Received: 11 October 2022 | Accepted: 16 February 2023
DOI: 10.1002/jeab.841

RESEARCH ARTICLE

Journal of the Experimental Analysis of Behavior
SEAB

Part, III: Increasing odor detection performance after training with progressively leaner schedules of odor prevalence

Mallory T. DeChant[1] | Edgar Aviles-Rosa[1] | Paola Prada-Tiedemann[2] | Nathaniel J. Hall[1]

[1]Department of Animal and Food Sciences, Texas Tech University, Lubbock, TX, United States

[2]Department of Environmental Toxicology, Texas Tech University, Lubbock, TX, United States

Correspondence
Mallory T. DeChant, Department of Animal and Food Sciences, Texas Tech University, Lubbock, TX, United States.
Email: mallory.dechant@ttu.edu

Funding information
U.S. Department of Homeland Security Science and Technology Directorate, Grant/Award Number: 70RSAT20CB0000010

Editor-in-Chief: Mark Galizio
Handing Editor: Mark Galizio

Abstract
Prior work has demonstrated canine search behavior and performance declines when challenged with infrequent target odors. The purpose of this study was to evaluate whether performance could be maintained in a low target odor prevalence context by explicitly training dogs through progressively leaner target odor schedules. In Experiment 1, nine control dogs were trained at 90% target prevalence rate. Nine experimental dogs were trained with progressively lower prevalence rates in 10% increments until reaching 20% prevalence with > 85% detection accuracy in the training context. Both groups were tested in the operational context at a 10% target odor prevalence. Experimental dogs had higher accuracy, hit percentage, and shorter search latency in the operational context compared with control dogs. In Experiment 2, twenty-three operational dogs were challenged with a target frequency of 10%, which resulted in 67% accuracy. Control dogs were then trained with 90% target frequency, whereas experimental dogs received a progressively decreasing target rate from 90% to 20%. The dogs were rechallenged with target frequencies of 10, 5, and 0%. Experimental dogs outperformed control dogs (93% vs. 82% accuracy) highlighting the effect of explicit training for infrequent targets.

KEYWORDS
canine vigilance, detection dogs, dog search behavior, low target odor prevalence, search extinction

INTRODUCTION

Dogs are trained for the detection of a wide range of materials such as explosives, narcotics, ignitable liquids (Abel et al., 2020; Jezierski et al., 2014; MacLean & Hare, 2018; Marks, 2002) providing critical detection capabilities for national security. Similar to a wide range of human safety workers, dogs are routinely asked to search in specific environments (i.e., airports, cargo shipments, or stadiums) for long durations, where the probability of finding a true target can be exceedingly low.

Prior research has demonstrated that dogs can be highly sensitive to contexts that are indicative of a low probability of encountering a target stimulus. For example, Gazit et al. (2005) led explosives detection dogs to search along two distinct paths (A and B). Along Path A, five targets were present, whereas Path B had no targets present (Gazit et al., 2005). The dogs quickly showed reduced search along Path B such that when explosives were hidden along this path, the dogs showed a substantial reduction in detection. Furthermore, Porritt et al. (2015) found similar results where a group of dogs that repeatedly searched an area with no target odors showed a substantial search decrement and detection in those areas. Most recently our lab has developed a laboratory model of this effect (Aviles-Rosa et al., 2023) using automated olfactometer equipment. Dogs showed substantial decrement (i.e., accuracy and search performance behavior) when the prevalence of the target odor was rapidly reduced from 90% in a training context to 10% in a simulated "operational" context. Together, these results suggests that dogs can be acutely sensitive to the prevalence rate of the target odor in the search environment and that low target odor frequency in the search environment leads to a decrement in search behaviors and detection performance.

One suggestion to mitigate this decrement in performance was from a follow up study by Porritt et al.

(2015), where the authors trained dogs to indicate to a nonexplosive, innocuous odor. When this innocuous odor was placed in the search path with no other explosives, the dogs' performance improved and was sufficient to maintain search behavior (Porritt et al., 2015). These results suggest that to improve dogs' detection performance, at least occasional reinforced targets are needed, which can be done via increased prevalence of target odors or by planting an innocuous odor in the search environment. However, providing operational finds to the dogs at a frequency sufficient to maintain search behavior may not always be feasible due to logistical and safety reasons.

Basic laboratory research on behavioral persistence under extinction conditions suggests that training with intermittent schedules of reinforcement leads to greater resistance to extinction (partial reinforcement effect; Chan & Harris, 2019; Harris et al., 2019; Haselgrove et al., 2004; Lewis, 1960; Nevin, 1988). Importantly, this effect seems to be maintained by animals learning about the presence of nonreinforced trials (Harris et al., 2019; Haselgrove et al., 2004) such that learning the presence of nonreinforced trials helps maintain responding when the animal is not reinforced under extinction or near-extinction conditions. This behavioral principle can be applied to detection dogs in an attempt to maintain search behavior in contexts with low target odor frequencies that resemble extinction procedures.

Detection dog behavior can be considered a behavior chain with two links. A discriminative stimulus evokes trained search behavior, leading to the appearance of a target odor. The target odor serves as a conditioned reinforcer for the search behavior and as a discriminative stimulus to occasion an alert behavior, which is reinforced with the terminal reinforcer (food or toy). When the search behavior is not reinforced with the presence of a target odor, search behavior and performance decreases (e.g., Aviles-Rosa et al., in press; DeChant et al., in press, Parts I and II this issue; Gazit et al., 2005; Porritt et al., 2015). Previous work in rodents on the use of intermittent schedules of reinforcement on resistance to extinction (Chan & Harris, 2019; Harris et al., 2019; Haselgrove et al., 2004) suggests that when applied to detection dogs, training dogs with an intermittent schedule of reinforcement (e.g., training with an odor frequency similar to what may be anticipated in an operational scenario) may improve performance.

Recent work in rodents developed an analog model for working dogs in which rodents had to manipulate one object (lever or chain) to simulate a dog's "search" behavior followed by making another response to simulate a dog's "alert" behavior (Thrailkill et al., 2016). This study found that intermittent reinforcement of the simulated search behavior (compared with continuous reinforcement) led to greater resistance to extinction of the search behavior. This finding further suggests that training dogs with a more intermittent frequency of target odor may lead to greater resistance to extinction in operational contexts.

Producing intermittent odor frequency for detection dog search can be accomplished by including higher frequencies of "blank" trials. A blank trial is a trial in which only distractor odors are available and the target odor is not presented. Interestingly, despite previous work suggesting that intermittent reinforcement (or higher frequency of blank trials) may help with dogs' persistence in operational conditions, this does not seem to be a standard training practice. Recent work from DeChant et al. (2020) included a survey of the typical practices of law enforcement/professional canine handlers and hobby/sport canine handlers. A total of 38 law enforcement/professional handlers participated. Although 95% of handlers reported training their dogs once a week or more frequently, 75% of handlers reported running one or fewer blank runs during training sessions (DeChant et al., 2020). Further, 45% of handlers reported running fewer than one blank run per training session, and 13% never conducted a blank run. Given that this practice conflicts with the preliminary scientific evidence, an important next step is to evaluate whether increasing the frequency of blank trials during training of dogs searching for an operationally relevant odor leads to increased resistance to search extinction and overall better performance in a simulated operational test. The purpose of this study was to evaluate whether detection performance could be maintained in a low-target-odor-prevalence context by explicitly training dogs through a progressively leaner schedule of targets. Experiment 1 first established this effect in a laboratory model and Experiment 2 replicated and extended this work in operationally deployed detection canines.

EXPERIMENT 1: EFFECT OF INTERMITTENT ODOR PREVALENCE TRAINING ON DETECTION PERFORMANCE IN A LABORATORY MODEL

Materials and methods

Animals

Eighteen mixed-breed dogs (the same dogs used in Parts I and II of this issue, split into three different cohorts of six dogs each) were used and housed at the Texas Tech University (TTU) Canine Olfaction Lab and were participating in a training program to increase adoptability. The dogs' backgrounds were unknown, but all were presumably naïve to detection training when first brought to TTU and showed adequate detection capabilities in Parts I (Aviles-Rosa et al., in press) and II (DeChant et al., in press) of this series. The dogs received twice daily walks, social enrichment, and training for adoption. All procedures used in both experiments were reviewed and approved by TTU Institutional Animal Care and Use Committee (IACUC #19093–10).

Apparatus and general procedures

The dogs were trained to operate the three-alternative choice automated olfactometer system described by Aviles-Rosa et al. (2021) 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, respectively. These two contexts model how dogs are frequently trained in one area but operationally work 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 context (i.e., the training context for Cohort 2 was the operational context for Cohorts 1 and 3).

At the start of each session, an experimenter brought the dog into the appropriate room and the dog worked off leash independently of the experimenter. The experimenter was always blind to the position of the odorants (and presence of the target odor) in the odor ports.

The dogs were trained to detect and alert to the odor of double-based smokeless powder (SP), which is a propellant that most explosive detection dogs are trained to detect. In addition, there were five different distractor odors (cotton gauze, plastic gloves, blank jar, limonene (–) 10^{-3} v/v dilution in mineral oil, and mineral oil). These distractors were chosen based on typical items an operational dog may encounter, items generally used in the laboratory (i.e., gloves), and another odorant (i.e., limonene) that served as a novel strong odor.

At the start of every trial, all three olfactometers were activated to present an odor to its respective odor port. If the target odor was programmed to appear for that trial, the port in which it would appear was determined pseudorandomly such that it appeared approximately equally in all three ports within a session. If a target was not programmed to appear, all three olfactometers would present one of the five distractor odors. If an olfactometer was not programmed to present a target, the distractor odor was selected at random from a uniform distribution with equal probability of each distractor. Distractor selection occurred independently for each olfactometer; thus, the same distractor could appear in multiple ports within the same trial.

Once all three olfactometers were activated, a panel that covered the three odor ports was raised by the computer. At this point, the handler would tell the dog to "search." Infrared beams measured each nose entry to all three odor ports. Dogs were required to make a 4-s continuous nose hold as an "alert" response. If no target odor was present, dogs were trained to make an "all-clear" response by investigating all three ports and then removing their nose from the apparatus for four consecutive seconds. The first response, an alert or all clear, immediately terminated the trial, which led to the panel cover lowering to cover the odor ports. The olfactometers then activated an odor purge and an odor exhaust evacuated odor out of the room for the 20-s intertrial interval (ITI). If a dog did not make a response (an alert or an all clear) within 45 s of the trial start, typically by failing to search any or all three of the ports, the trial was terminated and scored as a "timeout." Scoring of all responses was done by the computer via the infrared beam detection of nose entries to each port. If the dog made a correct alert to the port presenting the target, the computer sounded a "correct" tone and triggered an automated

TABLE 1 Measures of performance

Measurement	Definition	Calculation
Overall accuracy, %	Dog alerted to the correct port during an odor trial or did a correct all clear during a blank trial.	(Number of correct responses /40 trials) × 100
False alerts, %	Dog alerted to the incorrect port on any trial (odor or blank)	(Number of false alerts/ 40 trials) × 100
Timeout, %	Dog did not search all three ports or make a response within 45 seconds after the trial started	(Number of timeouts / 40 trials) × 100
Hits, %	Dog alert to the odor correctly on an odor present trial.	(Number of correct responses to odor / number of odor trials) × 100
Correct rejections, %	Dogs made a correct all clear during a blank trial.	(Number of correct all clears / number of blank trials) × 100
False all clears, %	Dog response was an all clear after sampling the port containing the target odor	(Number of false all clears after the dog sampled all ports during an odor trial / number of odor trials the dog sampled all ports) × 100
Latency, s	Time required for a dog to search at least one port after a trial started. If the dog did not search during a trial the trial latency was 45 s.	($\sum$ Latency / 40)

feeder. If the dog made an incorrect alert (i.e., a false alert), the trial terminated and the panel lowered, covering the odor ports, and the computer initiated the ITI. If a dog made a correct all-clear response, the panel lowered and the computer initiated the ITI. These responses were not reinforced to mimic how detection dogs typically continue searching after clearing an area without targets and when no reinforcers are delivered. If a dog made an all-clear response when a target was present, the panel was lowered and the ITI started (similar to an operational setting if an unknown target is missed). Table 1 summarizes the various trial outcomes and measures of performance.

Experimental design

Baseline period
All dogs started in a baseline period that consisted of four 40-trial sessions in both the operational and training contexts at a 90% odor prevalence rate (e.g., 10% of the trials were blanks/distractors). The dogs then entered either an experimental or control training period. The dogs were randomly assigned either to the control or experimental training based on experimental and control assignments from the preceding study (DeChant et al., in press, Part II). These assignments were maintained for this study. This was done to ensure the control group did not have prior experimental manipulations. Risks of potential experimental carryovers were planned to be addressed in Experiment 2.

Training period
Following the baseline period, each dog received two daily sessions in the training context during the "training period." The dogs were not trained in the operational context during the training period. This was done to evaluate whether modifying training in the training context could lead to changes in performance in the operational context. The control group dogs ($n = 9$) were trained at a 90% odor prevalence rate for the entire training period (i.e., two sessions per day). This was to reflect a scenario where dogs are exposed only to high odor prevalence rates during training. The experimental group ($n = 9$) had the target odor prevalence rate reduced by 10% after two sessions with an accuracy ≥ 85%. A criterion of 85% accuracy was selected to be high and statistically well above chance but a readily achievable criterion. The objective of this procedure was to provide the experimental group with a training schedule where the odor prevalence rate was systematically reduced. The experimental group started training at a target odor prevalence of 80% (e.g., 32 out of 40 trials contained a target odor). After two consecutive sessions with accuracy greater than or equal to 85%, the target odor prevalence was reduced by 10% until the target odor prevalence reached 20%. If accuracy was below the 85% training criterion, the dog was trained on the prior target odor prevalence rate until criterion was reached and then the target odor prevalence was reduced. For example, if a dog was below 85% training criterion at 60% target odor prevalence, the dog would be trained again at 70% target odor prevalence until accuracy was at 85% or greater. The dog would then be trained at 10% progressively lower target odor prevalence. Two experimental dogs (Dallas and Pumpkin) needed additional days (Dallas 9 and Pumpkin 8) of training, so one control dog, Jax, was yoked to the number of additional training days (9 days).

Testing period
Following the training period, all dogs received one session in the operational context and one session in the training context per day for 5 days. The operational context target odor prevalence was 10% for both the control and experimental groups. In the training context, the target odor prevalence for the control group was maintained at 90% and maintained at the trained level of 20% for the experimental group.

Statistical analysis

All data were analyzed separately by the three different periods (baseline, training, test). Because no visual or statistically significant differences were observed in the dogs' performance during the baseline or training periods, below we only present the statistical analyses conducted during the test period, which was the most relevant and significant part of the experiment. To assess the effect of training to low odor prevalence rate on dogs' performance, a logistic generalized linear mixed-effect model was fit with an interaction of group (experimental vs. control) and context (operational vs. training), their main effects, a main effect of session number, and a random intercept for each dog. The primary dependent variable was the trial accuracy outcome (1 for correct and 0 for incorrect). An otherwise identical model was fit for the dependent variables of timeouts, hits, false alerts, and false all clears. To assess the effect of training to low odor prevalence rate on latency, an otherwise identical linear model was fit where the dependent variable was the log-transformed latency. Statistical significance of fixed effects was evaluated using the ANOVA function in the car package in R (Fox & Weisberg, 2018). Significant interactions were further analyzed with Tukey-adjusted post hoc tests from the lsmeans package (Lenth, 2016) to assess for differences between groups for each context. The lmer package (Kuznetsova et al., 2017) in R (R version 3.5.1, www.r.project.org; R Core Team, 2018) was used to fit models.

Results

Figure 1 shows each dog's accuracy (proportion of correct responses that includes hits and correct rejections) by

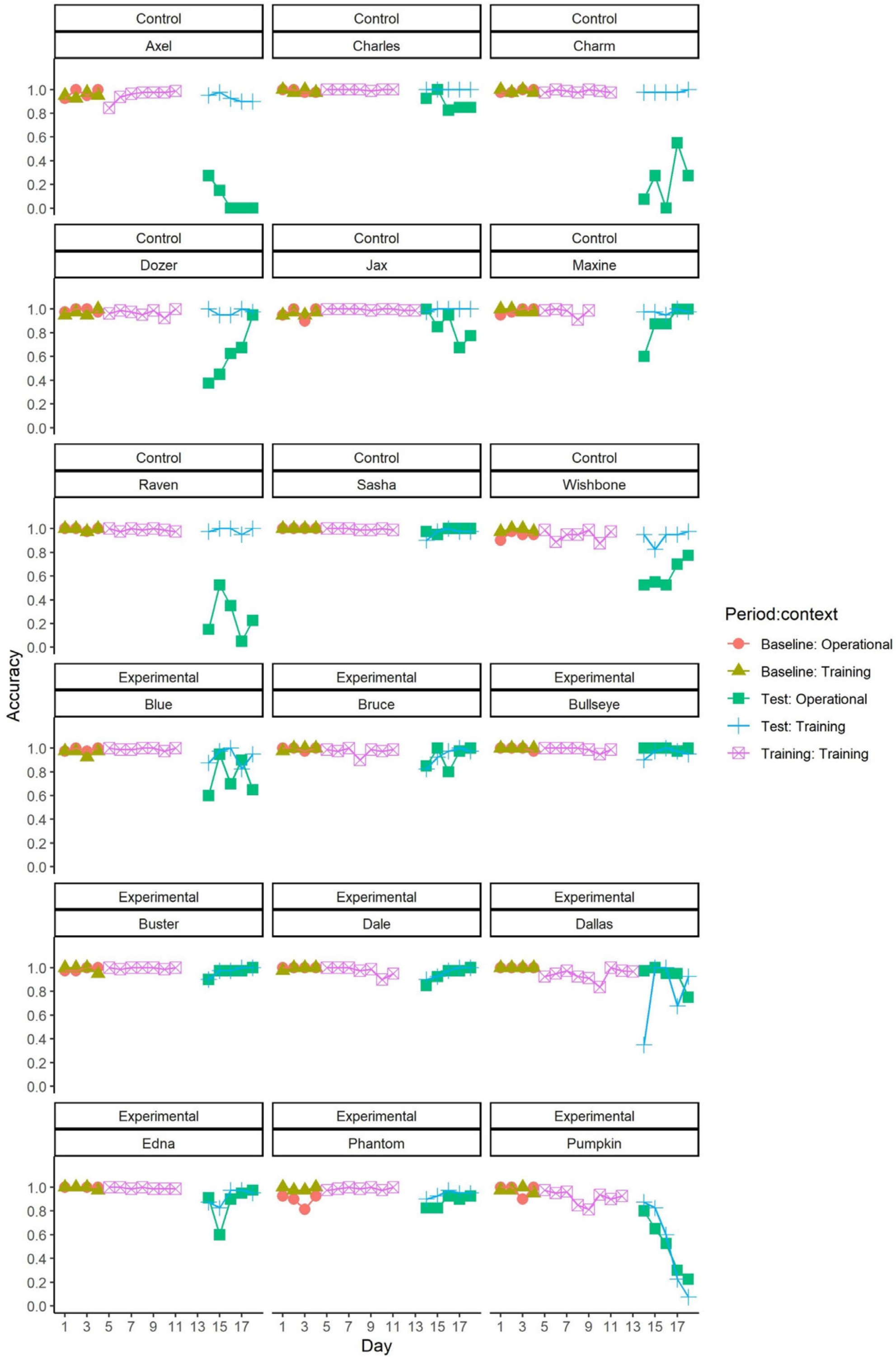

FIGURE 1 Individual dog accuracy (accuracy that includes hits and correct rejections) data for Experiment 1. Each point for the baseline and testing periods summarizes the proportion of correct trials out of 40 trials. For the training period, the day represents the average of two sessions (i.e., 80 trials) in the training context. Experimental dogs showed higher performance in the operational context, suggesting training to lower target odor prevalence had an effect on search decrement and performance. One dog, Pumpkin, was the only experimental dog that did not improve in accuracy during the test period in operational context. Three control dogs (Sasha, Maxine, Charles) only showed minor decrements in performance in the operational context.

TABLE 2 Mean ± *standard deviation of the* control and e*xperimental group overall performance during the* baseline, low training, and testing periods *during Experiment 1*

Period	Baseline				Low training		Test			
Context	Training		Operational		Training		Training		Operational	
Group	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental
Accuracy, %	98.05 ± 1.63	99.00 ± 1.20	97.91 ± 1.76	97.96 ± 4.26	97.88 ± 1.98	98.09 ± 1.86	97.05 ± 2.53	92.56 ± 5.73	60.05 ± 33.20	90.68 ± 7.34
False alert, %	00.69 ± 1.63	0.39 ± 0.74	0.69 ± 1.05	0.35 ± 0.523	0.76 ± 1.36	1.11 ± 1.51	1.77 ± 2.04	5.56 ± 5.79	4.61 ± 4.26	2.12 ± 2.58
Timeout, %	0.62 ± 0.69	0.39 ± 0.74	0.90 ± 1.36	1.45 ± 3.86	0.88 ± 1.01	0.682 ± 5.29	0.44 ± 0.68	1.68 ± 1.38	35.33 ± 30.49	7.12 ± 6.75
Hit, %	98.53 ± 1.27	99.14 ± 0.89	98.45 ± 1.46	97.95 ± 4.15	98.33 ± 1.62	99.47 ± 0.56	98.20 ± 1.63	97.62 ± 2.35	71.11 ± 32.86	98.12 ± 2.58
Correct rejection, %	93.75 ± 12.10	97.65 ± 4.65	93.05 ± 9.60	98.12 ± 5.30	93.78 ± 7.56	96.77 ± 3.57	86.66 ± 15.41	90.02 ± 8.84	58.82 ± 33.41	89.86 ± 8.06
False all clear, %	0.69 ± 0.92	0.214 ± 0.33	0.54 ± 0.77	0.25 ± 0.50	0.52 ± 0.60	0.22 ± 0.37	0.80 ± 0.68	0.77 ± 1.79	0.0 ± 0.0	0.62 ± 1.76
Latency, s	5.37 ± 1.73	4.63 ± 2.12	5.66 ± 1.89	5.25 ± 3.64	6.21 ± 2.08	4.65 ± 2.08	5.31 ± 1.95	4.61 ± 1.94	17.81 ± 11.47	6.08 ± 3.02

Note. Data do not include Pumpkin. Accuracy: Percentage of trials that a dog alerted to the target odor (i.e., hit) or called a correct all clear. False alert: Percentage of trials that a dog alerted to an incorrect port during a target odor or blank trial. Timeout: Percentage of trials that a dog did not alert or search all three odor ports within 45 s. Hit: Percentage of target odor containing trials that a dog alerted to the target odor. Correct rejection: Percentage of trials that did not contain the target odor that dogs did a correct all clear. False all clear: Percentage of trials dogs made an all-clear response on a target odor present trial. Latency: Time (in seconds) it took for the dog insert its nose into one odor port and start the search.

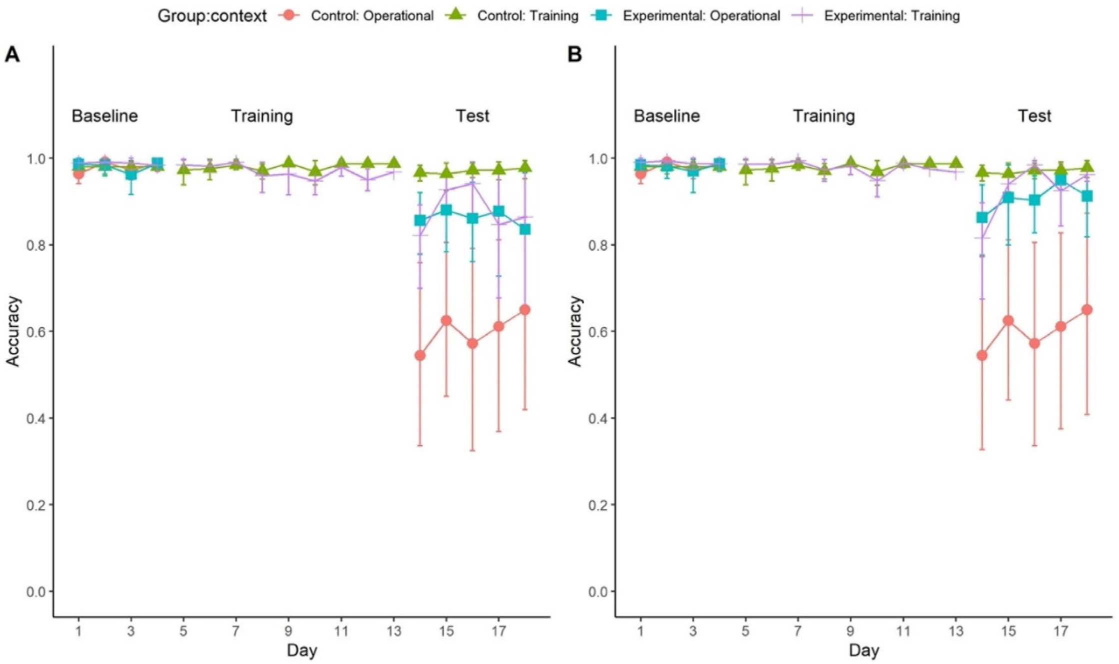

FIGURE 2 Panel **(A)**: Mean accuracy by group for Experiment 1 with all dogs included. Error bars represent 95% confidence interval of the proportion of trials that resulted in a correct response during Experiment 1. The experimental group maintained accuracy near the training criterion, except for the fifth day when it dropped, which was driven by one dog, Pumpkin, in the experimental group. Panel **(B)**: Mean accuracy by group for Experiment 1, excluding Pumpkin from the experimental group. Error bars represent 95% confidence interval of the proportion of trials that resulted in a correct response.

group (experimental and control) during the different periods of the experiment. Table 2 shows the mean percentage of dogs' accuracy, false alerts, timeouts, hits, correct rejections, false all clears and latency by group, context, and period.

Accuracy

During the testing period, the accuracy for the experimental group was higher (90.68%) than that for the control group (60.05%) in the operational context (see Table 2 and Figure 2A). Most dogs in the experimental group showed less of a decrement in the operational context compared with the training context. Interestingly one control dog, Sasha, maintained excellent performance in the operational context and two control dogs showed improvements across sessions (Dozer and Maxine). Nearly all experimental dogs, however, were able to maintain adequate performance in the operational context, with one clear exception (Pumpkin). Pumpkin showed a significant drop in performance in both the training and operational context during the test period, indicating global performance decrement (see Figure 1).

Statistical analysis showed no significant increase in performance across the testing sessions ($\chi^2 = 3.47$, $p = .06$). The interaction between group and context, however, was significant ($\chi^2 = 327.26$, $p < .001$), indicating that the dogs' performance between the training and operational context differed between the experimental and control group. Tukey-adjusted post hoc tests showed that the experimental group in the operational context had greater accuracy relative to the control group ($z = -2.56$, $p = .01$; experimental: 90.68% vs. control: 60.05%; see Table 2). However, the experimental group had lower accuracy ($z = 2.03$, $p = .002$) in the training context than did the control group (experimental: 92.56% vs. control: 97.05%).

However, the results for one dog, Pumpkin, were unique across the population in that he showed a complete drop in performance in both the operational and training context. This highlights that the experimental training may not be ideal for all dogs. However, because his poor performance was a singular outlier and a different trend from all other dogs (a decrement in performance for both training and operational contexts), we repeated the analysis without Pumpkin's data (see Figure 2B). All results were statistically identical except that without Pumpkin's substantial decline across sessions in both contexts, the main effect of session was now statistically significant ($\chi^2 = 40.28$, $p = < .001$). This indicates that accuracy overall increased with session number across all groups (except for Pumpkin). The remaining statistical analyses were therefore conducted excluding Pumpkin to avoid this dramatic decline to obscure trends among the other dogs. The data in Table 2 also exclude Pumpkin. Nonetheless, it is important to note that one of the nine experimental dogs showed complete extinction in both the training and operational context with the lowered prevalence of targets.

Latency

During the testing period, we assessed whether the treatment reduced latency to approach the panel for experimental dogs. In the operational context, experimental dogs showed the first nose poke within approximately 6 s, whereas control dogs' latency was up to 17 s. In the training context, both groups showed a latency around 5 s, which was close to the latency observed for the experimental dogs in the operational context (see Table 2). The mixed-effect statistical model indicated there was a main effect of context ($\chi^2 = 881.84$, $p < .001$), group ($\chi^2 = 3.88$, $p = .04$), session ($\chi^2 = 7.12$, $p < .001$), and the context-group interaction ($\chi^2 = 341.86$, $p < .001$). Post hoc tests for the interaction indicate that experimental group had a lower latency to search in the operational context compared with the control group ($z = 3.46$, $p < .001$). The search latency of both groups in the training context was not statistically different.

Timeouts

During the testing period, the timeout rate was higher for the control group (35.33%) compared with the experimental group (7.12%) in the operational context (see Table 2). The main effect of context ($\chi^2 = 189.39$, $p < .001$) and group ($\chi^2 = 4.28$, $p = .03$) and the context by group interaction ($\chi^2 = 84.94$, $p < .001$) on the probability that the dog made a timeout were statistically significant, but the main effect of session was not ($\chi^2 = 0.39$, $p = .53$). A post hoc test for the interaction indicated that the experimental group had a lower probability of a timeout (i.e., incomplete search; $z = 2.35$, $p = .02$) in the operational context compared with the control group. There was no difference between experimental and control dogs in the training context.

Hits

During the testing period, the proportion of odor trials that correctly ended in a hit (i.e., a correct alert) was higher for the experimental group (98.12%) compared with the control group (71.11%) in the operational context (see Table 2). The probability of a correct hit was predicted by a significant context (operational and training context) by group (experimental and control) interaction ($\chi^2 = 24.49$, $p < .001$) and by the main effect of context ($\chi^2 = 129.77$, $p < .001$). The hit rate of the experimental dogs in the operational context was ($z = -4.07$, $p < .001$) statistically higher than the hit rate of control dogs. The proportion of hits between groups in the training context was not statistically different ($z = -0.68$, $p = .49$). The effect of session ($\chi^2 = 3.20$, $p = .07$) and group ($\chi^2 = 1.50$, $p = .30$) were both nonsignificant.

Correct rejections

During the testing period, the correct-rejection rate was higher for the experimental group (89.86%) compared with the control group (58.82%) in the operational context (see Table 2). The frequency of correct rejections was predicted by the group (experimental and control) and context (operational and training) interaction ($\chi^2 = 57.68$, $p < .001$) and main effect of session ($\chi^2 = 39.69$, $p < .001$), where correct rejections increased across sessions. The data showed that experimental dogs had correct rejections similar to those of the control group in the training context ($z = 0.34$, $p = .64$), but the experimental dogs showed higher correct rejection than control dogs in the operational context ($z = 2.89$, $p < .01$).

False alerts

During the testing period, the false-alert frequency was slightly higher for the control group (4.61%) than for the experimental group (2.12%) in the operational context (see Table 2). The frequency of false alerts was predicted by a group (experimental and control) and context (operational and training) interaction ($\chi^2 = 49.78$, $p < .001$) and the main effect of session ($\chi^2 = 80.96$, $p < .001$), where false alerts decreased across sessions. The data showed that experimental dogs had a higher false-alert rate in the training context relative to the control ($z = -2.85$, $p = .004$). No statistically significant difference in the false-alert rate was observed in the operational context between groups. The number of times a dog made a false alert during the entire testing period (out of 7,236 times the dog made a response, i.e., not including timeouts) in both the training and operational context was 59 times for mineral oil, 38 times for limonene, 36 times for gloves, 49 times for cotton ball, and 69 times for blank jar.

False all clears

During the test period in the operational context, the control group had a 0.0% false-all-clear rate, whereas the experimental group had 0.62% false all clears (one false all clear out of 160 searches with target odor) during the testing period, showing no relevant difference in this metric (see Table 2). These data represent that both control and experimental dogs rarely made an all-clear response when a target odor was in fact present and the dog poked their nose into all three ports.

Discussion

These results from the laboratory model support that training dogs with systematically lower odor prevalence rates that more closely matched the simulated operational conditions led to moderately improved performance when challenged with the low prevalence rates in the operational condition. Unlike our prior attempts using noncontingent reinforcers or a Pavlovian conditioned stimulus to maintain search behavior (DeChant et al. in press), training dogs with systematically more intermittent schedules of odor prevalence did lead to improvements in performance that are noteworthy and important if translatable to deployed detection canines. This highlights the importance of dogs being conditioned to a high frequency of no-odor-present trials when there remain sufficient targets to maintain search.

High accuracy in the test condition after the low prevalence training was observed in eight of the nine experimental dogs. This suggests that the effect of the training is consistent and replicable among most of the dogs but that some dogs might have a lower tolerance to continuous training and testing under a low prevalence schedule. Experimental dogs had an increase in false alert rates during training context in the testing period, which suggests that some dogs may have started to show some response variability with lowered target frequencies. Such individual differences could be an interesting point for future investigation. If some dogs substantially tolerate more lean schedules of reinforcement, this could be a useful selection criterion for detection dogs that need to search with infrequent targets. Conversely, some dogs in the control group performed at high levels in the operational test. Given that these dogs participated in previous test sessions (e.g., Parts I and II), it is likely that the dogs may have had experience with intermittent odor frequency during prior tests, which was sufficient to prevent the performance decrement. Due to the randomization of when a target odor may appear, these dogs may have had sufficient training in prior studies with intermittent odor frequency to maintain performance. This is an unfortunate drawback to using the same participants across studies.

One interesting finding is that the false-all-clear rate was consistent throughout the study for all groups and conditions. This measure represents when the dog inserted their nose into the port that had the target odor and still called an all-clear response. This measure therefore largely reflects the odor occasioning of an alert, which remained largely unaffected by the odor prevalence rate. The performance decrement appears to be largely driven by reduction in search behavior (e.g., timeouts increasing).

There are several important limitations to the present study. First, to avoid potential confounding experimental effects, the dogs were not rerandomized into groups following prior experiences and they were not selected into group by performance (DeChant et al., in press). This leads to the potential that the effects observed may also relate to prior experiences. However, given that the previous manipulations had no effect on detection performance and that baseline performances were identical, this may have only been a minor limitation. Second, the present study used dogs that were purposely trained for this project, with only a few months of detection training. Operational dogs typically have years of prior training. Thus, it remains unclear whether the present results would translate to operational dogs.

To resolve some of the limitations of Experiment 1, Experiment 2 extended Experiment 1 with a larger cohort of operational dogs. Operational dogs were used because they regularly search contexts that have no or few targets present. The extension of Experiment 2 was to evaluate whether the effects of training with decreased frequency of target odors can lead to improvements in performance of certified operational dogs. However, several procedural modifications were necessary to make the study feasible for operational teams that had ongoing duties. The study timeline was condensed to 1 hr/day for 5 days, which necessitated procedural changes, which are noted in detail.

EXPERIMENTAL 2: EFFECT OF INTERMITTENT ODOR PREVALENCE TRAINING ON DETECTION PERFORMANCE OF OPERATIONAL DETECTION CANINES

Methods

Animals

Twenty-seven certified explosive detection dogs (12 Labrador Retrievers, six Belgian Malinois, two Dutch Shepherds, and seven German Shepherds) participated in this study. The dogs were from federal, county, city, and private organizations, and all had previously passed an independent certification exam. Two dogs (2a, 4b) were removed from the study due to illness not related to the study, one dog (8a) was removed due to failure to reach training criterion for the smokeless powder target within the allotted time, and one dog (13a) was removed due to the handler's personal emergency. Thus, 23 certified explosive detection dogs were used for analysis. The dogs' ages ranged from 3 to 8.5 years old, and all were handled by their regular handler who works the dog operationally. All procedures used in this experiment were reviewed and approved by TTU Institutional Animal Care and Use Committee (ACUC #19093–10).

Olfactometer training

The dogs were trained to search six individual Bluetooth automated olfactometers (see Figure 3). The number of

FIGURE 3 Olfactometer arrangement for Experiment 2. Six olfactometers were arranged in a semicircular pattern. Each olfactometer contained the target odor and five distractors. Handlers asked dogs to search the olfactometers, and when a dog made an alert (left image), the handler called the olfactometer indicated by the dog, which was then entered into the computer by the experimenter. The experimenter and handler were both blind to the location of the target odor. The computer would indicate correct responses with a "chirp" and incorrect responses with a buzzer. An all-clear response was marked with a simple end of trial tone.

olfactometers was increased from three to six to better match typically detection dog "line-up" procedures that we anticipated the dogs would be more familiar with. The olfactometers were arranged in a semicircular pattern approximately 0.25 m apart. Each olfactometer could present any of six different odors to the odor port. These odors included double base smokeless powder (obtained from agency bunker at facility) and five distractors (vinyl gloves, cotton balls, toothpaste, duct tape, and blank jar). The distractors were selected based on availability and are common items used as distractors for detection training. Each olfactometer was battery operated and controlled via a central computer connected via Bluetooth. The mechanism of odor generation and the algorithm to determine target odor prevalence, location of the target, and distractors presented were identical to those in Experiment 1. An experimenter operated the computer during the study.

At the start of every trial, a tone cued the handler to start a search. If a dog alerted to an olfactometer, the handler would call out the number of the olfactometer. The experimenter would then enter this number into the computer, and the computer would respond with a correct or incorrect tone or neutral end-of-trial tone. The neutral end-of-trial tone was used when the dog made an all-clear response to mimic typical operational conditions for an all-clear response. Thus, the experimenter and handler were blind to the location of the target odor. If the handler did not call a response (i.e., either alert or all clear) within 45 s, it was recorded as a timeout, which was considered an incorrect response.

The performance measures collected were identical to those of Experiment 1 (i.e., Table 1), except for latency. Latency was not collected because this was largely handler driven, not dog behavior (i.e., handlers decided when, after the tone, they would send their dog to search). One important deviation from Experiment 1 is that canine responses were dependent on their handler's interpretation of the dog's behavior rather than the computer-scored infrared beams. This was required because the operational dogs all had different types of alerts (e.g., sitting, laying down, staring etc.); thus, the handler did serve an additional and important role in this experiment.

Experimental design

Due to the extended time required for the dogs to complete Experiment 1, direct replication was not feasible for working with operational detection canine teams. The study was therefore modified to be completed within five working days with 1-hr sessions each day. To achieve this, we removed the within-subject comparison between contexts and minimized training time. Testing proceeded as follows: First, the dogs received a brief qualification session to confirm recognition and detection of the smokeless powder target and use of the six olfactometers followed by a pretraining test where the target odor frequency was reduced to 10%. This was completed on the first day of testing. Next, the dogs were randomly assigned to receive experimental or control training, which was restricted to 3 days. Last, the dogs received a posttraining test on the fifth day to reassess performance under infrequent target conditions.

Qualification

The initial qualification period consisted of one 10-trial session at 100% target odor prevalence. The qualification criterion was five consecutive correct trials or eight correct responses out of the 10 trials (e.g., accuracy ≥ 80%). The training criterion was reduced to 80% (compared with Experiment 1) due to the time constraint of 5 days for testing. The purpose of this qualification was simply to familiarize the dogs and handlers with the olfactometer and ensure appropriate odor detection for the target odor and that handlers understood how the olfactometers operated. Only the last five trials of the qualification period were retained for data analysis. The dogs progressed to the pretraining testing period after completion of the qualification period. Twenty-six of the 27 dogs completed the qualification period within 10 trials or less, indicating rapid transfer to the olfactometer boxes. The dog that did not meet qualification was not included. Throughout the study, three additional dogs (2a, 4b, 8a) discontinued participation due to unrelated conflicts from the study during the training period and were not included in the analyses. Thus, only 23 dogs are included in the analyses, figures, and tables.

Pretraining testing

Each dog received one 20-trial session for the pretraining testing period in which the frequency of a target odor appearing in one of the six boxes on a trial was lowered to 10% (e.g., 18 trials had distractor odors only and only on two trials did a box contain a target). Following the pretraining test, the dogs progressed to the training period and were randomly assigned into two training groups for the experimental design.

Training

Eleven dogs were randomly assigned to the control (a) group in which for the entire training period the target odor prevalence remained at 90%. The other 12 dogs were grouped into the experimental (b) group, where the target odor prevalence rate started at 90% and was systematically reduced by 10% after every session achieving an accuracy ≥ 80%. If a dog met the accuracy criterion in all sessions, the dog would reach a target prevalence rate of 20% prior to posttraining testing. If the dog did not reach accuracy of 80% or greater, it had an additional session at that odor prevalence until it met or exceeded the 80% accuracy criterion. The objective was to train the experimental group with a lower odor prevalence rate, whereas the control group received matched training sessions at a consistent target odor prevalence.

The training period was a total of 3 days with 2–3 sessions per day. The time between each session was at least 5 min to allow the dog to get water and rest for a longer period if needed. The timeline for this study was limited as such based on canine availability from operational duties and thus needed to be fit within these parameters. Ten experimental dogs (1b, 3b, 5b, 6b, 7b, 9b, 10b, 11b, 12b, and 13b) completed the training to 20% target prevalence, and the remaining two dogs (2b and 14b) reached 30% target odor prevalence during the training period. Dog 2b reached accuracy of 80% or greater at 30% target prevalence. Dog 14b did not reach 80% or greater accuracy at 30% target odor prevalence, and we unfortunately were limited in time to continue training and therefore progressed to the posttraining test.

Posttraining testing

All dogs received a total of three 20-trial sessions for the posttraining testing period. The first session was at 10% target odor prevalence, identical to the pretraining test. The second session was at 5% target odor prevalence, and the final session was at 0% target odor prevalence. The target prevalence was lowered in the last two sessions to provide an additional challenge for operational dogs that have been completing search tasks for an extensive period.

Video coding

An ethogram based on Part I (see Table 1 in Aviles-Rosa et al., in press) was modified to video code canine search

TABLE 3 Ethogram of the behaviors that were coded for each dog during Experiment 2

Behavior description	Behavior coded	Agreement
Search behavior	0. No search (dog did not search any olfactometers) 1. Drive-by search (dog hovered but did not fully insert nose) 2. Incomplete search (dog missed an olfactometer but trial completed) 3. Thorough search (dog searched all 6 olfactometers and/or made an alert)	Values other than "3" too infrequent
General behaviors	1. Paw at olfactometer 2. Vocalization during search 3. On leash or off leash 4. Handler asked dog to search one olfactometer again	1. Values too infrequent 2. Values too infrequent 3. ICC = 0.97 4. ICC = 0.53
Tail position	1. Wagging 2. Relaxed 3. Tucked 4. Aroused/Raised	% Agreement 96.70
Search pattern	1. Linear 2. Random 3. Skipped an olfactometer	% Agreement 91.03

behavior during the pretraining test and posttraining test sessions. In total, there were 92 videos, where 30% of the videos were double coded. The interobserver agreement is shown in Table 3. To permit analysis of the scored behaviors, a retention criterion was implemented such that a behavior needed to be observed at least 10 times (or 10 times for each level of a factor) out of the 1,840 coded trials (>0.5% frequency). In addition, the interclass correlation (ICC) for interobserver agreement needed to be >0.4, reflecting adequate agreement. Percentage of agreement was used to assess interobserver agreement for tail position and search pattern because of the infrequency of some behaviors (e.g., tail tucked and incomplete search, no search, and drive-by search). The video-coded behaviors were combined with accuracy derived from the olfactometers to identify whether any specific video-coded behaviors were associated with performance.

Statistical analysis

The analyses evaluated the effect of training assignment (control or experimental) on posttraining testing performance. To assess the effect of the experimental progressively lowered odor prevalence training on accuracy, a logistic generalized linear mixed-effect model was fit in which an interaction of group (experimental vs. control) and odor prevalence rate (10, 5, and 0%) was assessed. The dependent variable was the trial outcome (1 for correct and 0 for incorrect). This was also predicted by a random intercept for each dog. An otherwise identical fixed and mixed-effect model was fit to predict false alerts, timeouts, false all clears, and correct rejections. The lmer package (Kuznetsova et al., 2017) in R (R version 3.5.1, www.r.project.org; R Core Team, 2018) was used to fit models. For the 30% double-coded video behaviors, the ICC package in R was used for interobserver agreement. To assess whether any video coded behaviors predicted overall accuracy, a logistic generalized linear mixed-effect model in which accuracy (correct or incorrect response) was predicted by the behavioral measures that meet the retention criteria with a random effect of dog.

Results

Figure 4 shows individual dog performance by group and period. Figure 5 shows dog performance by group (experimental and control). Dogs in the experimental group had greater performance during the posttraining test than did control dogs, but otherwise both groups had identical performance in the qualification, pretraining test, and training periods. Table 4 shows the mean and standard deviation of accuracy, false alerts, timeouts, hits, and correct rejections, by group and period.

Accuracy

Figure 4 shows the accuracy by individual dog. Although some control dogs performed similarly to the experimental dogs, overall, the experimental group had greater accuracy (91.55%) in the posttraining test period compared with the control group (82.36%; see Figure 5 and Table 4). The interaction between group and odor prevalence was not statistically significant ($\chi^2 = 1.71$, df $= 2$, $p = .55$), indicating that the group difference was consistent for each low odor prevalence tested (10, 5, and 0%). The main effect of group on dogs' accuracy was significant ($\chi^2 = 4.08$, df $= 1$, $p = .04$). Thus, dogs in the experimental group that received the experimental training, overall, showed higher detection accuracy during all sessions of the posttest in comparison with the control group. The main effect of odor prevalence rate was not statistically significant ($\chi^2 = 4.89$, df $= 2$, $p = .086$).

False alerts

During the posttraining test, the control group had more false alerts (17.48%) compared with the experimental group (8.02%; see Table 4). The interaction between group and odor prevalence was not significant ($\chi^2 = 1.16$, df $= 2$, $p = .55$) for false-alert rate. The main effect of group on dogs' false-alert rate was statistically significant ($\chi^2 = 4.55$, df $= 1$, $p = .032$), and the main effect of odor prevalence rate was not significant ($\chi^2 = 4.29$, df $= 1$, $p = .11$). Thus, dogs in the experimental group showed fewer false alerts than did those in the control group.

Timeouts

During the posttraining test, the experimental group had a slightly higher timeout rate (0.13%) compared with the control group (0%), which is attributed to one dog stopping in the middle of searching to get water and running out of time (see Table 4). The interaction between group and odor prevalence was not significant ($\chi^2 = 0.0003$, $p = .98$) for the timeout rate during posttraining test. The main effect of group on dogs' timeout rate was not statistically significant ($\chi^2 = 0.0038$, $p = .095$), and the main effect of odor prevalence rate was not significant ($\chi^2 = 0.000$, $p = .099$). Therefore, the timeout rate remained similar for the two groups during posttraining test.

False all clears

During the posttraining test, the experimental group had slightly higher false-all-clear rates (5.55%) compared with the control group (3.03%; see Table 4). The interaction between group and odor prevalence was not significant ($\chi^2 = 0.0024$, df $= 1$, $p = .96$) for the false-all-clear rate.

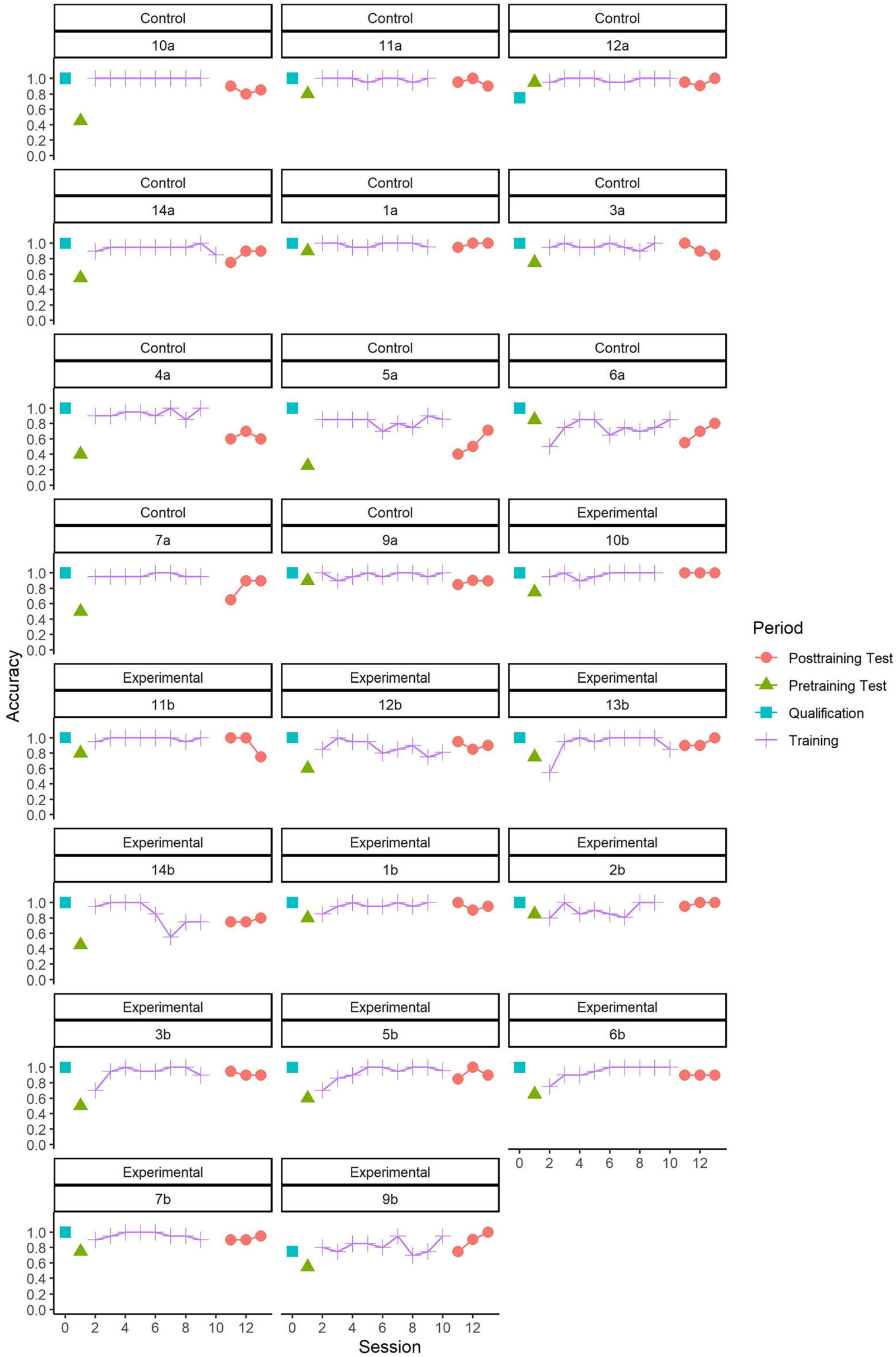

FIGURE 4 Individual dog accuracy (which includes hits and correct rejections) data for Experiment 2. The point for the qualification period summarizes the last five trials. Each point for the pretraining test, training, and posttraining test periods summarizes 20 trials for each session

The main effect of group on the dogs' false-all-clear rate was not statistically significant ($\chi^2 = 0.0039$, df $= 1$, $p = .949$), and the main effect of odor prevalence rate was not significant ($\chi^2 = 254$, df $= 1$, $p = .613$). Thus, the false-all-clear rate remained similar for the two groups during the posttraining test.

Correct rejections

During the posttraining test, the experimental group had higher correct-rejection rate (91.66%) compared with the control group (81.76%; see Table 4). The interaction between group and odor prevalence was not significant

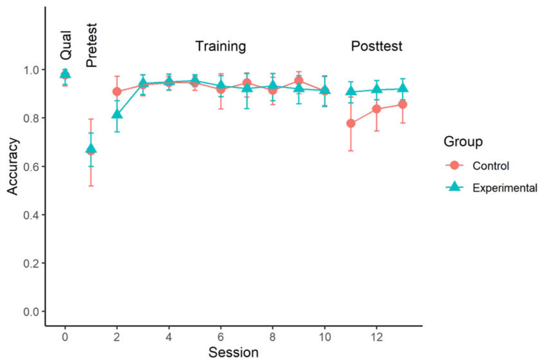

FIGURE 5 Overall accuracy. The average accuracy for both experimental and control group during the qualification period (Qual), pretraining test, training, and the posttraining test. Overall, the experimental group had greater performance during the posttraining test when compared with the control group. Error bars represent 95% confidence interval of the percentage of odor trials that resulted in a correct response.

(χ^2 = 1.50, df = 2, p = .47) for correct-rejection rate during the posttraining test. The main effect of group on the dogs' correct rejection rate was statistically significant (χ^2 = 4.02, df = 1, p = .044), and the main effect of odor prevalence rate was significant (χ^2 = 6.18, df = 2, p = .045) for correct-rejection rate. The experimental group had a higher correct-rejection rate compared with the control group.

Behavior video coding

Table 3 shows the behaviors excluded due to infrequency. A thorough search was scored for nearly all trials. This is an interesting deviation from our prior work that showed substantial decreases in the search score. Interestingly, however, an unaccounted variable in this measure is the frequency with which handers re-sent a dog to reinvestigate one olfactometer. This was scored 832 times over 174 trials, indicating that perhaps handlers are noticing reduced search and are re-sending dogs to missed olfactometers, which is causing all trials to end with a complete search. Overall, the number of times a handler asked a dog to reinvestigate the olfactometers was 434 times for the control group and 398 times for the experimental group.

The retained behaviors were therefore tail position, whether the dog was worked on or off leash, and whether the handler asked the dog to re-search a box. None of the retained behaviors statistically predicted trial accuracy (tail position: p = .58, on leash: p = .59, handler asked dog to search again: p = .41).

Discussion

Overall, the results of Experiment 2 help confirm the results of Experiment 1. The magnitude of the effects was a little lower, which is not unexpected given that the operational dogs have substantially longer histories working in varied environments and training practices. One limitation for Experiment 2 was the shortened time available to use the operational dogs. The dogs were needed for their regular operational work (searching venues). It was, therefore, not feasible to directly replicate Experiment 1. Another limitation for Experiment 2 was our decision to not include all-clear signals during the initial qualification period. We did not include this due to the time constraint, so the qualification was designed to be as short as possible to simply demonstrate that the dog could detect SP and use the olfactometers to search without issue. Nonetheless, in Experiment 2, the confirmation that the proposed training for increased intermittency of odor frequency led to improvements of already-operational dogs highlights the potential importance of this training consideration for detection canines.

GENERAL DISCUSSION

The effects of training dogs under progressively lower prevalence rates on detection accuracy were evaluated in this study under laboratory and applied conditions. The low-prevalence training improved the dogs' performance under both laboratory conditions and applied conditions with operational detection canines. Of all the mitigation strategies we evaluated (DeChant et al., in press, Part II), training dogs under a low-prevalence schedule seems to be the most effective method to maintain search performance with infrequent targets.

Our findings are in agreement with those of Thrailkill et al. (2016) who found that a simulated search behavior under a lean schedule of reinforcement was more resistant to extinction compared with the same behavior trained under a rich schedule of reinforcement. The training schedule used for the experimental groups placed search behaviors under an intermittent schedule of reinforcement, making it more resistant to lean schedules of odor appearance (e.g., partial reinforcement effect). Because within the experimental framework, poor performance (e.g., increase in false-alert and timeout rates) is driven by extinction of search behavior and not the odor-alert-reinforcer contingency, making search more resistant to extinction has important consequential effects on detection performance.

The average performance of the control group also improved in both Experiments 1 and 2. This suggests that in addition to systematically lowering the target odor schedule, the dogs may have adjusted to the intermittent reinforcement schedule over time. However, further research is needed to observe the performance for the control dogs under a similar scenario with an increase in sessions, as their performance may reach similar accuracy compared with the experimental dogs with more testing sessions.

TABLE 4 Mean ± standard deviation of the control and experimental group during the training, pretraining testing, and posttraining testing periods

	Pretraining test		Training		Posttraining test	
	Control[1]	Experimental[1]	Control[1]	Experimental[1]	Control[1]	Experimental[1]
Accuracy	66.36 ± 24.09	67.08 ± 13.04	90.11 ± 8.78	92.04 ± 4.73	82.36 ± 14.39	91.55 ± 5.78
False alert	30.90 ± 23.00	26.66 ± 13.20	5.96 ± 6.008	6.00 ± 4.91	17.48 ± 14.39	8.02 ± 5.84
Timeout	2.72 ± 4.67	6.25 ± 14.16	0.21 ± 0.37	0.28 ± 0.34	0.00 ± 0.00	0.13 ± 0.48
False-all-clear rate	0.00	0.00	4.99 ± 5.72	3.87 ± 2.70	3.03 ± 10.05	5.55 ± 12.97
Correct rejection	63.63 ± 25.01	64.35 ± 13.07	84.34 ± 13.90	90.78 ± 6.96	81.76 ± 15.35	91.66 ± 5.73

Note. [1]Mean (*SD*).

Interestingly, one noticeable difference between Experiment 1 and Experiment 2 was that operational dogs' response to infrequent targets was to engage in false alerts. This is expected extinction-induced response variability but differs from the increased number of timeouts we observed in Experiment 1. In Experiment 1, we observed only a minor increase in false alerts within the laboratory conditions, but we observed more timeouts and failure to search as the response to extinction of search behavior. These differences likely reflect that (1) laboratory dogs were previously naïve to detection work, with only 1 month of training, whereas operational dogs had years of training and experience and (2) operational dogs were largely tested on leash and walked by the handler. Laboratory dogs were worked off leash and were free to engage in other behaviors, but leashed dogs may be more likely to engage in search when directed on leash by the handler. This is further indicated by the fact that handlers re-sent their dogs to re-search or better search boxes on 174 trials. Perhaps without handler intervention, these would have been the timeouts observed in the laboratory conditions with minimal experimenter intervention. A future study observing the influence of handler intervention and detailed search would be interesting to conduct with operational dogs.

An additional important consideration between the laboratory and operational period is that substantially more data could be collected in the laboratory period. For Experiment 1, dogs were tested in low-prevalence conditions over 5-day periods, which is five times longer than what was done in Experiment 2. The shortened timeline for Experiment 2 was necessary because this would have required operational dogs to not be available to agencies for well over 2 weeks, which simply would not fit with day-to-day operational needs.

Interestingly, due to some of these differences, the behavioral data of the dogs became less informative in predicting accuracy than was observed in the laboratory condition in Part I (see Aviles-Rosa et al., in press). This seems to be driven primarily by the fact that handlers were successful in intervening when dog search was showing signs of decline and re-sent dogs to the boxes repeatedly. This made nearly all trials have a final search score of "thorough search." Interestingly, however, for the five trials in which the search score was not rated as thorough, overall accuracy was 40% (two misses, one false alert, and two correct all clears). Although not subject to statistical analysis, it does again highlight the importance of visually inspecting for thorough search behavior, reflecting the highest miss rate in the study that was observed when the search score was not thorough.

It appears there was not a clear predictor of an increased probability of a false alert. It is important to note, however, that the behavioral-coding agreement was on the borderline of acceptable. This was largely driven by the infrequency of the behaviors in the coded videos, making one or two disagreements have substantial effects on agreement scores. Nonetheless, agreement when a behavior did not occur was very high, indicating that the borderline agreement scores likely did not substantially affect the outcomes of this analysis.

Interestingly, across both experiments, the relation between the probability of an alert given the dog sampled the target odor box did not change. If the dogs encountered the target odor (sampled the box with the target), detection was very high (e.g., probability of a miss was low: false all clear). This suggests that although the first link of the behavior chain may have been undergoing extinction (search behavior), the second link was not interrupted (alert behavior). Rather, the disruption of the first link (search behavior) led to disruption in accessing the odor for alerting/detection, but the second link showed minimal disruption.

These results highlight that the intermittency of odor prevalence is a very important and relevant variable that requires attention for detection dogs. Systematically training dogs to a schedule of odor prevalence that is similar to operational conditions is important. However, this manipulation can likely only be used in conjunction with other efforts to reduce extinction of search behavior. Regardless of the intermittency training, extinction would occur in situations in which targets (e.g., explosives) are never encountered. Thus, the combination of intermittency training for odor prevalence as well as use of innocuous odors (i.e., Porritt et al., 2015) may be highly compatible approaches that will maintain search behavior for few target odors that can be feasible to maintain in operational conditions.

In conclusion, the present experiments showed that under infrequent target odor conditions, search behavior was disrupted for dogs in the control group, leading to poorer performance. Training experimental dogs through an incremental training program with increasingly more infrequent target odors led to improvement in overall performance and accuracy in Experiments 1 and 2. This highlights that the partial reinforcement effect can be used to increase resistance to extinction of search behaviors for working detection 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: 70RSAT20 CB0000010. 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 STATEMENT

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 (IACUC #19093–10).

ORCID

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

REFERENCES

Abel, R. J., Lunder, J. L., & Harynuk, J. J. (2020). A novel protocol for producing low-abundance targets to characterize the sensitivity limits of ignitable liquid detection canines. *Forensic Chemistry*, *18*, Article 100230. https://doi.org/10.1016/j.forc.2020.100230

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

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.*

Chan, C. K. J., & Harris, J. A. (2019). The partial reinforcement extinction effect: The proportion of trials reinforced during conditioning predicts the number of trials to extinction. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*(1), 43–58. https://doi.org/10.1037/xan0000190

DeChant, M. T., Ford, C., & Hall, N. J. (2020). Effect of handler knowledge of the detection task on canine search behavior and performance. *Frontiers in Veterinary Science*, *7*, Article 250. https://doi.org/10.3389/fvets.2020.00250

DeChant, M. T., Aviles-Rosa, E., Prada-Tiedemann, P. A., & Hall, N. J. (in press). A laboratory model of canine search vigilance decrement, II: Noncontingent reward and Pavlovian appetitive stimuli. *Journal of Experimental Analysis of Behavior*

Fox, J., & Weisberg, S. (2018). *An R companion to applied regression*. SAGE.

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

Harris, J. A., Kwok, D. W. S., & Gottlieb, D. A. (2019). The partial reinforcement extinction effect depends on learning about non-reinforced trials rather than reinforcement rate. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*(4), 485–501. https://doi.org/10.1037/xan0000220

Haselgrove, M., Aydin, A., & Pearce, J. M. (2004). A partial reinforcement extinction effect despite equal rates of reinforcement during Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*(3), 240–250. https://doi.org/10.1037/0097-7403.30.3.240

Jezierski, T., Adamkiewicz, E., Walczak, M., Sobczyńska, M., Górecka-Bruzda, A., Ensminger, J., & Papet, E. (2014). Efficacy of drug detection by fully-trained police dogs varies by breed, training level, type of drug and search environment. *Forensic Science International*, *237*, 112–118. https://doi.org/10.1016/j.forsciint.2014.01.013

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

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

Lewis, D. J. (1960). Partial reinforcement: A selective review of the literature since 1950. *Psychological Bulletin*, *57*(1), 1–28. https://doi.org/10.1037/h0040963

MacLean, E. L., & Hare, B. (2018). Enhanced selection of assistance and explosive detection dogs using cognitive measures. *Frontiers in Veterinary Science*, *5*, Article 236. https://doi.org/10.3389/fvets.2018.00236

Marks, A. (2002). Drug detection dogs and the growth of olfactory surveillance: Beyond the rule of law? *Surveillance & Society*, *4*(3), Article 3450. https://doi.org/10.24908/ss.v4i3.3450

Nevin, J. A. (1988). Behavioral momentum and the partial reinforcement effect. *Psychological Bulletin*, *103*(1), 44–56. https://doi.org/10.1037/0033-2909.103.1.44

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

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

How to cite this article: DeChant, M. T., Aviles-Rosa, E., Prada-Tiedemann, P., & Hall, N. J. (2023). Part, III: Increasing odor detection performance after training with progressively leaner schedules of odor prevalence. *Journal of the Experimental Analysis of Behavior*, 1–16. https://doi.org/10.1002/jeab.841