.e., the context effect during the renewal test).
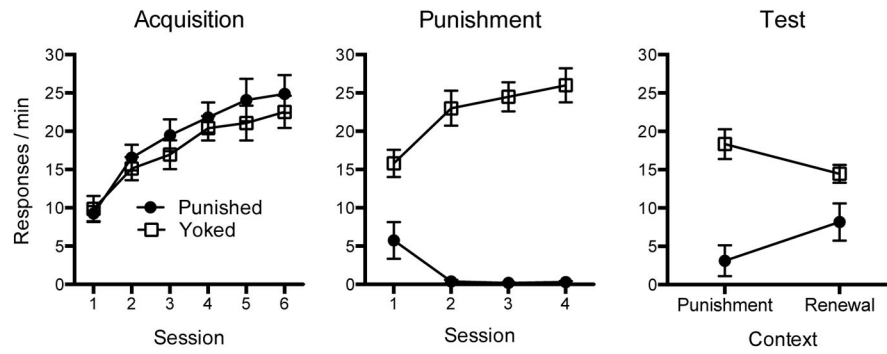
Figure 2.
Results of Experiment 2. Mean responding per minute during each 30-min acquisition session in Context A (left) and sessions in which punishment was introduced in Context B (middle). The right panel shows the mean responses per minute during each 10-min test session in the punishment (Context B) and renewal (Context C) contexts. Error bars represent *SEM*. Error bars are included to help interpret between-group differences, and are not relevant for interpreting within-subject differences (i.e., the context effect during the renewal test).
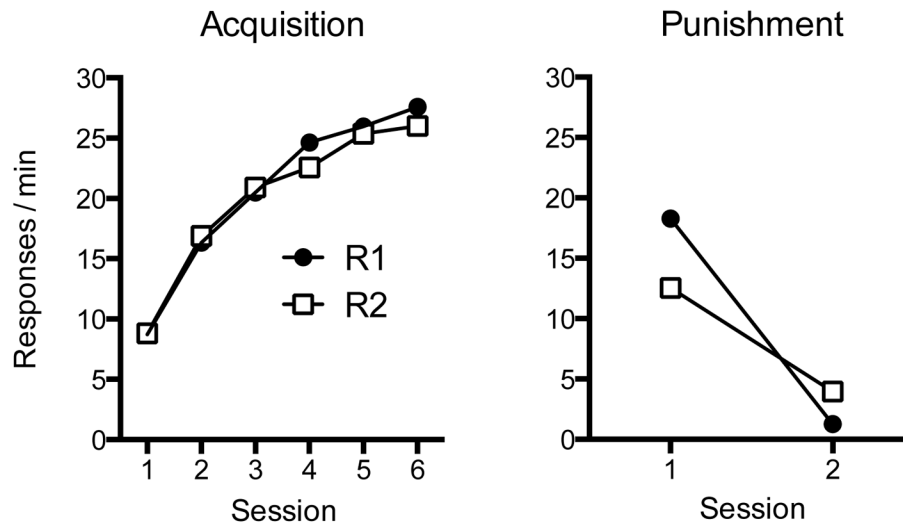
Figure 3.
Results of Experiment 3. Mean responses for R1 and R2 (lever or chain, counterbalanced) during each 30-min session of acquisition (left) and sessions in which punishment was added (right). Acquisition of R1 occurred in Context A and R2 in Context B; each response was punished in the opposite context. *SEMs* are not shown because all available comparisons are within-subject.
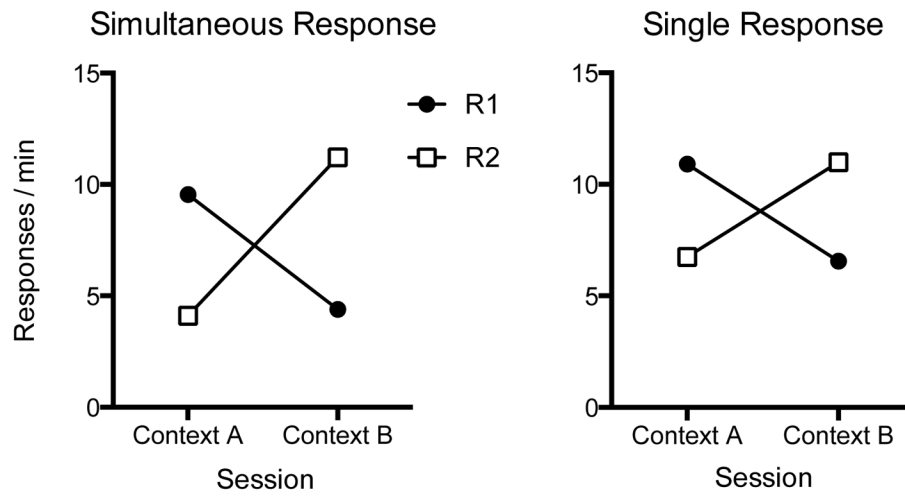
Figure 4.
Results of the Experiment 3 renewal tests. Left: Mean responses per minute during 10-min tests in each context when both responses were simultaneously available. Right: Mean responses per minute during four 10-min tests in which each response was tested separately in each context. *SEMs* are not shown because all available comparisons are within-subject.

Table 1

Design of Experiment 3

Acquisition	Punishment	Test 1	Test 2
A: R1: pellet	A: R2: pellet + shock	A: (R1- vs. R2-)	A: R1-; B: R2-
B: R2: pellet	B: R1: pellet + shock	B: (R1- vs. R2-)	A: R2-; B: R1-

Note: The design was within-subject. A and B refer to contexts. R1 and R2 refer to lever press or chain pull, counterbalanced. - = no programmed consequences. During acquisition and punishment sessions, pellets were always available on a VI 30-s schedule. In punishment sessions, 0.5-s, 0.6-mA shocks were also delivered on a VI 90-s schedule.