Received: 22 August 2022 | Accepted: 7 February 2023

DOI: 10.1002/jeab.838

RESEARCH ARTICLE

Journal of the Experimental Analysis of Behavior

SEAB SOCIETY FOR THE EXPERIMENTAL ANALYSIS OF BEHAVIOR

A laboratory model of canine search vigilance decrement, II: Noncontingent reward and Pavlovian appetitive stimuli

Mallory DeChant[1] | **Edgar O. Aviles Rosa**[1] | **Paola A. Prada-Tiedemann**[2] | **Nathaniel James Hall**[1]

[1]Department of Animal and Food Science, Texas Tech University, Lubbock, Texas, USA

[2]Department of Environmental Toxicology, Texas Tech University, Lubbock, Texas, USA

Correspondence
Nathaniel James Hall, Texas Tech University—Animal and Food Science, 1308 Indiana Avenue, Lubbock, Texas 79409, USA.
Email: nathaniel.j.hall@ttu.edu

Funding information
U.S. Department of Homeland Security Science and Technology Directorate (DHS S&T), Grant/Award Number: 70RSATT20CB0000010

Editor-in-Chief: Mark Galizio
Handling Editor: Mark Galizio

Abstract

Detection dogs have demonstrated reduced performance in operational settings when required to search in an environment where few to no target odors are present. This study's purpose was to increase detection dog accuracy using noncontingent reward (NCR) and Pavlovian stimuli associated with reward. Eighteen dogs were randomly spilt into two groups and received four 40-trial sessions in an operational and training context at 90% odor prevalence (baseline). Following baseline, in the operational context (now at 10% odor prevalence), experimental dogs received an NCR schedule consisting of delivering food rewards at the end of 66% of trials. After the NCR Test, dogs returned to baseline. During baseline, the experimental dogs received 10 days of delayed Pavlovian conditioning to a tone. During the test phase, the conditioned stimulus (tone) was presented to experimental dogs on average every two trials for 30 s in the operational context (now at 10% odor prevalence). Overall, NCR showed a nonsignificant trend for increased responding in the experimental group but tended to increase false alerts; therefore, a permutation of an NCR-like reward schedule may maintain search. The Pavlovian conditioned stimulus didn't decrease timeouts or improve accuracy, but a within-session analysis indicated that the dogs were more likely to time out and less likely to false alert when the tone was on than when it was off.

Detection dogs are required to perform long and difficult searches. Thus, it is critical that detection dogs remain attentive during the entirety of an operational task. However, most operational searches result in prolonged periods during which searching does not result in the discovery of the target odor (such as with explosives detection). Thus, in operational scenarios, searching may never yield a reinforcer, and as a result, search behaviors may undergo extinction, leading to a decline in performance (Aviles-Rosa et al., in press; Gazit et al., 2005; Porritt et al., 2015). For instance, Gazit et al. (2005) observed a decrement in search behavior when dogs repeatedly searched a path with no target odor. Similar results were also found by Porritt et al. (2015) where they observed a decrement in performance and search behavior after dogs searched an area with no target odors for 6 weeks. Both studies found that after dogs continuously searched an area with no target odor, they later had trouble finding a target odor when it was planted in the same area (Gazit et al., 2005; Porritt et al., 2015).

Recently, the authors developed a laboratory model to study this phenomenon (Aviles-Rosa et al., in press) We tested dogs in two different rooms (contexts). Each room was identical, but one room had high target odor frequency (90%), and the other, a low target odor frequency (10%). It was found that dogs had significantly lower performance and search-related behavior in the context with low target odor frequency (Aviles-Rosa et al., in press). Altogether, the available research suggests that a context with low target odor frequency reduces dogs' performance and search behavior (i.e., extinction of search behavior). These findings have significant implications for detection dogs. For instance, most explosive detection dogs continuously search an area (e.g., airports, cargo or passenger screening, etc.) where they rarely find a target odor. The infrequent appearance of a target odor in the operational context could result in a decrement in search behavior, and as a result dogs could potentially miss a target when it is present. Thus, the development of strategies that can help dogs maintain performance and search behavior in areas with low target odor frequency is of utmost importance.

Porritt et al. (2015) found that placing a noncontraband odor in the working environment was a good strategy to maintain search behavior and performance in operational scenarios. The authors demonstrated that planting a noncontraband odor (i.e., vanillin) in an area with no explosives significantly improved dogs' search behavior and performance (Porritt et al., 2015). Furthermore, dogs that constantly found a noncontraband target had better detection performance of an explosive when it was planted in the same area compared with dogs that searched the area and did not find a target for 6 weeks. Even when Porritt et al. (2015) found positive results in search behavior and performance by planting a noncontraband odor, this practice may not be feasible for all detection dog practices. For example, placing a noncontraband target in a populated area may cause unnecessary concern to the public. Furthermore, there may not be sufficient staff/logistics to place noncontraband target odors and some privately owned venues do not allow any targets (including noncontraband targets) to be placed, limiting what a handler may be able to do to mitigate search decline. Thus, there is still a need for additional methods that can help maintain search behavior longer in operational contexts with low target odor frequency.

One potential way to prevent extinction of search behavior and mitigate the performance decrement during low target prevalence searches may be to expose dogs to a noncontingent reward (NCR) schedule (reward presented independent of the presence or finding of a target odor). NCR is the delivery of rewards based on a fixed- or variable-time schedule and does not require a specific response by the participant (Marcus & Vollmer, 1996). Such a schedule for a detection dog could be the delivery of a toy or food according to a preprogrammed time schedule regardless of the dog finding a target or not. In pigeon and dog experimental models, there are conflicting results as to the increase in behavior such a schedule may yield (Lindblom & Jenkins, 1981; Pfaller-Sadovsky et al., 2019). Similar to NCR, when it is programmed in a positive reinforcement context, response reinstatement has demonstrated that extinction didn't result in erasure of the originally learned behavior association (Rescorla and Cunningham, 1977); however, response reinstatement is typically used in fear conditioning. For example, when reexposure to the unconditioned stimulus after extinction occurs, fear was reinstated even when the unconditioned stimulus and test were separated in time (Rescorla and Heth, 1975). Additionally, noncontingent reward may degrade the contingency for alerting on a target odor (reward is delivered in the absence of the required response), but behavioral momentum theory suggests that such a procedure could strengthen the persistence of behavior in the context in which the additional rewards are delivered through Pavlovian conditioning mechanisms (e.g. Lambert et al., 2016).

In a rodent model of detection dog behavior, Thrailkill et al. (2016) found that the use of NCR increased the number of responses in the first link of a two-link behavior chain that was created to model detection dog search when the first link (i.e., "search") underwent extinction. In this model, the researchers trained rats to engage in a behavior chain that involved lever pressing and a chain-pulling sequence (Thrailkill et al., 2016). First, a discriminative stimulus (e.g., panel light) signaled that the search response (e.g., chain pull) led to a second stimulus (e.g., insertion of lever) that set the occasion for a target response (e.g., lever press), which was reinforced by food (Thrailkill et al., 2016). Next, the chain pull did not lead to the presence of the lever (i.e., on extinction), and the rats that received the NCR schedule had greater resistance to extinction. The increase in behavior, however, was largely noted when the NCR was presented during training and extinction, thereby reducing discrimination between the two contexts. These findings occurred in a nonolfactory domain, but they nevertheless suggest that NCR may be a simple and viable way to prevent extinction or maintain dog search behavior in conditions under which the dog is unlikely to come across a target odorant.

Similar to the rat model, dog search behavior can be visualized as a behavior chain where the dog is signaled to search (first link), which results in the finding of a target odor. The target odor is a conditioned reinforcer for the search behavior and a discriminative stimulus that signals that the trained behavioral response (alert) will be reinforced (Thrailkill et al., 2016, 2018). Thus, as in the rat model, we hypothesize that an NCR schedule could maintain dog search behavior even in the absence of a target odor. Operationally, this could be as simple as a handler providing the dog a reward after every 5 min of search regardless of whether the dog successfully finds a target. Some canine handlers may already implement related procedures by giving the dog a "toy break," but it is unclear what effect this may have for canine performance.

Prior studies in rodents have also found that Pavlovian conditioned stimuli enhance motivation and increase behavioral responses during extinction (Cartoni et al., 2016; Holmes, et al., 2010). This phenomenon is called Pavlovian instrumental transfer (PIT), and it occurs when a Pavlovian conditioned stimulus associated with a reward increases an instrumental response for a different or the same reward under extinction (Cartoni et al., 2013). For example, rats are trained to associate a tone (conditioned stimulus) with a food reward such that when the tone is presented, food is delivered. Next, the rats are trained to press a lever (instrumental response) to receive a food reward. Then, the lever response is placed on extinction with or without the appearance of the previously conditioned tone. Results in rodents indicate that more lever presses are made in extinction in the presence of the tone than in the absence of the tone (Corbit & Balleine, 2011). These previous studies in rodents demonstrated an increase in behavioral responses during extinction with Pavlovian instrumental transfer, suggesting that PIT could be used as a mitigation strategy to maintain detection dog search behavior and performance in low target odor prevalence contexts.

The effects of NCR and PIT have been extensively studied in humans and rats (Holland, 2004; Meindl et al., 2021; Prevost et al., 2012; Thrailkill et al., 2016), and the literature suggests that both methods could be used to mitigate canine search decrement in contexts with low prevalence of target odors and that both methods could be easily implemented by detection dog handlers. As noted above, the authors recently developed a laboratory model of canine search behavior and demonstrated that dogs show a considerable decrement in search performance when the frequency of the target odor is reduced substantially below the level used in training (see Aviles-Rosa et al., in press, part I). Therefore, the purpose of these studies was to leverage this laboratory model to evaluate whether the use of NCR and PIT could mitigate canine performance decrement in a context with a low prevalence of odor targets.

EXPERIMENT 1: THE USE OF NONCONTINGENT REWARDS TO INCREASE DETECTION PERFORMANCE

Materials and methods

Animals

Eighteen mixed-breed dogs were used for these two experiments. These were also the participants of a previous study where a laboratory model was developed to study canine search decrement (Aviles-Rosa et al., in press). Dogs were tested in three independent cohorts of six dogs each. The dogs were sourced from local shelters and rescue organizations. The dogs were housed at the Texas Tech University (TTU) Canine Olfaction Lab, and they participated in a training program to increase adoptability. The dogs' backgrounds were unknown, but all were presumable naïve to detection training. The dogs also received two daily walks and training for adoption. All procedures used in both experiments were reviewed and approved by TTU Institutional Animal Care and Use Committee (ACUC #19093–10).

Apparatus

The dogs were trained to operate the three-alternative-choice automated olfactometer described by Aviles-Rosa et al. (2021). Briefly, each olfactometer was equipped with infrared sensors that measured the duration of the dog's nose in each port. Each olfactometer was connected to a separate odor port and was controlled by a computer for a completely automated device. The air flow of the odor lines of the olfactometers was set at 1 L/m and the continuous clean air line was at 2 L/m (1:3 odor dilution). This was identical for all the odors. To evaluate the effect of NCR on dog search behavior and performance, we used the laboratory model of canine search decrement described in the first part of this series (Aviles-Rosa et al., in press). Briefly, dogs were trained to operate the olfactometer in two different rooms or "contexts." These rooms were adjacent (shared a common wall) and were identical in size but were mirror images from a birds-eye view such that Room 1 had a door in the top right corner, with training equipment in the bottom left, whereas Room 2 had a door in the top left corner, with training equipment in the bottom right of the room. These rooms are referred to as the "training" and "operational" context. These two contexts model how dogs are frequently trained in one area but work and operate in another area. Cohorts 1 and 3 had the same room assignment for operational and training contexts; however, Cohort 2 had the opposite room assignment for the contexts (i.e., the training context for Cohort 2 was the operational context for Cohorts 1 and 3).

The dogs were trained to detect and alert to the odor of double-based smokeless powder (SP; Hodgdon H335) and ignore distractors (cotton gauze (Equate), nitril gloves (Med Pride, MPR-50504), blank/empty jar, limonene (10–3 v/v dilution in mineral oil; CAS # 5989-54-8), and food-grade mineral oil (Bluewater Chemgroup) in both contexts. SP is a common energetic that explosives detection dogs are trained to detect.

For each trial, either zero or one of the three odor ports presented the target odor of SP (the frequency of target odor presentation was explicitly manipulated and detailed below). The location of which port presented the target was pseudorandomized so that each port contained the target odor as equally as possible. Odor ports that were not programmed to present a target odor (i.e., two or three ports) presented one of the distractor odors (gauze, gloves, blank vial, lemon, or mineral oil) chosen at random from a uniform distribution of equal probability. The selection of distractors for each distractor odor port was independent; thus, it was possible for multiple ports to present the same distractor.

A correct alert consisted of holding their nose in the port containing SP for four consecutive seconds and was reinforced with a food reward. A false alert was defined as alerting to a port containing a distractor odor (incorrect response). During a blank trial (i.e., no target odor and all ports presenting a distractor), the dog had to search all three ports and not alert (remove the nose for four consecutive seconds) for the trial to be counted as correct ("all clear"); however, all-clear responses were not reinforced with a food reward. If the dog did not search all three ports or did not alert to any port within the 45 s, the trial was terminated and it was scored as incorrect (timeout). Half of the dogs ($n = 9$) were randomly assigned to receive the experimental treatments, and the other half served as control and did not receive treatment. Participant dogs had previously demonstrated a performance/search decrement when challenged with reduced target odor frequency in the operational context (Aviles-Rosa et al., in press). The dogs were assigned to the experimental/control groups before the beginning of Experiment 1 and independent of their performance decrement in our previous study. Dogs selected as the experimental group in Experiment 1 also served as the experimental group in Experiment 2. This was to ensure that the control group was never exposed to a prior mitigation strategy. For each cohort of six dogs (three cohorts total), three ($n = 3$) were assigned as control and three ($n = 3$) to experimental treatments.

Experimental design

The dogs started in a baseline period that consisted of four 40-trial sessions in both the operational and training context at a 90% odor prevalence rate (e.g., 10% of the trials were blank trials with no target odor). These four baseline sessions were the same 4 days of recovery from the immediately preceding study (Aviles-Rosa et al., in press). Each day, the dogs completed one session per context (operational and training). The order of contexts within a day (i.e., operational first or second) alternated daily throughout the duration of the study. After the baseline period, the dogs progressed to the test period that consisted of five 40-trial sessions in both the operational and training context. During the test period, the odor prevalence rate was 90% in the training context and was reduced to 10% (e.g., four out of 40 trials had target odor) in the operational context.

For the first cohort, the experimental dogs received a noncontingent reward schedule that consisted of delivering a "free" food reward delivered at the start of the intertrial interval and was independent of the dog's response during a trial. This schedule was selected to mimic an operational scenario in which it would be unknown whether a dog's response was correct but a handler may provide a reward to dogs on more frequent schedules whether the dog alerts or not. To prevent an increase in false alerts induced by the treatment in the first cohort ($n = 3$ experimental dogs), a 5-s delay was added after the end of the trial to deliver the noncontingent reward for the second and third cohort of experimental dogs.

Experimental dogs received a noncontingent reward schedule that consisted of delivering a "free" food reward at the end of 66% of the trials (e.g., food delivery was determined for the end of each trial by a random selection from a distribution weighted to deliver a reward 66% of trials; actual reward delivery was 63% of trials and led to reward delivery approximately every 1–2 min on average) within a session in the operational context. A high frequency of reward delivery was selected to maximize detection of a potential effect and maintain reward rates close to those observed in the training context. Additional NCR was not delivered in the training context given the already high rates of reward deliveries and to better mimic a situation where handlers may implement NCR specifically when on a long search. The control dogs ($n = 9$) did not receive the NCR schedule and were rewarded only when they alerted correctly to the target odor.

After completion of the test period, the dogs proceeded to a 10-trial control test followed by the recovery period that consisted of four 40-trial sessions in both the operational and training contexts. During recovery, the odor prevalence rate was 90% in both the training and operational contexts. This recovery period served to increase the dog's performance in preparation for Experiment 2. Additionally, dogs began Pavlovian conditioning one session per day in preparation for Experiment 2 (see Experiment 2 for details).

The control test was a single 10-trial session, conducted in the training room in which all odors from the olfactometer were unplugged, but otherwise run identically to a training session. The olfactometer program activated valves, recorded behavior, and reinforced responses identically to a training session at 90% target odor prevalence. This was done to verify that the dogs did not learn to identify the correct port via incidental cues unrelated to the target odor. This could occur, for example, if the dogs learned a distinctive "click" from the olfactometer solenoid valves that would indicate which port contained target odor or some other unrealized unintentional cue. Thus, if the dogs could use a cue *other than the target odor*, we would expect performance to remain high, indicating that they could identify the correct port in the absence of the target odor. However, it was expected that the dogs were only using olfactory cues and would therefore be unable to identify the target port and would respond by making either all-clear responses or false alerts at a rate that would not exceed chance performance. For analysis purposes, a "correct" response was defined as a dog identifying the olfactometer with the target odor valve activated (the programmed correct port), although no actual target odor was being delivered.

Statistical analysis

The data were analyzed for the baseline period to verify no or limited a priori differences in performance between groups. The primary dependent variable for baseline period performance was accuracy (correct vs. incorrect; binomial). We then analyzed differences between groups (experimental vs. control) and contexts (operational vs. training) and session number during the test period where odor prevalence was decreased to 10% in the operational context and an NCR intervention was provided to experimental dogs in that context. The primary dependent variables were accuracy (correct vs. incorrect; binomial), log-transformed latency (time from start of trial to first nose entry; continuous), hit rate (hit or miss when odor present, binomial), timeout probability (searched all

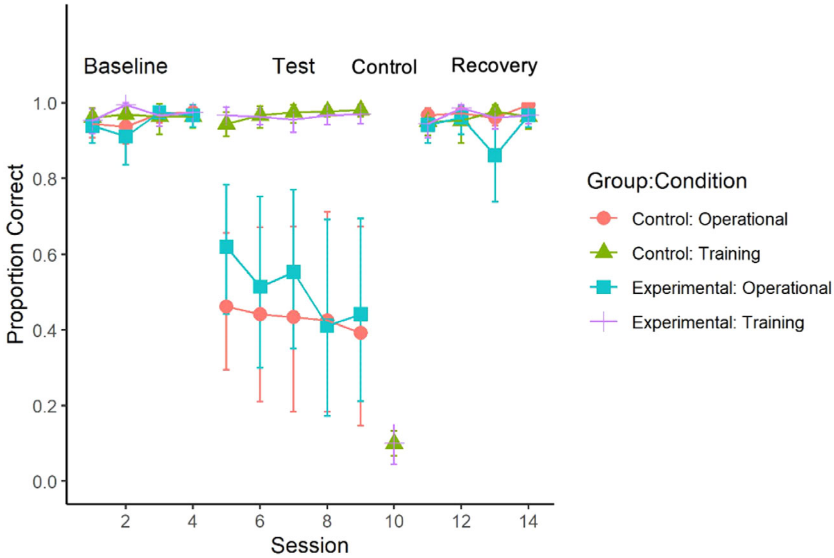

FIGURE 1 Mean proportion of correct responses by group for during the different phases of Experiment 1. The error bars represent 95% confidence interval of the proportion of trials that resulted in a correct response during Experiment 1 (i.e., a correct hit or correct all clear). Both the control and experimental groups showed similar results. Session 10 is the odor control, demonstrating that dogs cannot identify the correct odor port using unintentional cues in the absence of the target odor.

three ports or ended in timeout, binomial), and false-alert rate (trial ended in a false alert or not, binomial). Binomial variables were analyzed with logistic mixed-effect models, and continuous variables were analyzed with linear mixed-effect models.

All dependent variables were analyzed with the same model formula in which the dependent variable was predicted by the fixed effects of group (experimental vs. control), context (training or operational), the group by context interaction, and session. A random effect of dog ID (random intercept model) was included for all models. Statistical significance of fixed effects was evaluated using the analysis-of-variance function in the car package in R (Fox & Weisberg, 2019). Significant interactions were further analyzed with Tukey-adjusted post hoc tests from the lsmeans package (Russell, 2016) to assess for differences between groups for each context. Models were fit with the lmer package (Kuznetsova et al., 2017) in R (R version 3.5.1, www.r.project.org; R Core Team, 2018).

Results

Figure 1 shows mean and 95% confidence intervals of dog performance by group (experimental and control) and context for baseline, the test phase, the control test, and the recovery phase for Experiment 1. There appears to be a potential for minor elevation in performance in the operational context for the experimental dogs for the first test session (Session 5), but then performances become highly similar between groups. Figure 2 shows the same data at the individual participant level. This highlights that a few individual dogs (Charles, Jax, Sasha, Buster, Edna) showed minor disruption (irrespective of experimental group), whereas most dogs showed substantial disruption in performance. Table 1 shows the mean accuracy, false alerts, timeout rate, hit rate, correct-rejection rate, and false all-clear rate by group, context, and period for Experiment 1.

Baseline

During the baseline period, where the target prevalence was identical in the operational and training contexts, there was no significant interaction between group and context ($\chi^2 = 2.55$, $p = .11$) and no difference between groups ($\chi^2 = 0.07$, $p = .93$) on detection accuracy. However, both groups of dogs did show a significantly lower accuracy in the operational context ($\chi^2 = 9.63$, $p < .01$) and showed improvement across sessions ($\chi^2 = 7.96$, $p < .01$). This likely reflects the recovery from the previous experiment in the operational context because baseline started immediately after the low-prevalence challenge in part I (Aviles-Rosa et al., in press).

Importantly, however, both groups in both contexts were performing highly similarly in the last two sessions prior to the initiation of the decreased target frequency in the operational context and the initiation of the NCR treatment for the experimental dogs. When the logistic mixed-effects model was conducted over the last two baseline sessions (Sessions 3 and 4), there was no effect of group, context, or their interaction on accuracy (all $p > .40$), which is further shown by the mean accuracies (experimental dogs: operational 97.1% vs. training

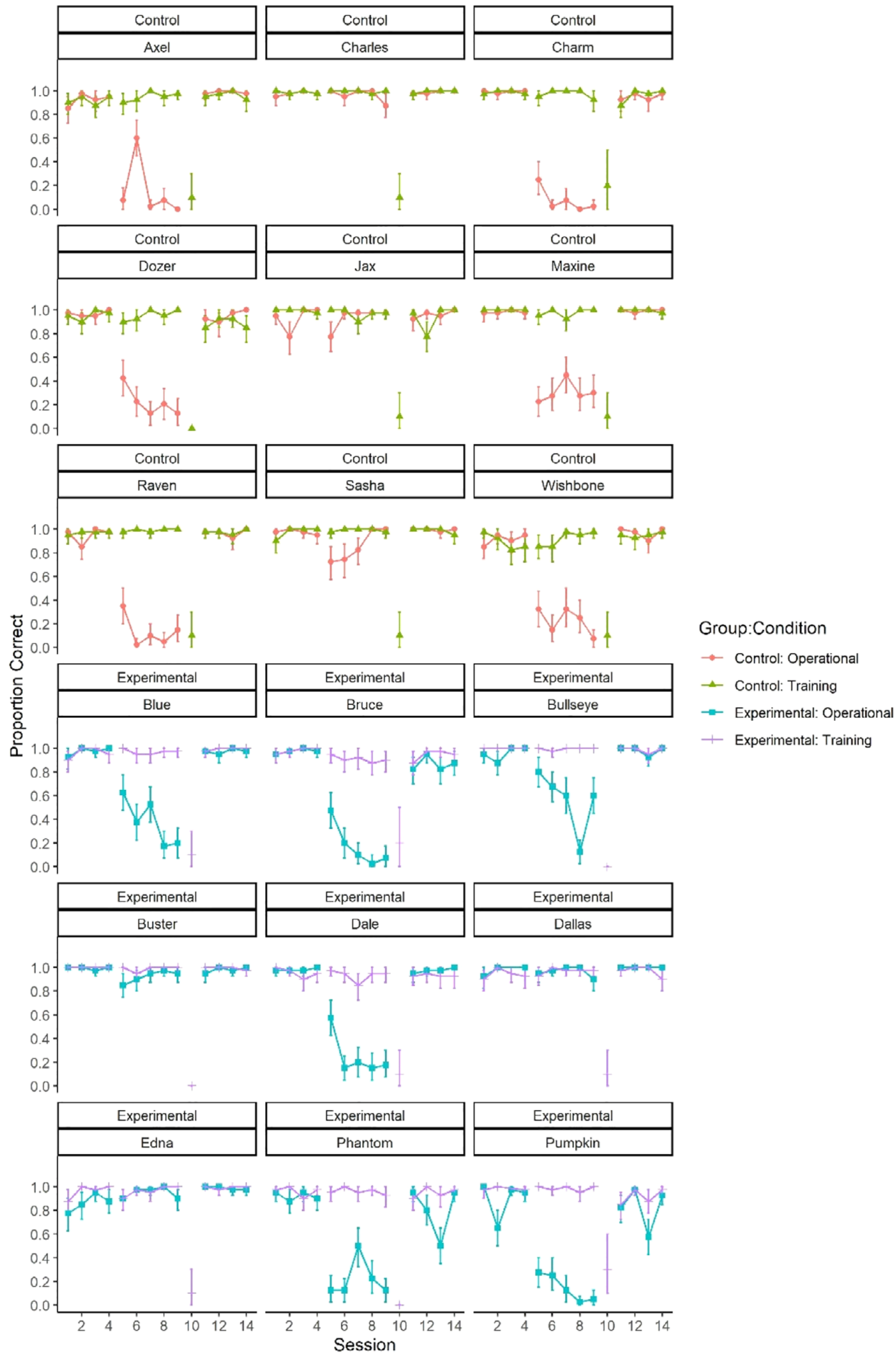

FIGURE 2 Individual dogs' mean proportion correct for each session, graphed similarly to Figure 1. The phases are as follows: baseline, test, odor control, and recovery. Most dogs show a substantial performance disruption in the operational context during the test phase. Similar disruption is seen for experimental and control dogs. Session 10 is the odor control, demonstrating that dogs cannot identify the correct odor port using unintentional cues in the absence of the target odor.

97.1%; control dogs: operational 97.3% vs. training 96.3%). Further analysis of the last two sessions for the dependent variables of false alerts, false all clears, and timeouts indicates that there was no difference in timeout rates or false all clears between group, context, or their interaction (all $p > .10$). There was, however, a minor increase in false alerts in the training context over the operational context ($\chi^2 = 5.15$, $p = .02$) but no effect of group ($\chi^2 = 0.15$, $p = .90$) or a context by group interaction ($\chi^2 = 0.28$, $p = .60$). False-alert rates for the training

TABLE 1 Mean ± standard deviation of the control and experimental group overall performance during the baseline, testing, and recovery phase during Experiment 1

Period	Baseline				Testing				Recovery			
Context	Training		Operational		Training		Operational		Training		Operational	
Group	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental
Accuracy, %	96.45 ± 3.61	97.24 ± 1.96	95.65 ± 2.85	94.79 ± 4.55	96.89 ± 2.39	96.44 ± 2.72	43.06 ± 37.33	50.77 ± 35.16	96.11 ± 3.37	96.25 ± 3.35	97.36 ± 1.76	93.26 ± 7.83
False, %	1.80 ± 3.26	0.97 ± 1.65	1.10 ± 1.81	1.29 ± 1.95	1.66 ± 1.78	2.61 ± 2.49	8.06 ± 5.59	26.00 ± 31.24	0.69 ± 1.10	1.20 ± 2.07	0.55 ± 0.85	2.2 ± 4.64
Timeout, %	0.69 ± 0.85	0.20 ± 0.44	1.65 ± 2.03	1.94 ± 3.35	0.88 ± 0.99	0.27 ± 0.26	48.75 ± 35.20	23.22 ± 26.17	1.59 ± 2.52	1.85 ± 3.11	1.31 ± 1.18	3.26 ± 5.91
Hit, %	97.22 ± 2.82	97.55 ± 2.00	96.01 ± 2.17	94.75 ± 4.37	97.34 ± 2.18	98.39 ± 1.09	12.22 ± 20.32	14.44 ± 19.59	96.62 ± 3.72	97.32 ± 2.60	97.63 ± 1.74	92.55 ± 8.24
Correct rejection, %	89.58 ± 16.23	94.44 ± 12.28	92.40 ± 14.24	95.13 ± 6.83	92.77 ± 7.12	78.88 ± 24.72	41.67 ± 37.93	49.01 ± 36.10	90.27 ± 11.73	85.18 ± 18.20	88.88 ± 8.71	84.02 ± 27.26
False all clear, %	1.15 ± 1.25	1.75 ± 1.51	1.78 ± 1.32	2.20 ± 2.47	0.62 ± 0.43	0.74 ± 0.78	3.43 ± 6.85	0 ± 0	1.79 ± 2.11	0.77 ± 1.12	0.85 ± 1.04	1.40 ± 1.28

Note. Each session was 40 trials. The target odor prevalence was 90% in both contexts during baseline and recovery, 90% in training context for testing, and 10% in operational context in testing. Accuracy: percentage of trials that a dog made a correct response by either alerting to the target odor or calling a correct all clear. False: percentage of trials that a dog alerted to an incorrect port during a target odor or blank trial. Timeout: percentage of trials the dog "timed out" without searching all three odor ports or alerting to a port. Hit: percentage of odor present trials that a dog alerted to the target odor. Correct rejection: percentage of blank (odor absent) trials dogs did a correct all clear. False all clear: percentage of trials dogs made an incorrect all-clear response when target odor was present.

context occurred on approximately 2% of trials, whereas false alerts occurred on approximately 1% of trials in the operational context, highlighting the minor difference in false-alert rates.

Accuracy

During the testing period, there was a significant decline in performance across the testing sessions ($\chi^2 = 28.28$, $p < .001$). Additionally, the interaction between group and context was significant ($\chi^2 = 7.58$, $p = .005$). Tukey-adjusted post hoc tests to analyze the interaction and test group differences in the two contexts indicated that for the operational context there was no difference between the control and experimental group (est = 0.41, $z = 0.48$, $p = .63$), highlighting a minimal effect of NCR on accuracy. Furthermore, there was no difference between the experimental and control group in the training context (est = 0.19, $z = 0.22$, $p = .82$). Nonetheless, Figure 1 indicates a potential elevation in accuracy in Session 1. To further evaluate a potential effect in Session 1, the accuracy rates for the experimental and control group in the operational context of Session 1 were compared. This comparison further indicated there was no main effect of NCR (i.e., group) on accuracy ($\chi^2 = 0.90$, $p = .34$).

Latency

During the test period, log-transformed latency showed increases across sessions ($\chi^2 = 30.78$, $p < .001$) and a group by context interaction ($\chi^2 = 146.06$, $p < .001$). Post hoc tests to break down the interaction indicated that there was no difference in latency between the control and experimental groups in the training context (est = 0.02, $z = 0.17$, $p = .86$), but there was a trend in that experimental dogs responded faster in the operational context than control dogs (est = 0.20, $z = 1.72$, $p = .08$) did, suggesting a potential improvement in approach behavior due to NCR.

Timeout probability

During the test period, there was no significant interaction of group and context ($p > .80$), but there was a main effect of session ($\chi^2 = 72.18$, $p < .001$) and context ($\chi^2 = 560$, $p < .001$), indicating that timeout probability increased across sessions more generally and that timeout probability was lower for sessions in the training context compared with sessions in the operational context. There was a trend-level effect for group ($\chi^2 = 2.74$, $p = .09$), indicating that timeout probability was lower in the experimental dogs than in control dogs, suggesting that NCR may have a trend effect on reducing timeouts in the absence of an effect on accuracy.

Hits

During the test period, there was no significant effect of session, group, or group by context interaction ($p > .30$). There was, however, a main effect of context where the hit rate was overall lower in the operational context than the training context ($\chi^2 = 423$, $p < .001$), but this did not differ between experimental and control dogs, indicating no effect of the NCR treatment.

False alerts

During the test period, there was a significant effect of session on the false-alert rate ($\chi^2 = 7.42$, $p < .007$), indicating that false alerts declined across sessions. In addition, there was a significant interaction between group and context ($\chi^2 = 33.48$, $p < .001$). Post hoc tests to analyze this interaction indicated that for the training context there was minimal difference in false-alarm rate between experimental and control dogs (est $= 0.45$, $z = 0.68$, $p = .50$). However, there was a trend for increased false alerts in the experimental dogs compared with control dogs in the operational context (est $= 1.10$, $z = 1.73$, $p = .08$). The control group had an 8.06% false-alert rate, whereas the experimental group had a 26% false-alert rate (Table 1). This suggests that the NCR treatment could have had a minor effect in increasing the false-alarm rate for experimental dogs. Due to the added 5-s delay for NCR delivery for Cohort 2 and 3, Cohort 1 was analyzed separately to see if the false-alert rate varied. For Cohort 1, the control group had an 8.16% false-alert rate, whereas the Experimental group had 49.66% false-alert rate during the testing period. For Cohorts 2 and 3 combined, the control group had an 8.06% false-alert rate, whereas the experimental group had an 18.93% false-alert rate during the testing period. The separation of cohorts with the 5-s delay of NCR reward delivery improved the false-alert rate for the experimental group in the testing period, but it still was higher than the control group.

False all clears

False all clears were defined as the dog "calling" an all clear when there was a target odor present. During the test period in the operational context, the control group had a 3.49% false-all-clear rate, whereas the experimental group had a 0% false all clear during the testing period. These data represent that when the experimental dogs were searching, they did not call an all clear unless there was no target odor present.

Error types as a proportion of errors

Figure 3 shows each error type (false alerts, timeouts, false all clears) as a proportion of the total errors made. In the training context, the type of error (as a proportion of overall errors) was relatively similar across baseline, test, and recovery for both experimental and control dogs. For the operational context, however, the test period showed a marked increase in timeouts as a proportion of errors for control dogs, whereas the experimental dogs had similar proportions of timeouts and false alerts. This supports the results from the raw numbers of timeouts and false alerts.

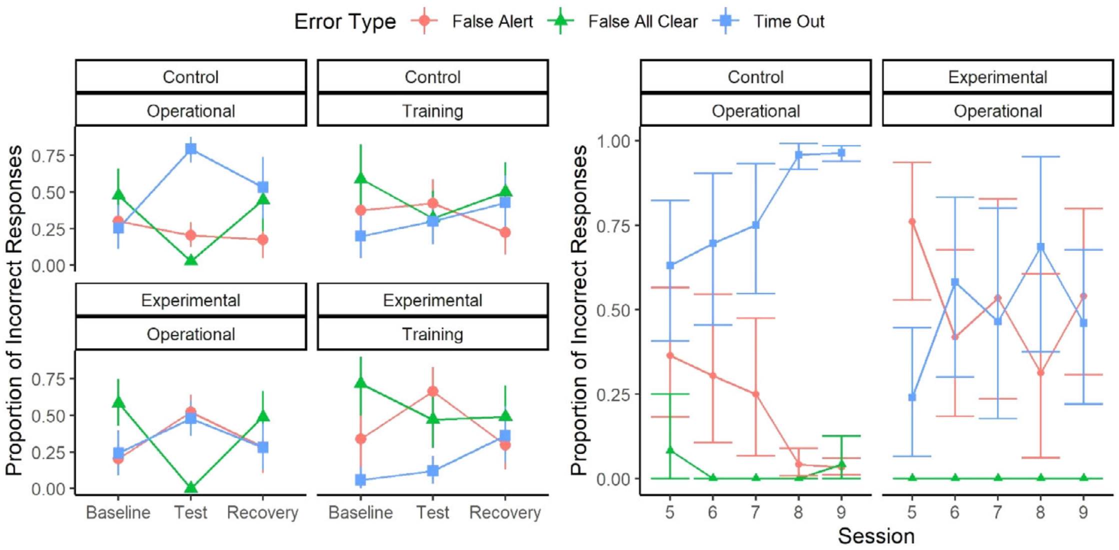

FIGURE 3 Error type frequency expressed as a proportion of errors made. Left: shows the false alerts, false all clears, and timeouts averaged across baseline, test, and recovery for each context and group. Right: shows the changes across session for the context in which errors were most frequent (i.e., the operational context) where approximately half of trials ended in an error. The points show the mean, and the error bars show the 95% confidence intervals.

False all clears declined to almost zero, indicating that very few misses occurred when the dogs examined all three ports. We did not perform further null-hypothesis significance testing due to the infrequent total number of errors during baseline and recovery periods as well as some error types almost never occurring in some conditions (i.e., false all clears).

Control test

The average accuracy of the dogs during the control test (one 10-trial session in which all target odors from the olfactometers were unplugged) was 10% and was well below chance. The 10% accuracy represents that on average one correct all-clear response occurred for the blank trial (1 out of 10 trials at 90% odor prevalence). This low accuracy shows that the dogs could not identify the olfactometer-programmed correct port above chance, even when the olfactometer valves were activated, if target odor is not presented. This result confirms that the dogs' performance was driven by the presence of the target and not unintentional cues from the olfactometer.

Discussion

The results from Experiment 1 demonstrate that NCR did not improve detection accuracy in the operational context. Although this experiment did not demonstrate an increase in detection accuracy with NCR, other publications have demonstrated an increase in behavior in other species. Thrailkill et al. (2016) concluded that delivery of NCR increased search persistence in rats provided that NCR had also been presented during training. Thrailkill et al. (2016) demonstrated that NCR during training contributed to the behavioral persistence under extinction because of the continued presentation of the reward stimulus and the association with learning the behavioral chain. The present experiment had a more complex search behavior chain than Thrailkill et al.'s (2016), and the NCR was only presented to the dogs in the operational context, making it easier to discriminate between high and low reward-frequency conditions. This procedure was chosen to better mimic how a dog may be trained and then deployed but does leave open the question as to whether NCR used during the training context as well may help with NCR treatment during the operational context. Future studies using NCR treatment in the training context during high (i.e., 90%) target odor prevalence may demonstrate a positive response that is more similar to that observed by Thrailkill et al. (2016). From an operational perspective, however, NCR may not be feasible in the field if the NCR may cause an increase in false alerts due to random NCR delivery, similar to the increase in false-alert rate in Cohort 1 during the operational context in the testing period. Other manipulations that would reduce the large discrepancy in reward frequency between the training and operational conditions, such as reducing the target-prevalence frequency in training, may also be important manipulations to test.

Interestingly, however, there were trend associations indicating some potential small effects of NCR. First, there was a trend for NCR to increase the probability of the dog making a response (reduced timeouts) and a trend for dogs to show shorter latency during the operational context. However, NCR also showed a trend to increase the false-alert rate. Together, these results do suggest that there may be a minor effect of NCR to increase search activity, but when there were infrequent targets, the dogs became more likely to false alert but not actually improve accuracy. Together, these trends suggest that the current version of NCR did not effectively improve detection dog performance, but do not rule out that a permutation of the procedure could be effective.

For example, when NCR is used as a behavioral treatment, it is common to not deliver rewards immediately after an undesirable behavior (Pfaller-Sadovsky et al., 2019). A similar practice of a 5-s delay of NCR after a trial ended was implemented, but this delay may not have been sufficient. Perhaps a permutation of the procedure in which NCR only follows an all-clear response and with reduced NCR delivery frequency, a greater and more consistent effect maybe observed. Such a procedure would no longer be "noncontingent" but would reinforce correct all-clear responses on an intermittent schedule regardless of accuracy. The only difference is that this modified NCR would potentially reinforce all-clear responses even if a dog missed a target (whereas reinforcing only correct all clears would not reinforce in this condition). Given that false all clears occurred so infrequently during the test phase, this may not be a substantial concern. This modified NCR procedure, however, would be an important procedure to test because if it were deployed operationally, a handler would not know if a dog missed a live target. Thus, this NCR procedure would better mimic an operational deployment where rewards are delivered in the absence of a found target but adequate search was completed. The data suggest that there could potentially be a benefit of such a procedure that warrants further evaluation in a future experiment. Experiment 2 took a different approach to mitigate performance decrement by looking at the effects of Pavlovian appetitive stimuli.

EXPERIMENT 2: PAVLOVIAN APPETITIVE STIMULI TO INCREASE SEARCH

Materials and methods

Animals

The same animals were used in Experiment 2 as in Experiment 1.

Apparatus

The same apparatus, odorants, contexts, and group assignments were used in Experiment 2 as in Experiment 1.

Experimental design

After the test period of Experiment 1, the dogs received five 40-trial sessions in the operational and training contexts at 90% odor prevalence rate over a period of 10 days (e.g., one session per day, where context alternated across days). The first four sessions served as the "recovery" period in Experiment 1. In addition, the experimental dogs received one daily Pavlovian conditioning session in the operational context separate from their daily detection training. The Pavlovian conditioning consisted of placing the experimental dogs in the operational context room for 1 hr. During this time, a tone (432 Hz; conditioned stimulus [CS]) was presented for 30 s at 2-min intervals for 60 min. When the tone was presented, an automated feeder was activated, and food was delivered multiple times while the tone was on. The control group did not receive Pavlovian conditioning. After 10 days of Pavlovian conditioning, every dog received two additional training sessions in the operational and training contexts as a "refresher" of the detection task. These two refresher sessions served as the baseline performance, and the tone was not presented during these sessions.

After completion of the baseline period, dogs started the testing period. The target odor prevalence in the training context was 90% and reduced to 10% in the operational context. The dogs received five 40-trial sessions in both the operational and training contexts during the testing period. During the test period in the operational context, the tone (CS) was presented pseudorandomly to the experimental dogs at approximately 2-min intervals or approximately every two trials. The tone started at the start of the trial and ended after 30 s or after the dog made a response terminating the trial. No food was delivered with the tone during testing. The tone was not presented to the control group.

Pavlovian conditioning was done in the operational context to facilitate transfer of conditioned responding to the operational context. Furthermore, the control group did not receive a control manipulation or presentations of the tone to mimic how such a procedure might be used in practice, comparing dogs that would receive the complete manipulation versus not receive the manipulation. Thus, potential effects of the tone are not limited to Pavlovian conditioning in this experiment (i.e., the tone presence alone may have an effect) but reflect the potential effect of the entire intervention compared with no intervention.

Statistical analysis

All data were analyzed with identical models and dependent variables as Experiment 1. In addition to these models, we conducted a within-subject analysis for the experimental group in which the same dependent variables were predicted by the tone presence versus absence for the experimental group in the operational context (where the tone alternated from being on and off).

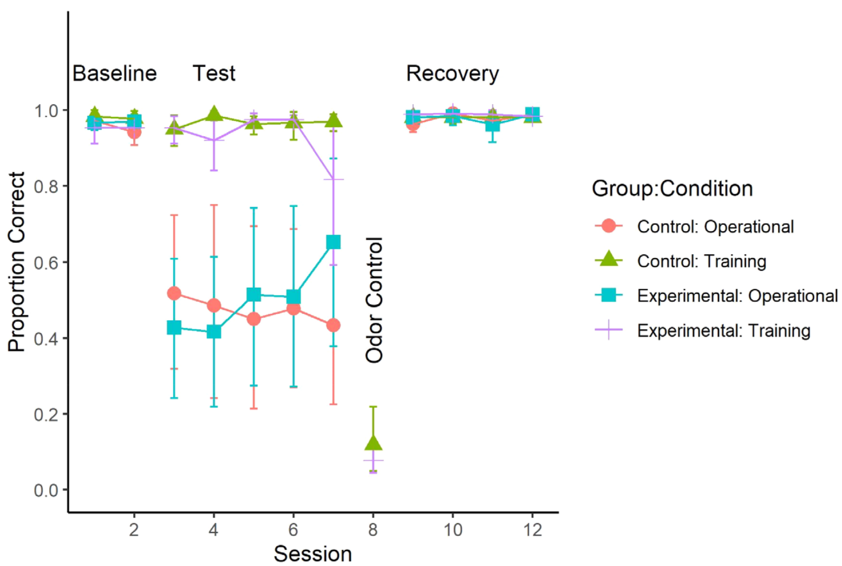

FIGURE 4 Mean proportion correct by group and context for Experiment 2. The error bars represent 95% confidence intervals of the proportion of odor trials that resulted in a correct response during Experiment 2. Both the control and experimental groups showed similar results. Session 8 is the "odor control," demonstrating that dogs cannot identify the correct odor port using unintentional cues when the target odor is not presented.

Results

Figure 4 shows dog performance by group (experimental and control). In the operational context, both groups had performance less than 75%, well below the training criterion of 85%, suggesting the appetitive conditioned stimulus did not maintain performance at our training criterion. Figure 5 shows individual dog performance by group (experimental and control). Similar to Experiment 1, some dogs were more resistant to disruption (i.e., Charles, Sasha,

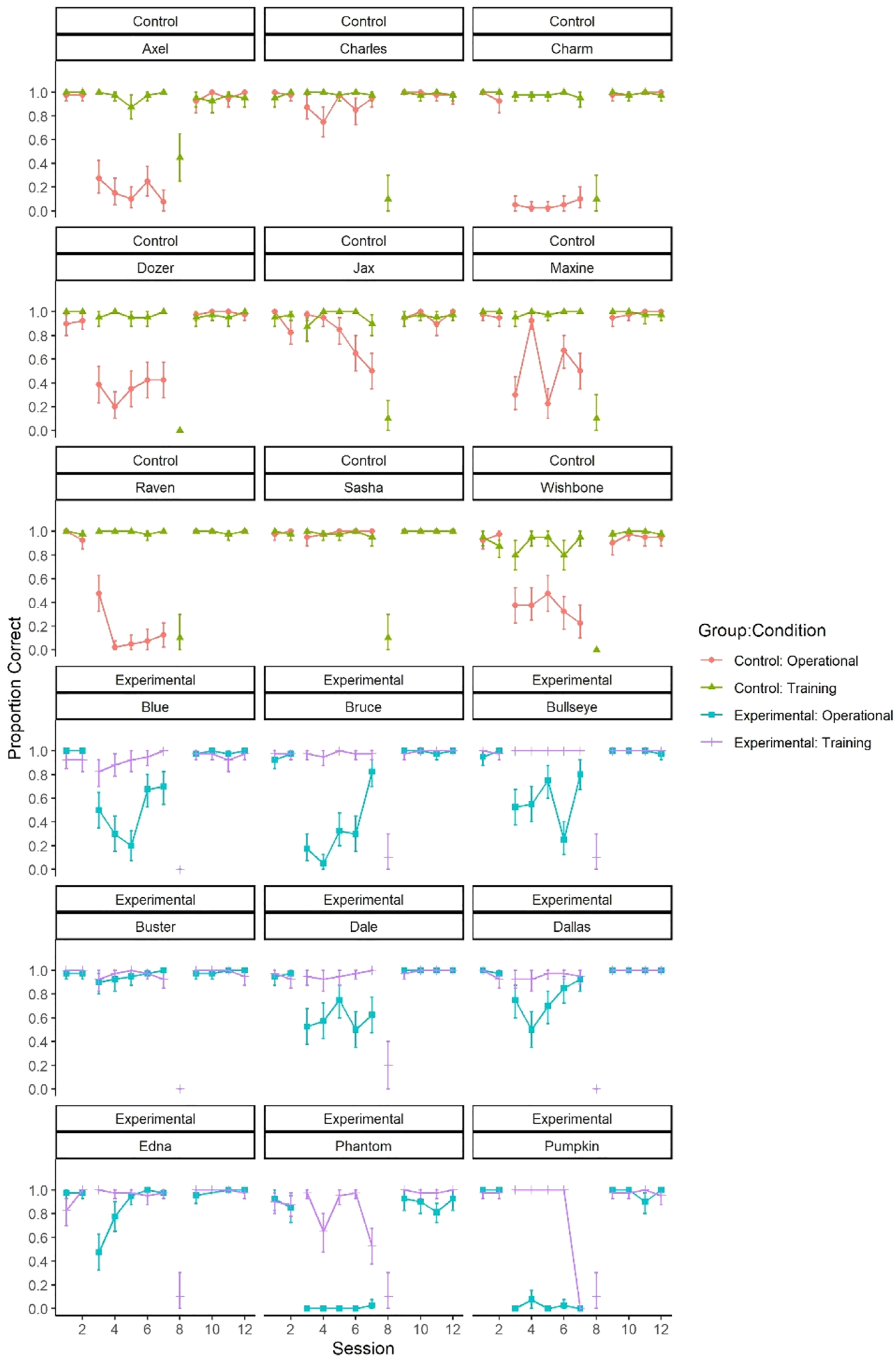

FIGURE 5 Individual dogs' mean proportion correct for each session, graphed similarly to Figure 4. The phases are as follows: baseline, test, odor control, recovery. Most dogs show a substantial performance disruption in the operational context during the test phase. Similar disruption is seen for experimental and control dogs. Session 8 is the odor control, demonstrating that dogs cannot identify the correct odor port using unintentional cues in the absence of the target odor.

TABLE 2 Mean ± standard deviation of the control and experimental group overall performance during the baseline, testing, and recovery phase during Experiment 2

Period	Baseline				Testing				Recovery			
Context	Training		Operational		Training		Operational		Training		Operational	
Group	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental
Accuracy, %	96.78 ± 2.42	95.24 ± 2.34	96.58 ± 2.37	92.73 ± 9.15	96.72 ± 3.15	92.79 ± 7.19	47.29 ± 34.10	50.38 ± 33.40	98.05 ± 1.63	98.83 ± 1.23	97.91 ± 1.76	97.73 ± 3.92
False, %	0.83 ± 1.18	2.00 ± 2.06	0.75 ± 0.74	1.94 ± 3.54	1.27 ± 1.92	2.10 ± 1.75	4.61 ± 4.89	6.16 ± 7.30	0.69 ± 1.63	0.34 ± 0.70	0.69 ± 1.05	0.38 ± 0.49
Timeout%	1.19 ± 1.45	1.35 ± 1.90	1.42 ± 1.40	4.12 ± 7.41	1.00 ± 1.19	4.21 ± 7.77	47.97 ± 33.59	43.38 ± 35.40	0.62 ± 0.69	0.62 ± 0.98	0.90 ± 1.36	1.60 ± 3.57
Hits, %	97.44 ± 2.39	96.18 ± 2.08	97.35 ± 2.23	93.51 ± 8.49	97.71 ± 2.44	93.53 ± 7.57	56.66 ± 34.00	63.88 ± 37.97	98.53 ± 1.27	99.16 ± 0.83	98.45 ± 1.46	97.82 ± 3.82
Correct rejections, %	90.87 ± 11.02	86.80 ± 14.21	89.68 ± 7.24	85.71 ± 20.51	87.77 ± 13.25	86.11 ± 10.24	46.25 ± 34.28	48.88 ± 33.06	93.75 ± 12.10	95.83 ± 6.98	93.05 ± 9.60	96.94 ± 6.09
False all clear, %	1.33 ± 1.34	1.56 ± 1.45	1.38 ± 1.44	1.42 ± 1.19	1.12 ± 0.93	1.02 ± 1.25	1.66 ± 3.53	0.73 ± 2.07	0.69 ± 0.92	0.21 ± 0.32	0.54 ± 0.77	0.30 ± 0.49

Note. Each session was 40 trials. The target odor prevalence was 90% in both contexts during baseline and recovery, 90% in training context for testing, and 10% in operational context in testing. Accuracy: percentage of trials that a dog made a correct response by either alerting to the target odor or calling a correct all clear. False: percentage of trials that a dog alerted to an incorrect port during a target odor or blank trial. Timeout: percentage of trials the dog "timed out" without searching all three odor ports or alerting to a port. Hit: percentage of odor present trials that a dog alert to the target odor. Correct rejection: percentage of blank (odor absent) trials dogs did a correct all clear. False all clear: percentage of trials dogs made an incorrect all-clear response when target odor was present.

and Buster), which seemed unrelated to experimental group. Most dogs, however, showed a substantial drop in performance in the operational context. Table 2 shows the mean accuracy rate, false alerts, timeout rate, hit rate, correct-rejection rate, and false-all-clear rate by group, context, and period for Experiment 2.

Baseline

During the baseline period (the two refresher sessions prior to the test period), there was a significant interaction between group and context ($\chi^2 = 8.57$, $p = .003$) and the main effect of session was not significant ($\chi^2 = 2.34$, $p = .13$). To analyze the interaction between group and context, Tukey-adjusted post hoc tests were conducted. There was no difference between the experimental and control group in the operational context (est = 0.22, $z = 0.56$, $p = .57$). There was, however, a significant difference between the control (98% accuracy) and experimental group (96% accuracy) in the training context, where control dogs showed slightly better performance ($z = 2.42$, $p = .02$). This may be due to the 10, 1-hr Pavlovian training sessions the experimental dogs received immediately prior to the two refresher sessions. Nonetheless, both groups of dogs had highly similar performance in the training context (within 2% accuracy) and did not differ statistically in the operational context (see Figure 4).

Accuracy

During the testing period, the interaction between group and context for accuracy was significant ($\chi^2 = 8.53$, $p = .003$). Tukey-adjusted post hoc tests were conducted to analyze this interaction. Post hoc tests indicated that for the operational context there was no difference between the control and experimental groups (est = 0.06, $z = 0.07$, $p = .94$). Furthermore, there was no difference between the experimental and control group in the training context (est = 0.69, $z = 0.88$, $p = .38$), indicating no clear improvement due to the CS tone. The decrease in performance for the experimental group during the last session in the training room during the testing period was driven by two dogs, Pumpkin and Phantom; all the other dogs had similar performance in the last sessions when compared with the other four sessions (see Figure 5). In addition, there was a slight nonsignificant increase in performance in the last session for the experimental group in the operational context during the testing period that was driven by Bruce and Bullseye. It is possible that the experimental condition of the Pavlovian appetitive stimuli had a positive effect for Bruce and Bullseye; however, it is also possible the increase in accuracy could be a learning effect of repeated experience of the operational task, owing to the presence of reinforced targets present and not necessarily an effect of tone (see Figure 5).

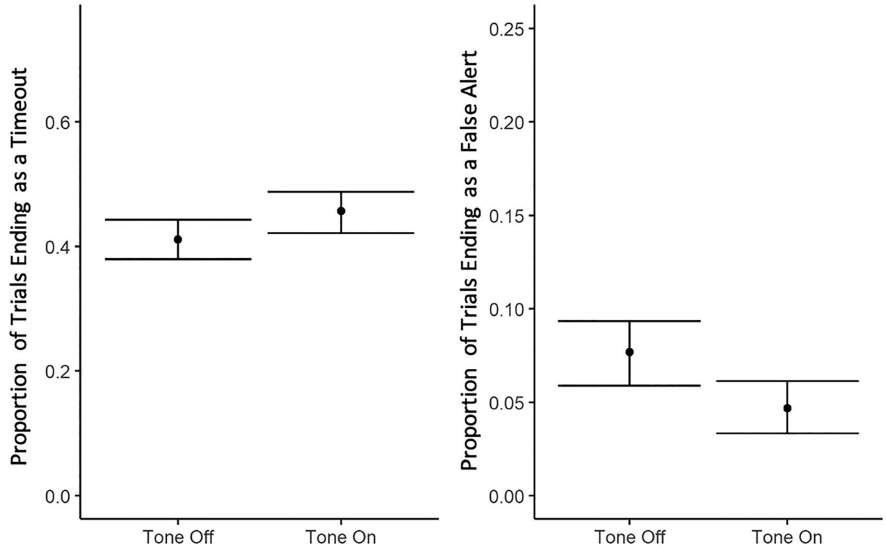

FIGURE 6 Within-subject comparison of trials with the tone on and off for experimental dogs in the operational context. Error bars show the 95% confidence interval, and points show the mean. Trials that started with the tone CS were more likely to end in a timeout and less likely to end in a false alert.

To further assess this potential effect of the tone on accuracy, the accuracy for the experimental group in the operational context on trials in which the tone was present was compared with trials in which the tone was off. This within-subject analysis similarly revealed no effect of tone on detection accuracy ($\chi^2 = 0.62$, $p = .43$). Furthermore, when only the last test session where increases were most apparent for the experimental group was analyzed, we found there was still no significant difference in accuracy between when the tone was on versus off ($\chi^2 = 2.38$, $p = .12$). Furthermore, accuracy was 68% when the tone was off compared with 63% when the tone was on.

Latency

During the testing period, whether the treatment reduced latency to approach the panel for experimental dogs was assessed. There was a significant group by context interaction ($\chi^2 = 58.46$, $p < .001$) and the effect of session was not statistically significant ($\chi^2 = 1.55$, $p = .21$). Post hoc tests for the interaction indicated that latency was not different between the control and experimental group for the operational context ($z = 0.64$, $p = .53$) or between groups for the training context ($z = 0.39$, $p = .70$). Next, a within-subject analysis was conducted analyzing specific trials within the session in which the tone was present or absent. Again, the presence of the tone had no effect on latency to approach the panel ($\chi^2 = 1.59$, $p = .19$).

Timeout probability

During the testing period, responding was not predicted by context and group interaction ($\chi^2 = 0.59$, $p = .44$). There was not a significant effect of session ($\chi^2 = 0.32$, $p = .56$) or group ($\chi^2 = 0.03$, $p = .86$); however, there was a strong association with context ($\chi^2 = 748$, $p < .001$). Thus, the probability of a timeout was similar across the testing phase within each context and substantially higher in the operational context. A within-session analysis of the experimental dogs in operational context indicated that there was an increase in the timeout probability for trials in which the tone was on versus off ($\chi^2 = 6.96$, $p < .01$). When the tone was on, the dogs showed a timeout frequency of 45% of trials; whereas, when the tone was off, timeouts were reduced to 42% of trials (see Figure 6).

Hits

The probability of a correct hit was predicted by a significant context by group interaction ($\chi^2 = 11.93$, $p < .001$) and session ($\chi^2 = 9.82$, $p < .001$), where hit rate decreased across sessions. Post hoc tests, however, indicate that hit rate was not different between the control and experimental group in the operational context ($z = 0.41$, $p = .68$) or training context ($z = 1.44$, $p = .15$). The within-subject analysis within a session indicated that there was no difference in the hit rate whether the tone was on or off ($\chi^2 = 0.98$, $p = .32$).

False alerts

The frequency of false alerts was not predicted by a group and context interaction ($\chi^2 = 0.42$, $p = .51$) or group ($\chi^2 = 0.64$, $p = .42$). However, there was a significant effect of session ($\chi^2 = 11.93$, $p < .001$) and context

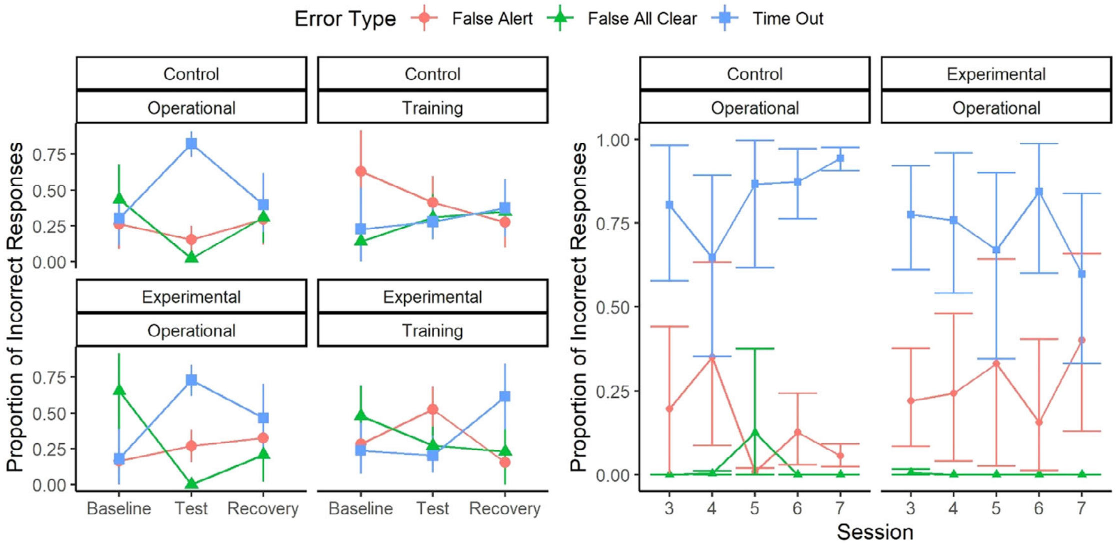

FIGURE 7 Error-type frequency expressed as a proportion of errors made for Experiment 2. Left: shows the false alerts, false all clears, and timeouts averaged across baseline, test, and recovery for each context and group. Right: shows the changes across session for the context in which errors were most frequent (i.e., the operational context) where approximately half of trials ended in an error. The points show the mean, and the error bars show the 95% confidence intervals. Timeouts increased substantially in the operational environment during the test phase, and this was similar across experimental and control dogs.

(χ^2 = 68.47, p < .001). The within-subject analysis, however, indicated that on trials in which the tone was present, the probability of a false alert was significantly lower than on trials in which it was absent (4.6% tone on vs. 7.6% tone off; χ^2 = 7.67, p < .001; see Figure 6).

Error types as a proportion of errors

Figure 7 shows the error types expressed as a proportion of errors made. During test, a vast majority of errors made for both experimental and control dogs were timeouts. Unlike Experiment 1, with the NCR treatment, there was no evidence of increases in false alerts or reduction in timeouts. The experimental treatment appeared to have little influence on the types of errors made; both groups showed a substantial increase in timeouts compared with false alerts and false all clears. We did not perform further null-hypothesis significance testing due to the infrequent total number of errors during baseline and recovery periods as well as some error types almost never occurring in some conditions (i.e., false all clears).

Control test

Nearly identical to Experiment 1, the average accuracy of the dogs during the control test was 11.5% and was well below chance. The control test again confirms that the dogs cannot identify the olfactometer-programmed correct port in the absence of target odor, indicating that the dogs' performance was driven by the presence of the target odor and not unintentional cues from the olfactometer.

Discussion

Overall, the appetitive CS tone in PIT did not improve measures of detection dog performance. The only positive finding was that the within-session analysis indicated that the false-alarm rate was significantly lower for trials in which the tone was on versus off. This occurred simultaneously with an increase in timeouts, yet there was no overall improvement in accuracy when the tone was on compared with off. One possible explanation for this is that when the tone turned on at the start of the trial, the dogs engaged in tone-directed behavior (looking toward the tone source, i.e., sign tracking; Hearst & Jenkins, 1974), which was incompatible with a false alert. A sign-tracking response involves designation of attention to the CS, which transforms the CS from a predictor of the unconditioned stimulus (US) into a reward stimulus (Robinson & Flagel, 2009). Thus, the tone may have acted to reduce false alerts, or responding at all, by generating a conditioned response to the tone that was incompatible for an alert, thereby decreasing false alerts and increasing timeouts. Nonetheless, overall accuracy and latency was not influenced by the tone, suggesting that it had no positive influence on detection dog performance.

Similar to what was observed in Experiment 1, the performance decrease in the operational context was almost entirely related to poor search behavior (increased timeouts) and not the extinction of an alert response when a target was detected. This is highlighted by the average percentage of false-all-clear rates (the dog investigating all ports and calling an all clear when target odor is present; Table 2). The average false-all-clear rate in the operational context for the experimental group was 0.7%

and for the control group was 1.66%, meaning that when the dogs did fully search the odor ports, almost no targets were missed regardless of the conditions. This suggests that most errors were therefore driven by incomplete search.

Although this experiment did not demonstrate that an appetitive stimulus with PIT can improve detection performance in low target odor prevalence environments in dogs, elevations of responding in extinction have been demonstrated in other animal models. In previous studies that used rodents (Campese et al., 2017; Galarce et al., 2007; Holland, 2004) and humans (Cartoni et al., 2013, 2015; Prevost et al., 2012; Trick et al., 2011) researchers were successful in demonstrating that an appetitive conditioned stimulus in Pavlovian instrumental transfer enhanced the response rate in extinction. One possible explanation for the lack of effect on latency or timeouts in Experiment 2 may be a species difference or the fact that our search task is more complex than the single operant typically used in PIT studies. The search behavior chain employed in this study requires investigation of three ports, and perhaps general motivational increases that are sufficient for increases in single-operant lever pressing may not be sufficient to complete a three-port search task.

GENERAL DISCUSSION

Overall, Experiments 1 and 2 found similar performance decrements when the target odor prevalence was decreased, replicating the effect observed in previous studies (Aviles-Rosa et al., in press; Gazit et al., 2005; Porritt et al., 2015). Importantly, Experiments 1 and 2 found that NCR and PIT did not improve detection dog accuracy and other performance metrics under these conditions in an operationally relevant manner for detection dog handlers.

Nonetheless, in Experiment 1 we did find trends suggesting that NCR may increase some search-related behavior (latency and reduced timeouts), but this also led to increased false alerts. We found when a 5-s delay after a trial ended was implemented, the dogs did not false alert as frequently compared with Cohort 1 when the reward was delivered immediately at the start of the intertrial interval. However, this delay may not have been sufficient and future studies manipulating the delay are needed. One potential permutation of the NCR may be rewarding following an all-clear response; however, this would be contingent on a particular response (i.e., all clear) and not technically follow NCR methods (i.e., reward is noncontingent). In addition, this permutation of NCR would reward a dog even if a dog missed a target. This modified NCR procedure, however, would be an important procedure to test because it is unknown whether rewarding a dog for an all-clear response would change performance substantially. In addition, this modified NCR would need to be studied in operational dogs because if such a procedure would be deployed operationally, a handler would not know if a dog missed a live target, which could be detrimental. The results do suggest that there could be a permutation of the NCR procedure that may produce operationally relevant performance increases, but this requires additional study.

In Experiment 2, PIT did not lead to increased detection performance in our experimental context. The only observed effect was within session, wherein false alerts decreased and timeouts increased when the tone was on compared with off. We hypothesized that the tone may have caused incompatible sign-tracking responses that inhibited false alerts (such as looking at the speaker) and search responses altogether; however, we did not video-code the behavior of the dogs during the Pavlovian conditioning and test periods. This was an anecdotal experimenter observation that requires further testing/analysis to confirm. It would be interesting to record whether the tone led to orienting responses to the speaker, which led to the decrease in responses/increases in timeouts.

We also noted a trend for increasing performance in the experimental group toward the end of the test period in Experiment 2 (see Figure 4), although this difference did not reach statistical significance. Any improvement in the experimental dogs in Experiment 2 did not appear to be related to the tone CS because within-session performance was numerically lower for trials where the tone was present versus absent. One potential rationale for the increasing performance of experimental dogs overall is that these dogs had more exposure to changing contingences across Experiments 1 and 2 compared with control dogs. Because the same dogs were used as experimental and controls from Experiment 1 to 2, more generalized experience with varying contingencies in the operational context may have increased their sensitivity to the occasional presence of target odor.

There are further notable limitations of these experiments. First, the level of Pavlovian conditioning was not directly confirmed prior to testing. Even though the conditioning period consisted of 10 days, which is within the range of a previous study (Hall et al., 2015), a test was not conducted to confirm the level of Pavlovian conditioning. However, given the decrease in the false-alert rate and increase in the timeout rate with tone presence, we hypothesize this was not a primary limitation.

For both experiments, a primary limitation is that for logistical and cost reasons there was a fixed order in the experimental conditions and a lack of parametric manipulations. We only tested one frequency of NCR and PIT Pavlovian conditioning parameters. This leaves open the potential that different parameters might have found an effect. In addition, overlapping of Pavlovian conditioning for Experiment 2 and recovery from Experiment 1 may have had unintended consequences. In the two refresher baseline sessions in Experiment 2, there remained a minor difference in accuracy between the control and experimental group in the training context (98% vs. 96%). This was likely a carryover effect from Experiment 1 or an incidental consequence of the Pavlovian conditioning phase in Experiment 2. Nonetheless, if PIT had a substantial effect, it was likely that we would have still been able to detect such an effect given the substantial decline in performance from baseline for both groups. Nonetheless the consistent order for the two experiments was selected (rather than randomizing across cohorts to reduce carryover effects) because the sample size (18 dogs) was insufficient to include as an additional variable.

Another important limitation is that the level of training of these dogs does not closely match that for operational dogs. Operational dogs typically have greater than 1 year of weekly training, whereas the dogs in the present experiments had only one month of training. This limitation was acceptable to allow for rapid assessment of several procedures to mitigate potential performance decrements. If 18 dogs were required to be trained for greater than one year, the study would unlikely be feasible. Nonetheless, this does suggest that the magnitude of the effects maybe different in operational dogs that have substantially longer histories of training for odor detection.

A final consideration is that the absolute magnitude of decrease in performance observed in the operational room may be in part related to a contrast effect between the training and operational room target odor prevalence. By having the dogs rotate between these conditions within the same day, it's possible the drop in performance may have been greater than what would be expected if there were not such a consistent contrast to the training room and suggests that the magnitude of effect may not be as great in dogs that frequently deploy in new operational environments.

In conclusion, NCR and PIT did not improve detection performance in an operationally relevant way. NCR showed potentially relevant trends, but modification to the NCR procedure is likely necessary to reduce the potential for false alerts or other negative consequences. This warrants further study and analysis, but PIT did not yield any trends indicative that it could be a useful intervention when dogs must search in low target odor environments.

ACKNOWLEDGMENTS

We thank the undergraduate students for assisting in data collection and husbandry of the dogs.

FUNDING STATEMENT

This research was funded by a contract from the United States Department of Homeland security (DHS) Science and Technology Directorate (S&T), Contract: 70RSAT20CB0000010. Any opinions, findings, and conclusion or recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the sponsor.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ETHICS STATEMENT

All procedures used in both experiments were reviewed and approved by TTU Institutional Animal Care and Use Committee (ACUC #19093–10).

ORCID

Edgar O. Aviles Rosa https://orcid.org/0000-0002-9551-771X

REFERENCES

Aviles-Rosa, E. O., DeChant, M. T., Prada-Tiedemann, P. A., & Hall, N. J. (in press). A laboratory model of canine search vigilance decrement, I. *Journal of Experimental Analysis of Behavior.*

Aviles-Rosa, E. O., Gallegos, S. F., Prada-Tiedemann, P. A., & Hall, N. J. (2021). An automated canine line-up for detection dog research. *Frontiers in Veterinary Science*, *8*, Article 775381. https://doi.org/10.3389/fvets.2021.775381

Campese, V. D., Kim, I. T., Rojas, G., & LeDoux, J. E. (2017). Pavlovian extinction and recovery effects in aversive Pavlovian to instrumental transfer. *Frontiers in Behavioral Neuroscience*, *11*. https://doi.org/10.3389/fnbeh.2017.00179

Cartoni, E., Balleine, B., & Baldassarre, G. (2016). Appetitive Pavlovian-instrumental transfer: A review. *Neuroscience & Biobehavioral Reviews*, *71*, 829–848.

Cartoni, E., Puglisi-Allegra, S., & Baldassarre, G. (2013). The three principles of action: A Pavlovian-instrumental transfer hypothesis. *Frontiers in Behavioral Neuroscience*, *7*, Article 153. https://doi.org/10.3389/fnbeh.2013.00153

Corbit, L. H., & Balleine, B. W. (2011). The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *Journal of Neuroscience*, *31*(33), 11786–11794. https://doi.org/10.1523/JNEUROSCI.2711-11.2011

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Sage.

Galarce, E. M., Crombag, H. S., & Holland, P. C. (2007). Reinforcer-specificity of appetitive and consummatory behavior of rats after Pavlovian conditioning with food reinforcers. *Physiology & Behavior*, *91*(1), 95–105. https://doi.org/10.1016/j.physbeh.2007.01.021

Gazit, I., Goldblatt, A., & Terkel, J. (2005). The role of context specificity in learning: The effects of training context on explosives detection in dogs. *Animal Cognition*, *8*(3), 143–150. https://doi.org/10.1007/s10071-004-0236-9

Holland, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*(2), 104–117. https://doi.org/10.1037/0097-7403.30.2.104

Holmes, N. M., Marchand, A. R., & Coutureau, E. (2010). Pavlovian to instrumental transfer: A neurobehavioural perspective. *Neuroscience & Biobehavioral Reviews*, *34*(8), 1277–1295.

Lambert, J. M., Bloom, S. E., Samaha, A. L., Dayton, E., & Kunnavatana, S. S. (2016). Effects of noncontingent reinforcement on the persistence and resurgence of mild aggression. *The Psychological Record*, *66*(2), 283–289. https://doi.org/10.1007/s40732-016-0170-7

Lindblom, L. L., & Jenkins, H. M. (1981). Responses eliminated by noncontingent or negatively contingent reinforcement recover in extinction. *Journal of Experimental Psychology*, *7*(2), 175–190.

Marcus, B. A., & Vollmer, T. R. (1996). Combining noncontingent reinforcement and differential reinforcement schedules as treatment for aberrant behavior. *Journal of Applied Behavior Analysis*, *29*(1), 43–51. https://doi.org/10.1901/jaba.1996.29-43

Meindl, J. N., Ivy, J. W., Glodowski, K. R., & Noordin, K. (2021). Applying standards of effectiveness to noncontingent reinforcement: A systematic literature review. *Behavior Modification*, *45*(4), 619–640. https://doi.org/10.1177/0145445519865073

Pfaller-Sadovsky, N., Arnott, G., & Hurtado-Parrado, C. (2019). Using principles from applied behaviour analysis to address an undesired behaviour: Functional analysis and treatment of jumping up in companion dogs. *Animals*, *9*(12), Article 1091. https://doi.org/10.3390/ani9121091

Porritt, F., Shapiro, M., Waggoner, P., Mitchell, E., Thomson, T., Nicklin, S., & Kacelnik, A. (2015). Performance decline by search dogs in repetitive tasks, and mitigation strategies. *Applied Animal Behaviour Science*, *166*, 112–122. https://doi.org/10.1016/j.applanim.2015.02.013

Prevost, C., Liljeholm, M., Tyszka, J. M., & O'Doherty, J. P. (2012). Neural correlates of specific and general Pavlovian-to-instrumental transfer within human amygdalar subregions: A high-resolution fMRI study. *Journal of Neuroscience*, *32*(24), 8383–8390.

Rescorla, R. A., & Cunningham, C. L. (1977). The erasure of reinstated fear. *Animal Learning & Behavior*, *5*, 386–394.

Rescorla, R. A., & Heth, C. D. (1975). Reinstatement of fear to an extinguished conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, *1*(1), 88–96.

Robinson, T. E., & Flagel, S. B. (2009). Dissociating the predictive and incentive motivational properties of reward-related cues through the study of individual differences. *Biological Psychiatry*, *65*(10), 869–873. https://doi.org/10.1016/j.biopsych.2008.09.006

Russell, V. L. (2016). Least-squares meaans: The R package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33.

Thrailkill, E. A., Kacelnik, A., Porritt, F., & Bouton, M. E. (2016). Increasing the persistence of a heterogeneous behavior chain: Studies of extinction in a rat model of search behavior of working dogs. *Behavioural Processes*, *129*, 44–53. https://doi.org/10.1016/j.beproc.2016.05.009

Thrailkill, E. A., Porritt, F., Kacelnik, A., & Bouton, M. E. (2018). Maintaining performance in searching dogs: Evidence from a rat model that training to detect a second (irrelevant) stimulus can maintain search and detection responding. *Behavioural Processes*, *157*, 161–170. https://doi.org/10.1016/j.beproc.2018.09.012

Trick, L., Hogarth, L., & Duka, T. (2011). Prediction and uncertainty in human Pavlovian to instrumental transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 757–765. https://doi.org/10.1037/a0022310

Yager, L. M., & Robinson, T. E. (2010). Cue-induced reinstatement of food seeking in rats that differ in their propensity to attribute incentive salience to food cues. *Behavioural Brain Research*, *214*(1), 30–34. https://doi.org/10.1016/j.bbr.2010.04.021

How to cite this article: DeChant, M., Aviles Rosa, E. O., Prada-Tiedemann, P. A., & Hall, N. J. (2023). A laboratory model of canine search vigilance decrement, II: Noncontingent reward and Pavlovian appetitive stimuli. *Journal of the Experimental Analysis of Behavior*, 1–17. https://doi.org/10.1002/jeab.838