



# Retrieval practice after multiple context changes, but not long retention intervals, reduces the impact of a final context change on instrumental behavior

Sydney Trask<sup>1</sup> · Mark E. Bouton<sup>1</sup>

Published online: 12 December 2017  
© Psychonomic Society, Inc. 2017

**Abstract** Recent evidence from this laboratory suggests that a context switch after operant learning consistently results in a decrement in responding. One way to reduce this decrement is to train the response in multiple contexts. One interpretation of this result, rooted in stimulus sampling theory, is that conditioning of a greater number of common stimulus elements arising from more contexts causes better generalization to new contexts. An alternative explanation is that each change of context causes more effortful retrieval, and practice involving effortful retrieval results in learning that is better able to transfer to new situations. The current experiments were designed to differentiate between these two explanations for the first time in an animal learning and memory task. Experiment 1 demonstrated that the detrimental impact of a context change on an instrumental nose-poking response can be reduced by training the response in multiple contexts. Experiment 2 then found that a training procedure which inserted extended retention intervals between successive training sessions did not reduce the detrimental impact of a final context change. This occurred even though the inserted retention intervals had a detrimental impact on responding (and, thus, presumably retrieval) similar to the effect that context switches had in Experiment 1. Together, the results suggest that effortful retrieval practice may not be sufficient to reduce the negative impact of a context change on instrumental behavior. A common elements explanation which supposes that physical and temporal contextual cues do not overlap may account for the findings more readily.

**Keywords** Context · Instrumental learning · Operant conditioning · Variable training · Retrieval practice

Recent evidence suggests that operant behavior is highly dependent on the context (i.e., the physical environment where learning takes place, differing in tactile, visual, and olfactory properties) in which it is learned. For example, Bouton, Todd, Vurbic, and Winterbauer (2011) found that rats reinforced for lever pressing in Context A and then extinguished in Context B responded less in extinction than animals whose responding was extinguished in Context A. This context-change effect is a general phenomenon that occurs after training with both free-operant (Bouton et al., 2011; Trask, Shipman, Green, & Bouton, 2017) and discriminated-operant (Bouton, Todd, & León, 2014; Thraillkill, Trott, Zerr, & Bouton, 2016) procedures,

using either ratio and interval schedules of reinforcement, and after minimal or extended training (Thraillkill & Bouton, 2015). Similar results have been shown in delay matching-to-sample tests in humans (Dibbets, Maes, & Vossen, 2002). The results suggest the context is an important part of the associative structure that underlies instrumental learning (Bouton & Todd, 2014; Gámez, León, & Rosas, 2017; Thraillkill & Bouton, 2015; Trask & Bouton, 2014).

Developing methods that encourage instrumental behavior to transfer across contexts would have positive implications for therapies that encourage development of new behaviors. As just one example, in functional communication training, children with behavioral problems are reinforced for exhibiting prosocial behaviors and not reinforced for disruptive behaviors (e.g., Sprague & Horner, 1992). Finding ways to encourage better transfer of prosocial behavior to novel situations might improve the effectiveness of the treatment.

Todd, Winterbauer, and Bouton (2012) hypothesized that instrumental training in multiple contexts should encourage generalization between contexts by conditioning more contextual elements (the specific sensory cues that comprise the

---

✉ Sydney Trask  
sydney.trask@uvm.edu

✉ Mark E. Bouton  
mark.bouton@uvm.edu

<sup>1</sup> Department of Psychological Science, University of Vermont, Burlington, VT 05404-0134, USA

context), increasing the odds of encountering those or similar elements in a new context. The idea has roots in Estes's stimulus-sampling theory (Estes, 1955a, 1955b). They tested it using an ABC renewal paradigm, in which responding is acquired in one context (A), then extinguished in a second context (B). Responding then returns (or “renews”) when tested in a novel Context C. This effect likely depends (in part) on transfer of responding learned in Context A to Context C. In their experiment, before extinction in Context B, lever-pressing was reinforced in either one context (A) or two contexts (A and D). Following extinction, animals were tested in both the extinction context and a new context (C). While both groups showed ABC renewal, it was larger in animals that received acquisition training in two contexts, suggesting they showed better excitatory generalization to novel contexts. Further, rats that had received training in two contexts showed better transfer to the extinction context than those that had received acquisition training in only one context.

Contextual control of instrumental responding parallels research on human memory demonstrating that recall of word lists or pairings can also be impaired by context change (e.g., Godden & Baddeley, 1975; Smith, 1979; Smith, Glenburg, & Bjork, 1978; Smith & Vela, 2001). Moreover, human recall can also improve with practice that involves training in variable environments (Smith et al., 1978). Recently, Smith and Handy (2014) had participants learn several face–name pairs where each pairing was superimposed on a unique “video context” (moving scenery behind the image of the face and name during initial study). During subsequent retrieval practice trials, participants saw each face and wrote down the associated name. The face was always superimposed on the original context or was always presented on a new context. Initially, the context changes impaired performance: Participants in the varied condition showed worse name recall when shown the face throughout acquisition than participants who retrieved the name in the same context. However, when both groups were tested for recall of the name given the face in a new context, recall was higher if practiced in the varied condition. Further research showed that actively recalling the name given the face after context change, rather than simply being given the face–name pairing again (where no retrieval is needed), was necessary for this effect to occur. Thus, difficult retrieval of the item was needed in order to improve transfer, and experience with more contextual elements was not sufficient. The authors suggested conditions that created more difficult learning improved transfer to new situations.

The results were said to be consistent with the “new theory of disuse” (Bjork & Bjork, 1992), which suggests memory strength is determined by two independent processes: retrieval strength (how easily accessible the memory is) and storage strength (how well the item is represented in long-term memory). When a learned item is retrieved, its storage strength increases. Retrieval strength is dictated primarily by the match

between the cues currently present and the cues during the original learning, but it is also higher if an item was recently recalled. The greater the mismatch between the cues, the weaker retrieval strength will be. Items are thus more difficult to recall in a shifted context rather than in the context where the item was originally learned. However, the act of retrieval in a shifted context (where cues present during encoding are less available) leads to a greater increase in storage strength than retrieval in the acquisition context (where retrieval strength is high and the item is easily accessible). Smith and Handy (2014) argued that effortful recall after context changes increased storage strength for those items, effectively “decontextualizing” the information (see also Smith & Handy, 2016).

Perhaps effortful retrievals during training likewise enhanced transfer across contexts in Todd et al.'s (2012) instrumental learning experiments. The present research therefore aimed to test this possibility. In Experiment 1, operant responding was tested in a relatively novel context following acquisition in multiple contexts or in one context. In Experiment 2, operant responding was analogously tested in a relatively novel context following training that involved either long (14-day) or short (1-day) retention intervals inserted between training sessions. We hypothesized that any training procedure requiring more effortful retrievals might produce operant behavior that would transfer better to new situations.

## Experiment 1

Experiment 1 tested the hypothesis that training in multiple contexts will result in responding that transfers better to a novel context than equivalent training in one context, similar to Todd et al. (2012). However, the current experiment assessed responding following training in three contexts (rather than two) and used a different operant response, namely, a nose-poke response that required the rat to merely place its snout into a divot or depression in the chamber wall. Nose poking might depend less on idiosyncratic position and postural adjustments necessary to apply the right leverage to a lever, which is one possible explanation for the context-change effect in other responses (but see Bouton et al., 2014). Further, all animals were tested in both familiar and relatively novel contexts. Two groups of rats received six daily sessions in which nose poking produced a food-pellet reward. For Group Constant, these sessions occurred in the same context, Context A. But for Group Varied, the first two sessions were in Context A, the second two were in Context B, and the final two were in Context C. Responding was then tested (in extinction) in the most recently reinforced context and in a new context, D. We predicted reduced responding in Group Varied relative to Group Constant after each context switch in acquisition. But we also predicted better transfer of responding (indicated by higher levels of responding) to

Context D in Group Varied relative to animals that received training in one context.

## Method

### Subjects

The subjects were 16 naïve female Wistar rats purchased from Charles River Laboratories (St. Constance, Quebec), 75 to 90 days old at the start of the experiment. They were individually housed in suspended wire mesh cages and maintained on a 16:8-h light:dark cycle. Training sessions occurred during the light phase. Rats were food deprived to 80% of their initial weights throughout experimentation.

### Apparatus

Four sets of four operant conditioning chambers, housed in separate rooms of the laboratory, served as four contexts. The chambers were fully counterbalanced, each serving equally as Context A, B, C, or D. Each was in a sound-attenuation chamber and was of the same model (Med Associates model ENV-008-VP, St. Albans, VT). Boxes are described in detail in Trask and Bouton (2016). The nose-poke detector (2.5-cm diameter, 2-cm deep) was 6.3 cm above the floor directly to the right of the food magazine. Entries into the nose-poke detector broke an infrared beam and recorded a response.

Each set of boxes served as a distinct context. In one set, the floor was made of stainless steel grids (0.48-cm diameter) staggered so odd-numbered and even-numbered grids were mounted in two separate planes, one 0.5 cm above the other. A dish containing 5 ml of RiteAid lemon cleaner (RiteAid Corporation, Harrisburg, PA) was placed outside the chamber near the front wall. In the second set, one side wall and the ceiling had black diagonal stripes, 3.8 cm wide and 3.8 cm apart. The grids of the floor were mounted on the same plane and were spaced 1.6 cm apart. Five ml of Pine-Sol (Clorox Co., Oakland, CA) was placed in a dish outside the chamber. The third set had a 1.5-cm vertical gray stripe down the center of one side wall. The grids of the floor were spaced 1.6 cm apart. Five ml of white vinegar (PriceChopper Inc., Schenectady, NY) was placed in a dish outside the chamber. The fourth set had no adornment of the side walls. The floor grids had different diameters (0.5 cm and 1.3 cm, spaced 1.6 cm apart). One g of Vicks VapoRub (Procter & Gamble, Cincinnati, OH) was placed in a dish outside the chamber.

The reinforcer was a 45-mg grain-based rodent food pellet (5-TUM:181156; TestDiet, Richmond, IN, USA). Chambers were controlled by computer equipment in an adjacent room.

### Procedure

**Magazine training** Rats received 30-min sessions of magazine training in each context over the first 2 days of experimentation. On Day 1, half the animals in each group received sessions in Contexts A and D (counterbalanced so half were first in A, half in D) and half in Contexts B and C (counterbalanced so half were first in B, half in C). On Day 2, the rats received the opposite training. Daily sessions were separated by 1 hr. In each, approximately 60 reinforcers were delivered on a random time 30-s schedule. Nose-poke detectors were removed from the chambers.

**Acquisition** Over each of the next 6 days, rats received one session of instrumental training in which nose poking was reinforced on a variable interval 30-s schedule. Sessions began with the first nose poke and ended after 60 reinforcers had been earned (approximately 30 min). For Group Constant, all sessions were in Context A. For Group Varied, sessions were in Context A on the first 2 days, Context B on the next 2, and Context C on the final 2 days of acquisition. No hand shaping was necessary.

**Test** On the following day, all rats received two 5-min extinction tests. One test took place in Context D, where nose poking had not been trained in either group. The other test took place in the last context in which responding had been reinforced for each group (A for Group Constant, C for Group Varied). Testing order was counterbalanced, so half of the animals in each group were tested first in the familiar context and half were tested in the Context D. Food pellets were never presented.

**Data treatment** Data were subjected to analysis of variance (ANOVA), with a rejection criterion of  $p < .05$ . Partial eta squared was calculated as a measure of effect size on significant effects, and 95% confidence intervals (CIs) were calculated around the effect size.

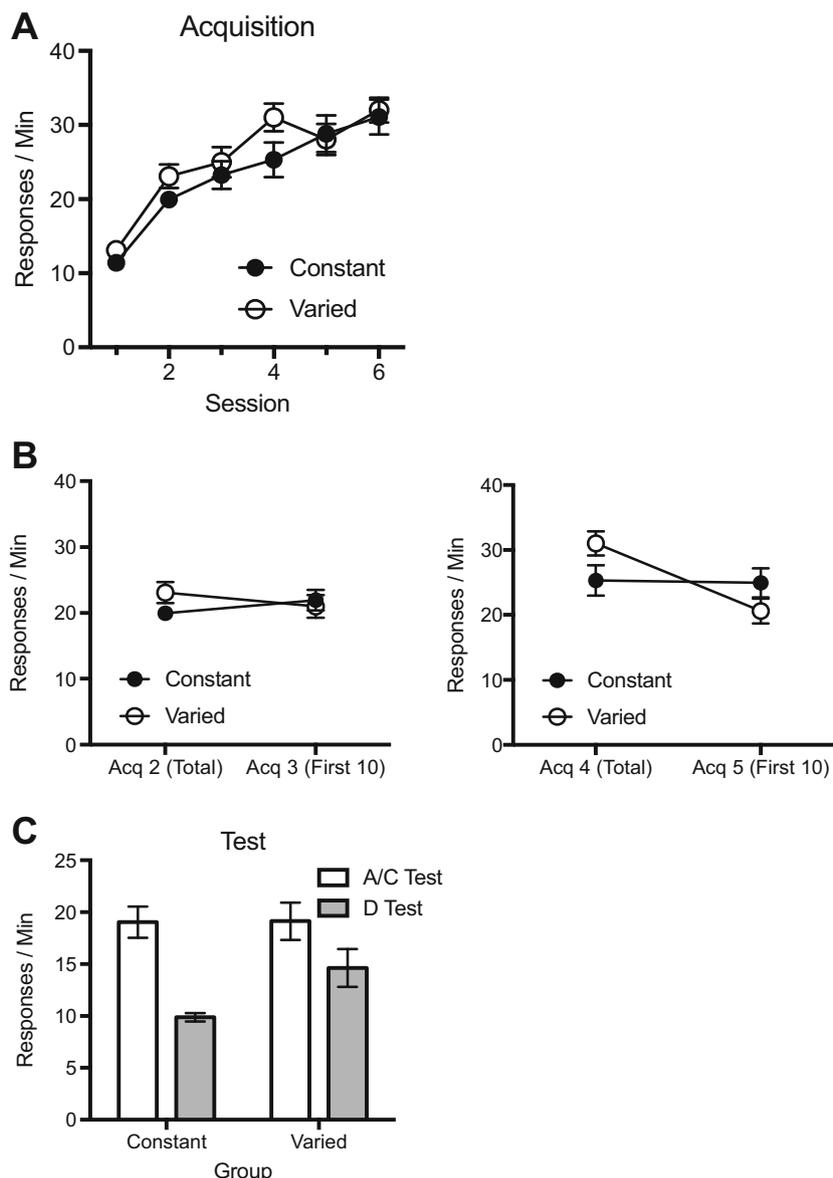
## Results and discussion

Throughout acquisition, animals in Group Constant gradually increased responding throughout the sessions. While this was also generally true of animals in Group Varied, on Day 3 (first training session after a switch to Context B) and Day 5 (first training session after a switch to Context C), animals showed a decrease in responding at the beginning of the session relative to the last session, even though responding was still reinforced. During the test, while both groups showed a significant context change effect when tested in D relative to their last reinforced context (A for Group Constant, C for Group Varied), the effect was more pronounced in Group Constant.

**Acquisition**

Mean responding in acquisition is summarized in Fig. 1a. A 2 (group) × 6 (session) ANOVA confirmed a main effect of session,  $F(5, 70) = 71.69$ ,  $MSE = 10.58$ ,  $p < .001$ ,  $\eta_p^2 = .84$ , 95% CI [.75, .87], and no main effect of group,  $F < 1$ , or interaction,  $F(5, 70) = 1.82$ ,  $p = .12$ . To isolate the immediate impact of a context change on responding in Group Varied, we focused on the first 10 minutes of responding after context changes relative to the preceding session (see Fig. 1b). The entire session mean was used to avoid any bias of higher rates toward the end of sessions. A 2 (group) × 2 (session) ANOVA assessed responding from Acquisition 2 to the first 10 minutes

of Acquisition 3 (when Group Varied changed from A to B). There was an interaction,  $F(1, 14) = 16.61$ ,  $MSE = 2.04$ ,  $p = .001$ ,  $\eta_p^2 = .54$ , 95% CI [.13, .73], but no effect of session or group,  $F_s < 1$ . Follow-up comparisons indicated that Group Varied decreased responding from Acquisition 2 to the first 10 minutes of Acquisition 3,  $F(1, 14) = 8.60$ ,  $p < .05$ ,  $\eta_p^2 = .38$ , 95% CI [.02, .62]. In contrast, Group Constant increased its responding,  $F(1, 14) = 8.01$ ,  $p < .05$ ,  $\eta_p^2 = .36$ , 95% CI [.02, .61]. A similar 2 (group) × 2 (session) ANOVA on responding from Acquisition 4 to the first 10 minutes of Acquisition 5 (see Fig. 1b) found an effect of session,  $F(1, 14) = 25.23$ ,  $MSE = 9.25$ ,  $p < .001$ ,  $\eta_p^2 = .64$ , 95% CI [.25, .79], an interaction,  $F(1, 14) = 21.83$ ,  $MSE = 9.25$ ,  $p < .001$ ,  $\eta_p^2 = .61$ , 95% CI



**Fig. 1** Results of Experiment 1. **a** Responding over sessions of acquisition. **b** Effects of the physical context changes during acquisition (group varied). Panels compare mean responding during each session that

preceded a context switch with mean responding during the first 10 min of the session that followed the switch. **c** Responding during testing. Error bars indicate standard error of the mean. Note changes in y axes

[.20, .77], but no effect of group,  $F < 1$ . Follow-up comparisons found that Group Varied showed a decrease in responding from Acquisition 4 to the first 10 minutes of Acquisition 5,  $F(1, 14) = 47.00, p < .001, \eta_p^2 = .77, 95\% \text{ CI } [.44, .86]$ . Group Constant did not,  $F < 1$ . Thus, context switches during training temporarily impaired nose-poke performance in Group Varied and are consistent with the idea that difficult retrievals were involved.

## Test

Test data are depicted in Fig. 1c. A 2 (group)  $\times$  2 (test: Context A/C vs. D) ANOVA found a significant main effect of testing context,  $F(1, 14) = 50.74, MSE = 7.35, p < .001, \eta_p^2 = .78, 95\% \text{ CI } [.46, .87]$ , but no main effect of group,  $F(1, 14) = 1.60, p = .23$ . The Group  $\times$  Session interaction, however, was significant,  $F(1, 14) = 5.89, MSE = 7.35, p < .05, \eta_p^2 = .30, 95\% \text{ CI } [.00, .56]$ . Follow-up comparisons revealed that both Group Varied,  $F(1, 14) = 11.03, p < .01, \eta_p^2 = .44, 95\% \text{ CI } [.06, .66]$ , and Group Constant,  $F(1, 14) = 45.94, p < .001, \eta_p^2 = .77, 95\% \text{ CI } [.43, .86]$ , showed reduced responding in Context D. While there was no group difference during the A/C test,  $F < 1$ , Group Varied responded significantly more than Group Constant in Context D,  $F(1, 14) = 6.17, p < .05, \eta_p^2 = .31, 95\% \text{ CI } [.00, .57]$ . When included as a factor in the test ANOVA, the identities of the four physical contexts used in the experiment did not interact with the other effects, largest  $F = 3.87, p > .05$ .

The results clearly suggest that training of the nose-poke response in multiple contexts temporarily hurt performance after each context switch during training but in the long run resulted in better transfer of responding to a new context than did training in a single context.

## Experiment 2

The passage of time and physical context change often have parallel effects on learning (Bouton, 1993). For example, latent inhibition (the finding in Pavlovian conditioning that a preexposed stimulus is more difficult to condition than a novel stimulus; e.g., Lubow & Moore, 1959) is attenuated by inserting either a change of context or a retention interval between the preexposure and conditioning phases (e.g., Rosas & Bouton, 1997, 1998). Similarly, extinction performance is lost following a context change (renewal; Bouton & Bolles, 1979; Bouton et al., 2011) or a retention interval (spontaneous recovery; Pavlov, 1927; Shaham, Adamson, Grocki, & Corrigan, 1997). Renewal and spontaneous recovery can both be reduced by presenting a retrieval cue from the extinction session (Brooks & Bouton, 1993, 1994), supporting the view that both phenomena result from failure to retrieve extinction.

Thus, there are grounds for thinking that inserting a retention interval between consecutive sessions of operant training—instead of changing the context, as in Experiment 1—should impair retrieval during training. If the effortful-retrieval-practice explanation of Experiment 1 is correct, then performance impaired by retention intervals, like context changes, should reduce the impact of a final context change on behavior. Experiment 2 tested this idea using methods that were otherwise identical to Experiment 1's. Two groups received six sessions of operant training before a test of responding in the training context and a new context. For Group Short-Interval, the acquisition sessions occurred on consecutive days, that is, with 24-hr intervals between them. For Group Long-Interval, a 14-day retention interval was inserted between Sessions 2 and 3 and Sessions 4 and 5 (the points in training where Group Varied had received context switches in Experiment 1). The long retention intervals were expected to make retrieval difficult, which should theoretically reduce the overall impact of the final context-change effect at test.

## Method

### Subjects

The subjects were 24 naïve female Wistar rats, maintained as in Experiment 1.

### Apparatus

Because the identity of the four boxes that provided contexts in Experiment 1 had no impact on performance (see above), only the lemon-scented and pine-scented boxes described above served as the two contexts. The same reinforcer was used.

**Magazine training** On the first day, all rats received two 30-min sessions of magazine training, one in each context, counterbalanced so that half of the animals were first in A and half in B. The sessions were separated by approximately 2 hrs. Events with the sessions were exactly as were those in Experiment 1. Nose-poke detectors were not present.

**Acquisition** Rats then received six sessions of nose-poke training, one a day, as in Experiment 1. Group Short-Interval was given all acquisition sessions on consecutive days. In Group Long-Interval, acquisition was conducted on consecutive days for Sessions 1 and 2, 3 and 4, and 5 and 6. However, a 14-day retention interval was inserted between Sessions 2 and 3 and acquisition Sessions 4 and 5. The start of acquisition was staggered between groups so that they finished acquisition on the same day. Thus, by the test, the two groups were the same age, had equivalent food deprivation, and had received

the same amount of handling overall. For all animals, acquisition sessions took place in Context A.

**Test** On the following day, rats were given two 5-min extinction tests, one in Context A and one in Context B (order counterbalanced).

## Results and discussion

Both groups gradually increased responding over the acquisition sessions. However, during Sessions 3 and 5 (the sessions that followed the 14-day retention intervals for Group Long-Interval), animals in Group Long-Interval, but not Group Short-Interval, showed a decrease in responding at the beginning of the session relative to responding during the previous one. During the test, the groups showed a similar decrease in responding when switched from Context A to Context B.

## Acquisition

Responding during acquisition is shown in Fig. 2a. A 2 (group)  $\times$  6 (session) ANOVA confirmed an effect of session,  $F(5, 110) = 42.44$ ,  $MSE = 12.55$ ,  $p < .001$ ,  $\eta_p^2 = .66$ , 95% CI [.54, .72], and an interaction,  $F(5, 110) = 2.46$ ,  $MSE = 12.55$ ,  $p < .05$ ,  $\eta_p^2 = .10$ , 95% CI [.00, .18], consistent with the impairment caused by retention intervals in Group Long-Interval. There was no effect of group,  $F < 1$ . As before, we focused on responding during the first 10 mins of the sessions that followed the long retention intervals (see Fig. 2b) in order to assess how a retention interval immediately impacted reinforced performance. A 2 (group)  $\times$  2 (session) ANOVA compared the mean responding from Acquisition 2 to the first 10 minutes of Acquisition 3. This revealed an interaction,  $F(1, 22) = 14.74$ ,  $MSE = 3.67$ ,  $p = .001$ ,  $\eta_p^2 = .40$ , 95% CI [.09, .61], but no effect of session,  $F(1, 22) = 2.64$ ,  $MSE = 3.67$ ,  $p = .11$ , or group,  $F < 1$ . Follow-up comparisons showed that Group Long-Interval decreased across this time,  $F(1, 22) = 14.94$ ,  $p = .001$ ,  $\eta_p^2 = .40$ , 95% CI [.09, .61]. Group Short-Interval did not,  $F(1, 22) = 2.45$ ,  $p > .05$ . A 2 (group)  $\times$  2 (session) ANOVA assessed responding from Acquisition 4 to the first 10 minutes of Acquisition 5 also found an effect of session,  $F(1, 22) = 22.42$ ,  $MSE = 5.67$ ,  $p < .001$ ,  $\eta_p^2 = .51$ , 95% CI [.18, .68], an interaction,  $F(1, 22) = 18.99$ ,  $MSE = 5.67$ ,  $p < .001$ ,  $\eta_p^2 = .46$ , 95% CI [.14, .65], and no effect of group,  $F < 1$ . Group Long-Interval again showed a decrease in responding,  $F(1, 22) = 41.34$ ,  $p < .001$ ,  $\eta_p^2 = .65$ , 95% CI [.36, .78]; Group Short-Interval did not,  $F < 1$ . These results indicate that a 14-day retention interval impaired performance when inserted during acquisition.

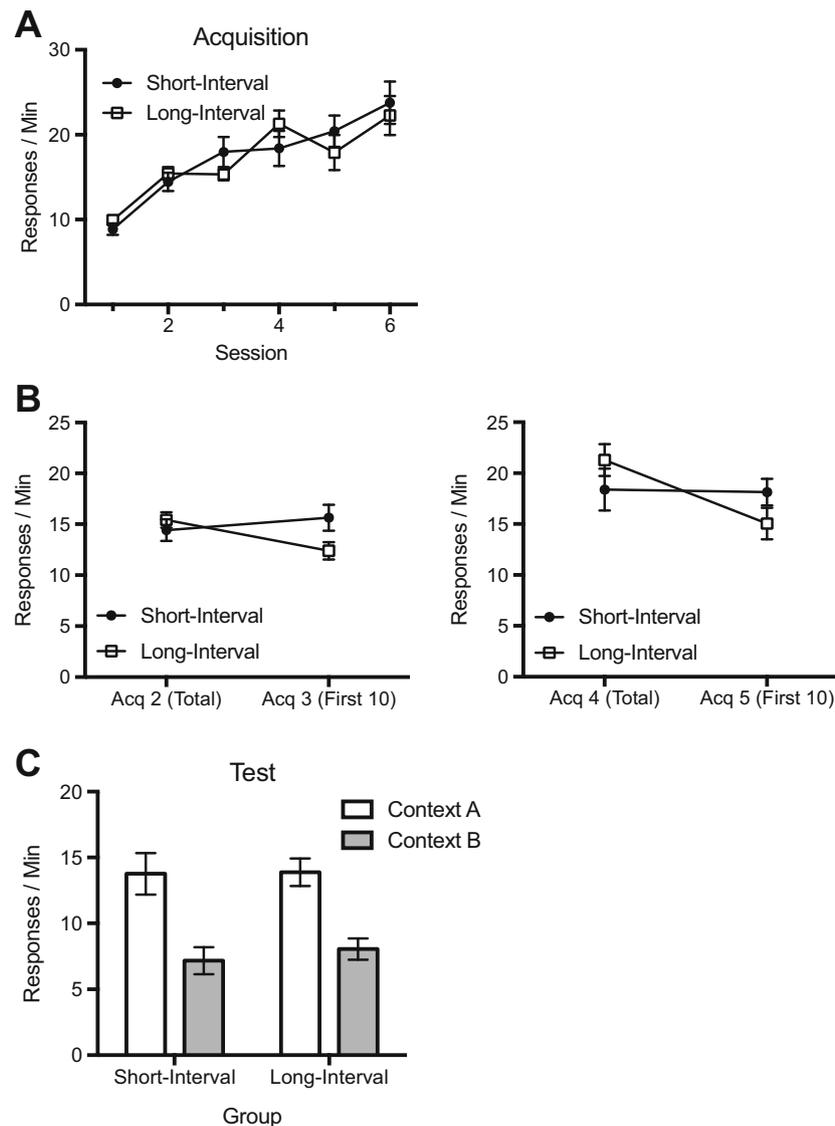
## Test

The test results are depicted in Fig. 2c. A 2 (group)  $\times$  2 (test: Context A vs. B) found a main effect of testing context,  $F(1, 22) = 56.62$ ,  $MSE = 8.19$ ,  $p < .001$ ,  $\eta_p^2 = .72$ , 95% CI [.46, .82], but no effect of group or interaction,  $F_s < 1$ . Follow-up comparisons showed both Group Long-Interval,  $F(1, 22) = 24.93$ ,  $p < .001$ ,  $\eta_p^2 = .53$ , 95% CI [.21, .70], and Group Short-Interval,  $F(1, 22) = 31.91$ ,  $p < .001$ ,  $\eta_p^2 = .59$ , 95% CI [.28, .74], responded more in A than in B. There were no differences between groups in either context,  $F_s < 1$ . Thus, training with longer retention intervals between sessions did not weaken the final context-change effect.

## General discussion

The goal of the present experiments was to separate two explanations of why operant training in multiple contexts can decrease the response-attenuating effect of a final context change: a view that emphasizes the positive effects of effortful retrieval practice (e.g., the new theory of disuse) and a view emphasizing generalization due to the conditioning of common elements. Both context change (Experiment 1) and long retention intervals (Experiment 2) inserted between sessions of operant training briefly impaired performance of the response. This result is consistent with the view that both manipulations made retrieval difficult. However, despite this similarity, only training in multiple contexts resulted in better transfer of the response to a new context; training across longer retention intervals had no impact on the context-change effect. Perhaps, contrary to the new theory of disuse as applied by Smith and Handy (2014), conditions that increase retrieval difficulty during learning do not necessarily improve retrieval of the response in new conditions. The present results also importantly expand the generality of the context specificity of operant behavior to the nose-poke response.

The effect of training in variable contexts demonstrated in Experiment 1 is analogous to results in Pavlovian fear extinction, which suggest that extinction in multiple contexts can enhance the generalization of extinction to new contexts (i.e., attenuate renewal) in both humans (Shiban, Pauli, & Mühlberger, 2013; Vansteenwegen et al., 2007) and rats (Gunther, Denniston, & Miller, 1998; Thomas, Vurbic, & Novak, 2009; but see Bouton, García-Gutiérrez, Zilski, & Moody, 2006). For example, in one experiment, Vansteenwegen et al. (2007) found that exposure to a spider video in individuals with arachnophobia in three contexts reduced renewal of fear (as measured by a skin conductance response) in a novel context relative to equivalent exposure to the spider video in a single context. Those results, like the present ones, suggest that learned performance transfers better to a new context if learning occurs in multiple contexts.



**Fig. 2** Results of Experiment 2. **a** Responding over sessions of acquisition. **b** Effects of the 14-day retention intervals inserted during acquisition (Group Long-Interval). Panels compare mean responding during each session that preceded the retention interval with mean

responding during the first 10 min of the session that followed the interval. **c** Responding during testing. Error bars indicate standard error of the mean. Note changes in y axes

Bjork and Bjork (1992) suggest that memory is dictated by two processes: storage strength and retrieval strength. The difference between these processes might be analogous to the distinction between learning and performance, where storage strength represents what is known and retrieval strength represents how easily that item is recalled during a test. Items can be high in storage strength but inaccessible due to low retrieval strength. Each time a memory is retrieved, its storage strength (and retrieval strength) increase. Retrieval strength, however, is dictated by how well cues during retrieval match those from acquisition and will be higher for recently recalled items. The larger the mismatch between these cues, the weaker the retrieval strength and more difficult retrieval will be. A memory that is retrieved despite a more effortful

retrieval receives a greater increase in storage strength than an easily retrieved memory. Smith and Handy (2014) suggested that when recall is more difficult during training, storage strength will increase. Increases in storage strength do not necessarily correspond with increases in retrieval strength. Instead, the act of effortful retrieval will result in an easier retrieval of that item in the future, creating less dependence on match between cues at encoding and retrieval. Without alterations, the theory inaccurately predicts that training across retention intervals results in performance that is more likely to transfer across contexts.

The results may be handled, however, by other theories that assume sampled contextual elements gain excitatory strength during acquisition (e.g., Estes, 1955b). Conditioning in

multiple contexts would increase the number of contextual elements that acquire excitatory strength, thereby increasing the chance that a novel context might contain shared elements that are already conditioned. Estes (1955a, 1955b) did not distinguish between elements arising from the environment versus the passage of time (the passage of time was assumed to correlate with fluctuation of physical cues such as temperature, humidity, and posture). However, it is not necessary to assume that elements corresponding to features of the physical context influence elements that might correspond to time. For example, some theories of timing have assumed the presence of a hypothetical cascading series of purely temporal stimuli that might be called temporal elements (Bouton & Hendrix, 2011; Desmond & Moore, 1988; see Kehoe, Horne, Macrae, & Horne, 1994, for one review). From this perspective, the results of Experiment 2 suggest that temporal elements and contextual elements might have little impact on one another despite their parallel effects on behavior. While context changes and retention intervals had similar effects on acquisition performance in Experiments 1 and 2, only context changes introduced during training reduced the negative impact of final testing in a new context.

A second explanation is that operant performance might rely on something analogous to transfer-appropriate processing in humans (Morris, Bransford, & Franks, 1977). That is, test performance is improved when cognitive operations used during encoding resemble those that are required during testing. The context switches during acquisition in Experiment 1 matched the final test context switch, but the retention intervals manipulated in Experiment 2 did not. A transfer-appropriate processing account predicts the current results and suggests that the retention interval manipulation of Experiment 2 would be effective if testing occurred following a retention interval.

Yet another explanation might not invoke memory processes per se. In the operant situation, physical context change might disrupt the animal's ability to perform a response that was specifically adapted to the environment in which it was learned. An animal trained in multiple contexts might contrastingly develop a form of responding that is less disrupted by context change. However, the easy nose-poke response (snout placement within a divot in the wall) studied here would seem to be less vulnerable to such a performance process than a response, like lever pressing, which might depend on more subtle aspects of the animal's position with respect to the lever and leverage. A long tradition in animal learning theory has also argued that context change affects in animal learning tasks mirror those studied in human memory (e.g., Millin & Riccio, 2004; Spear, 1978). That perspective is supported by studies that show retrieval cues can weaken or reduce the detrimental impact of a context change on performance. For example, Gordon, McCracken, Dess-Beech, and Mowrer (1981) demonstrated that the detrimental impact of a

context change on passive avoidance performance can be alleviated by providing a cue associated with acquisition (exposure to a "cueing box" identical to the start chamber in a passive-avoidance chamber). Further, the renewal effect can be reduced by presenting a cue associated with extinction at the time of testing (Brooks & Bouton, 1994). Thus, context change effects in animal learning experiments can often be attributed to memory retrieval deficits.

Like human memory, operant performance in rats can be context-dependent, and the present results further extend the parallel between what we know about memory and operant behavior. They also confirm that both memory and operant behavior can become less context dependent if training is given in multiple situations. However, in the present studies, this "decontextualization" of performance seemed limited to physical context changes during training. We have suggested that this result might imply that difficult retrievals are not sufficient to make an instrumental behavior context free.

**Author note** This research was supported by Grant RO1 DA 033123 from the National Institute on Drug Abuse to M.E.B. Send correspondence to either author, sydney.trask@uvm.edu or mark.bouton@uvm.edu. We thank Scott Schepers, Michael Steinfield, and Eric Thrailkill for their comments on the manuscript.

## References

- Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes*. Hillsdale, NJ: Erlbaum.
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, *114*, 80–99. <https://doi.org/10.1037/0033-2909.114.1.80>
- Bouton, M. E., & Bolles, R. C. (1979). Contextual control of the extinction of conditioned fear. *Learning and Motivation*, *10*, 445–466. [https://doi.org/10.1016/0023-9690\(79\)90057-2](https://doi.org/10.1016/0023-9690(79)90057-2)
- Bouton, M. E., García-Gutiérrez, A., Zilski, J., & Moody, E. W. (2006). Extinction in multiple contexts does not necessarily make extinction less vulnerable to relapse. *Behaviour Research and Therapy*, *44*, 983–994. <https://doi.org/10.1016/j.brat.2005.07.007>
- Bouton, M. E., & Hendrix, M. C. (2011). Intertrial interval as a contextual stimulus: Further analysis of a novel asymmetry in temporal discrimination learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*, 79–93. <https://doi.org/10.1037/a0021214>
- Bouton, M. E., & Todd, T. P. (2014). A fundamental role for context in instrumental learning and extinction. *Behavioural Processes*, *104*, 13–19. <https://doi.org/10.1016/j.beproc.2014.02.012>
- Bouton, M. E., Todd, T. P., & León, S. P. (2014). Contextual control of discriminated operant behavior. *Journal of Experimental Psychology: Animal Behavior Processes*, *40*, 92–105. <https://doi.org/10.1037/xan0000002>
- Bouton, M. E., Todd, T. P., Vurbic, D., & Winterbauer, N. E. (2011). Renewal after the extinction of free operant behavior. *Learning & Behavior*, *39*, 57–67. <https://doi.org/10.3758/s13420-011-0018-6>

- Brooks, D. C., & Bouton, M. E. (1993). A retrieval cue for extinction attenuates spontaneous recovery. *Journal of Experimental Psychology: Animal Behavior Processes*, *19*, 77–89. <https://doi.org/10.1037/0097-7403.19.1.77>
- Brooks, D. C., & Bouton, M. E. (1994). A retrieval cue attenuates response recovery (renewal) caused by a return to the conditioning context. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*, 366–379. <https://doi.org/10.1037/0097-7403.20.4.366>
- Desmond, J. E., & Moore, J. W. (1988). Adaptive timing in neural networks: The conditioned response. *Biological Cybernetics*, *58*, 405–415
- Dibbets, P., Maes, J. H., & Vossen, J. M. (2002). Contextual dependencies in a stimulus equivalence paradigm. *Quarterly Journal of Experimental Psychology: B*, *55*, 97–119. <https://doi.org/10.1080/02724990143000180>
- Estes, W. K. (1955a). Statistical theory of distributional phenomena in learning. *Psychological Review*, *62*, 369–377. <https://doi.org/10.1037/h0046888>
- Estes, W. K. (1955b). Statistical theory of spontaneous recovery and regression. *Psychological Review*, *62*, 145–154. <https://doi.org/10.1037/h0048509>
- Gómez, A. M., León, S. P., & Rosas, J. M. (2017). Roles of context in acquisition of human instrumental learning: Implications for the understanding of the mechanisms underlying context-switch effects. *Learning & Behavior*, *45*, 211–227. <https://doi.org/10.3758/s13420-016-0256-8>
- Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, *66*, 325–331. <https://doi.org/10.1111/j.2044-8295.1975.tb01468.x>
- Gunther, L. M., Denniston, J. C., & Miller, R. R. (1998). Conducting exposure treatment in multiple contexts can prevent relapse. *Behaviour Research and Therapy*, *36*, 75–91. [https://doi.org/10.1016/S0005-7967\(97\)10019-5](https://doi.org/10.1016/S0005-7967(97)10019-5)
- Kehoe, E. J., Horne, A. J., Home, P. S., & Macrae, M. (1994). Summation and configuration between and within sensory modalities in classical conditioning of the rabbit. *Learning & Behavior*, *22*, 19–26.
- Lubow, R. E., & Moore, A. U. (1959). Latent inhibition: the effect of nonreinforced pre-exposure to the conditional stimulus. *Journal of Comparative and Physiological Psychology*, *52*, 415.
- Millin, P. M., & Riccio, D. C. (2004). Is the context shift effect a case of retrieval failure? The effects of retrieval enhancing treatments on forgetting under altered stimulus conditions in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*, 325–334. <https://doi.org/10.1037/0097-7403.30.4.325>
- Pavlov, I. P. (1927). *Conditioned reflexes*. London, UK: Oxford University Press.
- Rosas, J. M., & Bouton, M. E. (1997). Additivity of the effects of retention interval and context change on latent inhibition: Toward resolution of the context forgetting paradox. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*, 283–294. <https://doi.org/10.1037/0097-7403.23.3.283>
- Rosas, J. M., & Bouton, M. E. (1998). Context change and retention interval can have additive, rather than interactive, effects after taste aversion extinction. *Psychonomic Bulletin & Review*, *5*, 79–83. <https://doi.org/10.3758/BF03209459>
- Shaham, Y., Adamson, L. K., Grocki, S., & Corrigan, W. A. (1997). Reinstatement and spontaneous recovery of nicotine seeking in rats. *Psychopharmacology*, *130*, 396–403. <https://doi.org/10.1007/s002130050379>
- Shiban, Y., Pauli, P., & Mühlberger, A. (2013). Effect of multiple context exposure on renewal in spider phobia. *Behaviour Research and Therapy*, *51*, 68–74. <https://doi.org/10.1016/j.brat.2012.10.007>
- Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory*, *5*, 460–471.
- Smith, S. M., Glenburg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, *6*, 342–353. <https://doi.org/10.3758/BF03197465>
- Smith, S. M., & Handy, J. D. (2014). Effects of varied and constant environmental contexts on acquisition and retention. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *40*, 1582–1593. <https://doi.org/10.1037/xlm0000019>
- Smith, S. M., & Handy, J. D. (2016). The crutch of context-dependency: Effects of contextual support and constancy on acquisition and retention. *Memory*, *24*, 1134–1141. <https://doi.org/10.1080/09658211.2015.1071852>
- Smith, S. M., & Vela, E. (2001). Environment context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review*, *8*, 203–220. <https://doi.org/10.3758/BF03196157>
- Spear, N. E. (1978). *The processing of memories: Forgetting and retention*. Hillsdale, NJ: Erlbaum.
- Sprague, J. R., & Horner, R. H. (1992). Covariation within functional response classes: Implications for treatment of severe problem behavior. *Journal of Applied Behavior Analysis*, *25*, 735–745. <https://doi.org/10.1901/jaba.1992.25-735>
- Thraillkill, E. A., & Bouton, M. E. (2015). Contextual control of instrumental actions and habits. *Journal of Experimental Psychology: Animal Learning and Cognition*, *41*, 69–80. <https://doi.org/10.1037/xan0000045>
- Thraillkill, E. A., Trott, J. M., Zerr, C. L., & Bouton, M. E. (2016). Contextual control of chained instrumental behaviors. *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*, 401–414. <https://doi.org/10.1037/xan0000112>
- Thomas, B. L., Vurbic, D., & Novak, C. (2009). Extensive extinction in multiple contexts eliminates the renewal of conditioned fear in rats. *Learning and Motivation*, *40*, 147–159. <https://doi.org/10.1016/j.lmot.2008.10.002>
- Todd, T. P., Winterbauer, N. E., & Bouton, M. E. (2012). Effects of the amount of acquisition and contextual generalization on the renewal of instrumental behavior after extinction. *Learning & Behavior*, *40*, 145–157. <https://doi.org/10.3758/s13420-011-0051-5>
- Trask, S., & Bouton, M. E. (2014). Contextual control of operant behavior: Evidence for hierarchical associations in instrumental learning. *Learning & Behavior*, *42*, 281–288. <https://doi.org/10.3758/s13420-014-0145-y>
- Trask, S., & Bouton, M. E. (2016). Discriminative properties of the reinforcer can be used to attenuate the renewal of extinguished operant behavior. *Learning & Behavior*, *44*, 151–161. <https://doi.org/10.3758/s13420-015-0195-9>
- Trask, S., Shipman, M. L., Green, J. T., & Bouton, M. E. (2017). Inactivation of the prelimbic cortex attenuates context-dependent operant responding. *The Journal of Neuroscience*, *37*(9), 2317–2324. <https://doi.org/10.1523/JNEUROSCI.3361-16.2017>
- Vansteenwegen, D., Vervliet, B., Iberico, C., Baeyens, F., Van den Bergh, O., & Hermans, D. (2007). The repeated confrontation with videotapes of spiders in multiple contexts attenuates renewal of fear in spider-anxious students. *Behaviour Research and Therapy*, *45*, 1169–1179. <https://doi.org/10.1016/j.brat.2006.08.023>