Received: 1 August 2022 | Accepted: 13 January 2023

DOI: 10.1002/jeab.832

RESEARCH ARTICLE

Journal of the Experimental Analysis of Behavior

SEAB

A laboratory model of canine search vigilance decrement, I

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

[1]Canine Olfaction Research and Education Laboratory, Department of Animal and Food Sciences, Box 42141, Texas Tech University, Lubbock, TX, United States

[2]Forensic analytical chemistry and odor profiling laboratory, Department of Environmental Toxicology, Box 1163, Texas Tech University, Lubbock, TX, United States

Correspondence
Nathaniel J. Hall, Canine Olfaction Research and Education Laboratory, Department of Animal and Food Sciences, Box 42141, Texas Tech University, Lubbock, TX 79409-214, USA.
Email: nathaniel.j.hall@ttu.edu

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

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

Abstract

Previous studies have found that infrequent targets can reduce dogs' vigilance. The purpose of this study was to develop a laboratory model to evaluate the effects of infrequent targets on dogs' search behavior and performance. Dogs ($n = 18$) were trained to detect smokeless powder in an automated olfactometer in two distinct rooms ("operational" and "training"). During baseline, the dogs received five daily sessions at a high target odor frequency (90%) in both rooms. Subsequently, the frequency of the target odor was decreased to 10% only in the "operational" room but remained at 90% in the training room. Last, the odor prevalence was returned to 90% in both rooms. All dogs showed a significant decrement in detection performance in the operational room when the target odor frequency was decreased but simultaneusly mantained high performance in the training room. This decrement was largely due to decreases in adequate search behavior. All dogs recovered performance when the odor frequency was increased again to 90%. Trial accuracy was associated with tail position, search score, latency, and duration of environmentally directed behaviors. The data show that low target odor prevalence significantly reduced search behavior and performance and that there are behaviors that can be used by handlers to assess their dog's search state.

KEYWORDS

behavior chain, detection dog, learned inattention, olfaction, search behaviors, vigilance decrement

Detection dogs have been trained to find a wide variety of hidden and cryptic targets such as invasive insects and weeds (Aviles-Rosa, Nita, et al., 2022; Hanigan & Smith, 2014), narcotics (Furton et al., 2002; Jantorno et al., 2020), explosives (Aviles-Rosa, McGuinness, et al., 2021; Lazarowski et al., 2021; Lazarowski & Dorman, 2014) and even missing persons (Jinn et al., 2020). In doing so, dogs are tasked to search a given environment to find and respond to their trained target odor, perhaps not unlike animals that must search for prey and productive areas to forage. Importantly, operational detection dogs are tasked to perform extensive searches often in areas with very low target densities, potentially leading to near extinction conditions for search behavior (e.g., dogs rarely find a target when deployed). Existing literature suggests that the low probability of occurrence of an event (e.g., finding a target odor) can induce a reduction in dogs' vigilance state (Gazit et al., 2005; Porritt et al., 2015). Vigilance refers to a state of readiness or alertness to detect and respond to a stimulus (Ballard, 1996; Hancock, 2013), and sustained attention refers to the ability of an individual to maintain a vigilance state for an extended period (Ballard, 1996). In humans, the ability of an individual to remain vigilant during a repetitive detection task is affected by different factors such as the difficulty of the task, the type of discrimination test, duration of the task, and event rate (Freeman et al., 2004; Hancock, 2013; Mackworht, 1948; Matthews et al., 1993; See et al., 1995). The inability to maintain vigilance during a task will then result in a reduction in performance characterized by an increase in stimulus misses and a reduction in the detection or hit rate (Ballard, 1996). Although it varies with different factors, in humans, a decrement in performance has been observed after 15–30 minutes on a repetitive task (Mackworht, 1948).

The effect of different factors on individuals' ability to maintain proficient detection during repetitive tasks has been extensively studied in humans during the past decades (Adams, 1987; Helton & Russell, 2011; Mackworht, 1948; See et al., 1995). However, little is known for detection dogs. Detection dogs perform

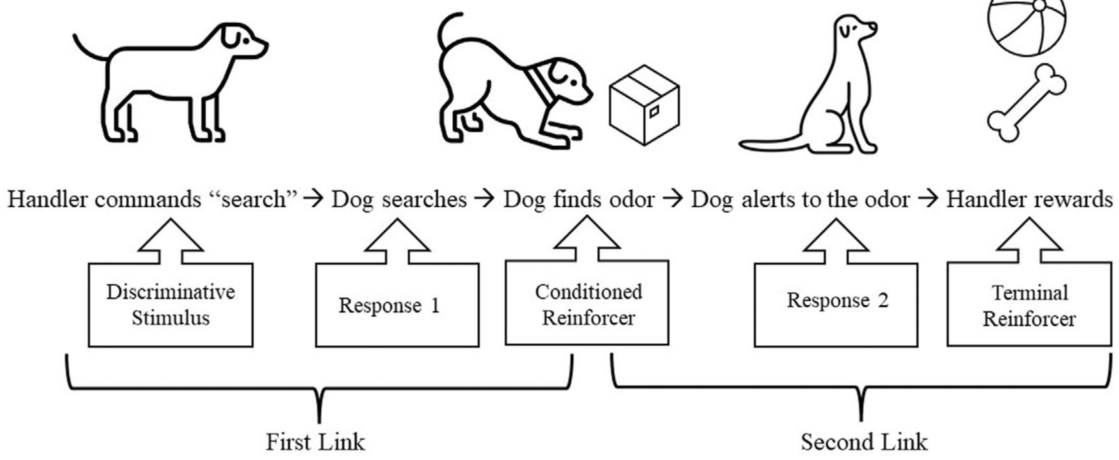

FIGURE 1 An illustration of the detection task visualized as a two-part behavior chain. The chain starts when the handler commands the dog to search. This serves as a discriminative stimulus that initiates the first link of the chain. The dog responds to the command by searching for the target odor (Response 1). The finding of the target odor serves as a conditioned reinforcer for the search behavior and as a discriminative stimulus that occasions the trained behavioral response (Response 2). The chain finalizes when the handler delivers a toy or food as the terminal reinforcer. Adapted from DeChant (2021) with permission of the author.

repetitive searches to areas, vehicles, luggage, and even persons for about 45 min or more at a time. More importantly, however, the probability that a detection dog encounters a target odor (e.g., explosive, narcotics, invasive species, etc.) during a search is usually low. The existing literature suggests that the long searches together with low occurrence of an event (e.g., the finding of a target odor) during a search could result in performance decrements (Gazit et al., 2005; Porritt et al., 2015). The goal of this experiment was to develop a laboratory model that mimics how detection dogs are often trained and deployed in varying places, with differing target odor densities, to evaluate how dogs' detection performance changes with target odor frequency. The development of such a model would be useful for evaluating methods to maintain performance when target density is low for operational searches.

One conceptualization of the canine detection task is that of a two-part behavior chain (Thrailkill et al., 2016; Thrailkill et al., 2018). In the first link of the chain, a dog is cued to "search" by the handler's command or the context itself (see Figure 1). The dog then engages in searching behaviors, which leads to encountering the target odor. The target odor serves as a conditioned reinforcer for the search behavior and as a discriminative stimulus that occasions the trained final alert response such as a sit, nose hold, or lie down (Porritt et al., 2015; Thrailkill et al., 2016; Thrailkill et al., 2018). The final alert is then reinforced with a primary or terminal reinforcer such as a toy or food (Figure 1). Thus, when the probability of encountering the target odor is very low, search behavior may undergo extinction because it is rarely reinforced by the appearance of the target odor. For this reason, a decrement or complete extinction of dog search behavior is to be expected in places with a low probability of finding a target odor. Importantly, this reduction or extinction in search behavior should not affect the relationship between the target odor occasioning a final trained response (i.e., the second link), as this contingency is still reinforced by the terminal reinforcer. When evaluating detection dog performance from a behavioral-chain perspective in places where the probability of encountering a target odor is low, we hypothesize that detection dogs may have poor detection performance due to a possible decrement or extinction in search behavior, with minimal effects on the second link, alerting when encountering the target odor.

Recent studies suggest that there can be a detection decrement in operational detection dogs when the presence of a target is infrequent or absent. For instance, Gazit et al. (2005) tested dogs in two similar paths. One path contained five targets (Path A) and the other path (Path B) contained no targets. They found that the dogs were able to discriminate between both paths, and a significant reduction in search-related behaviors was observed within the first session in Path B (Gazit et al., 2005). Reduction in search behavior was noticed as an increase in the amount of time that the dogs spent walking slowly instead of trotting or running the path. The dogs' detection performance and search behavior remained high in Path A, even when their search behavior was significantly lower in Path B (the path with no target). This indicated that the dogs were able to discriminate between paths and that they performed and behaved differently in the path that they learned that contained no targets in comparison with the path with which they had a prior reinforcement history with targets.

The same study found that when dogs continuously searched a path with no explosive (Path B), they later had difficulty detecting an explosive when present in the same path (Gazit et al., 2005). These results suggest that dogs can discriminate two visually similar contexts and that search behavior and performance may decrease specifically in the context or place associated with a low target odor frequency while simultaneously being high in a context with a higher target odor frequency.

Similar results were also found by Porrit et al. (2015). In their study they had three experimental groups of dogs. One group of dogs searched an area 36 times and never found a target odor, whereas the other two groups always found either an explosive or a nonexplosive target odor in the same area (Porritt et al., 2015). Similar to Gazit et al. (2005), they reported that within the first session, dogs in the experimental group with zero finds searched the area less thoroughly compared with the other two groups of dogs that always found a target. Immediately after the completion of the 36 searches, explosives were planted in the same area to evaluate how the treatment affected the dogs' capability of finding an explosive when it was presented in the same area. Half of the dogs that had previously searched the area with no targets failed to detect the explosives when placed in the same area, but no difference was found between dogs that had experience finding explosives or an unrelated odor in the same area. Dogs with zero prior finds needed six searches where explosives were planted before their performance was similar to the performance of the other groups (Porritt et al., 2015). The study also found a significant correlation between search behavior and detection rate, where dogs that performed more thorough searches had better detection rates (Porritt et al., 2015). The findings of both studies suggest that continuous testing or training with low target odor prevalence decreases search behaviors and that this decrement in search behavior has negative consequences for dogs' detection performance. The search decrement observed in the search areas with infrequent target odors is of importance for detection dogs, particularly to explosive detection dogs due to the way they are traditionally trained. For instance, detection dogs are frequently trained in areas where the target odors are presented at a high rate, such as a training facility. A previous survey found that most handlers do one or zero blank runs during training (DeChant et al., 2020). However, in operational scenarios, dogs will rarely find an explosive. Thus, an explosive detection dog could show high performance and search during training scenarios, but this may not be representative of their performance in the operational scenarios where the probability of finding a target odor is low.

Altogether the existing literature suggests an important performance effect on dogs when they are working in areas with infrequent targets that warrants further investigation. However, studying this phenomenon in operational settings can be expensive and challenging. Thus, the purpose of this study was to develop a laboratory model that replicates detection dog working conditions and that could induce the expected and observed search and performance decrement described in previous studies. To do this, dogs were tested in two different rooms: one with a high target odor prevalence (training room) and one with a low target odor prevalence (operational room) in an ABA reversal design. The goal was to model how detection dogs are frequently trained in one area or location at a high target rate but are deployed in another area or location with a low target rate. Having a laboratory model with which to study decrements in search behaviors and detection performance will allow us to further understand this phenomenon in a controlled setting and use this model to evaluate different procedures that could mitigate search behaviors and performance decrements in places with low target odor densities.

MATERIALS AND METHODS

Subjects

Eighteen mixed-breed dogs participated in this study (five spayed females and 13 neutered males). These 18 dogs were tested in three separate cohorts of six dogs. Participants were selected from local shelters and rescue organizations with the goal of providing additional training for adoption. The dogs were selected based on their food motivation, size (20–30 kg), age (<8 years), and boldness (e.g., were not afraid and approached the experimenter and the odor port during the selection trials). The dogs received 25% of their daily food ration in the morning (~ 0800) and the remaining in the afternoon (~ 1600) after training or testing. The dogs had free access to water in their kennels and during the training or testing session. In addition, the dogs received two walks or play sessions daily and received additional training for adoption ("sit," "down," "stay," loose leash walking, cooperative care for husbandry procedures (e.g., nail trim), and counterconditioning for any fear-related behaviors).

Apparatus

All training and testing were conducted at the Texas Tech University Canine Olfaction Research and Education (CORE) lab. All procedures and animal handling were approved by the Texas Tech University Institutional Animal Care and Use committee (protocol # 19093–10). The dogs were trained to use the automated olfactometer described and validated by Aviles-Rosa, Gallegos, et al. (2021). Briefly, the apparatus consisted of a panel with three odor ports, each one connected to an independent olfactometer (Figure 2). Each olfactometer was equipped with six solenoid valves connected to six odor vials. Of

FIGURE 2 Picture of the training and operational rooms. The olfactometer, size, and room size and color were identical. The only difference was the orientation of the olfactometer relative to the door. For the first and last cohort, Room A was the operational room, and for the second cohort Room B was the operational. The dogs' performance was not affected by the room selected as the operational or training room.

the six vials, one was the target odor and the others were distractor odors. Odor presentation and randomization was conducted by a computer program that ensured the target odor was presented equally across all three ports. In a given trial, the computer randomized the odor to be presented in each port and activated the corresponding valve in each olfactometer. Filtered air (1 L/m) carried the headspace of the odor vial into a PTFE manifold where it was diluted with a continuous airline (2 L/m) and subsequently carried to the odor port. The air dilution used (33% odor dilution) was identical for the target and distractor odors. Each odor port was equipped with an infrared sensor (IR) at the front of the port that measured if a dog sampled a port and the duration of the nose hold. The IRs were used to record responses during a trial (see training session below). A trial within a session started with lifting the panel that covered the ports (raised by a stepper motor). The experimenter then gave the dog the search command, and the dogs performed an off-leash search. The dogs had 45 s to search the apparatus and make a response. If the dog did not respond within 45 s, the trial was terminated and scored as a "timeout." The computer program saved the dogs' responses within a session in a comma-separated values file.

Training

The initial training to the apparatus consisted of teaching the dogs to search all three ports, alert to the port containing food (hot dogs) odor, and ignore (not alert) five nontarget odors/distractors. The odor of cotton gauze (Equate), nitril glove (MedPride, MPR-50504), food grade mineral oil (Bluewater Chemgroup), limonene (10^{-3} v/v dilution in mineral oil; CAS # 5989-54-8), and clean air were used as distractors in this study. These distractors were selected because they are common laboratory components that are used to prepare odorants and include a novel strong odor (limonene). The dogs were trained to alert to the port containing the target odor by holding their nose in the port for 4 s. To train the 4-s hold, we used positive reinforcement of successive approximations (shaping). Initially, the dogs were reinforced (with food or a treat) for just sampling or introducing their nose in the port containing the target odor. Subsequently, if a dog's performance was >85% correct responses in a 40-trial session, the nose-hold criterion was increased by 0.5-s steps in subsequent sessions until reaching the 4-s criterion. The nose-hold duration was recorded by the computer using the IR sensors in each port. By using the IRs, the apparatus was also able to record dog responses, search latency, the number of times a dog searched a port, and the time a dog sniffed a port during each trial. For instance, the apparatus measured search latency as the time from the initiation of a trial (panel raising) until the IR beam of any port was broken. The olfactometer program algorithm was also built to measure the amount of time the IR beam was broken and preprogram to score a response if the IR beam was broken for 4 s. For instance, if a dog sniffed Port 1 for 1 s and Port 3 for 4 s the computer would record all nose insertions and that the dog alerted to Port 3. During initial training, the dogs received one or two 40-trial training sessions a day where the target odor was hotdog at a prevalence rate of 100% (one of the three ports contained hot dog odor for all trials) until they reached the nose-hold criterion of 2 s. Most dogs completed initial training to the apparatus within 20 training sessions (< 10 days of training). The average number of training session needed to reach the nose-hold criterion was 10.83 ± 1.08 SE (Max = 22; Min = 5).

After the dogs were searching the apparatus independently with a nose hold of 2 s, the target odor was changed to 10 g of double-base smokeless powder (Hodgdon H335; referred to as SP subsequently) in a 40-mL vial. With SP as the target odor, the nose-hold criterion was subsequently gradually increased to 4 s (the final criterion). The dogs continued training to SP as described in our previous work (Aviles-Rosa, Gallegos, et al., 2021). The dogs received two 40-trial training sessions daily only in the training room where the target odor frequency was 90% (e.g., the target odor was present in 36 of the 40 trials) until their performance was >85% correct responses in two consecutive double-blind (e.g., the handler was blind to the location of the target and responses were recorded by the computer) sessions. Performance greater or equal to 85% correct responses was considered our training criterion. Although the number of training sessions varied with individuals based on performance, all 18 dogs met the training criterion within 10 days of training (~ 15 training sessions) with SP as the target odor. A detailed description of the training progression of 12 of the 18 dogs used can be found elsewhere (Aviles-Rosa, Gallegos, et al., 2021).

After dogs reached training criterion in two consecutive double-blind sessions in the training room, they were trained on an identical olfactometer in a different room. This room was denoted as the operational room. Both rooms were almost identical, but the position of the olfactometer within the room was different to facilitate room discrimination by the dogs (Figure 2). For the first and last cohort (12 dogs), Room A was the operational room. For the second cohort, Room B was the operational room to counterbalance any potential room effects. We assigned rooms as operational or training by cohort rather than by dog to facilitate training and testing. During training to the operational room, the dogs received one 40-trial double-blind session in the training and operational rooms daily. The order in which each room was tested was alternated every day. During training, in both rooms, a training session contained 10% (e.g., four out of 40 trials) blank trials with no target odor present in any of the ports; 90% of trials contained a target in one of the three ports. The position of the blank trials within a session was randomized by the computer program, and on average, there was one blank trial every 10 trials. During an odor-present trial, one of the ports contained SP and the other two ports contained distractor odors (selected randomly by the olfactometer program with equal probability). A correct response during an odor trial was noted if the dog correctly held its nose for 4 s in the port containing SP odor. A correct response for a blank trial was noted if the dog searched all three ports but did not alert to any of them after four continuous seconds of removing their nose from all ports (i.e., an "all-clear" response). Correct responses to the target odor were reinforced with a treat (continuous schedule of reinforcement), but correct all-clear responses were not reinforced but simply led to the next trial. Correct all clears were not reinforced because this is a common practice in detection dog training. The trial duration was set at 45 s. If the dog failed to alert or make an all-clear response within 45 s after the initiation of a trial, the trial was terminated and the trial was scored as a "timeout" and counted as incorrect. Each dog was required to detect SP with an accuracy >85% on two consecutive double-blind sessions in both rooms before they could start the experiment. Transitioning training to the operational room was quick, and the dogs were able to transfer training from the training room within 5–8 sessions for all dogs.

Experimental design

After meeting the training criterion in both rooms, the dogs received one double-blind session in each room daily for the duration of the experiment (14 days). Sessions were scheduled at least 2 hr apart from each other to prevent fatigue. Double-blinded conditions were ensured by the olfactometer because the trial randomization, odor presentation, and data collection were done by the computer program. The olfactometer indicated the initiation of a trial to the experimenter by lifting the panel covering the odor ports. The experimenter then gave the command "search" to the dog so that they could search all three ports. If the dog did not search all ports after the first command, the handlers repeated the command "search" only one more time within a trial. If a dog alerted to the correct port, the computer marked the correct response with a "bleep" and activated an automated feeder that delivered a treat to reinforce the correct response. Incorrect responses, all-clears, and timeouts resulted in the termination of the trial without reward. At the end of the trial, the odor valves turned off and the panel slowly covered the ports for a 20-s intertrial interval. The next trial started immediately after the 20-s intertrial interval.

Baseline period

After reaching training criterion in both rooms, the dogs entered the baseline period where they received five 40-trial sessions in the operational and training room at a target odor prevalence of 90% (e.g., 10% of the trials were blank). Each dog received two sessions daily, one in the operational room and one in the training room. The

TABLE 1 Ethogram with behaviors that were coded for each dog during Experiment 1 in the testing period in both the operational and training rooms

Behavior description	Behavior coded	Interobserver agreement
Dog approached three-odor port panel (event)		
Tail Position	1. Relaxed	
	2. Tucked	Cohen's Kappa = .60
	3. Raised/Aroused	
	4. Wagging	
Licking odor port	1. Yes	
	0. No	Too infrequent
Yawning during odor port search	1. Yes	
	0. No	Too infrequent
Vocalization during searching	1. Yes	
	0. No	Too infrequent
Look back at handler during searching	1. Yes	
	0. No	ICC < .40
Search behavior and interaction with odor ports	0. No approach 1. "Drive by" search and no nose insert 2. Bumped or poked panel	
	3. Searched ports but not all 4. Detailed search	ICC = .90
Panting (duration)		Too infrequent
Lying down (duration)		Too infrequent
Sitting (duration)		Too infrequent
Self-grooming (duration)	1. Licking	
	2. Scratching	ICC = .89
Environment directed behavior (duration)	1. Jumping on handler	
	2. Mouthing handler	
	3. Jumping on environment/wall	ICC = .99
Latency (duration)	1. Time from start of trial to first nose insert	Coded by computer

Note. Description of behavior category that was coded has the specific behavior identified for that category listed in "Behavior coded." Interobserver agreement is shown from double-coded videos. Behaviors that were scored less than 1% of trials or with interobserver agreement < .40 were not analyzed and are listed as too infrequent.

order in which each room was tested was alternated every day. The purpose of the baseline period was to ensure that the dogs' performance was similar in both rooms when the target odor prevalence was the same. Ensuring that performance was similar in both rooms when the target odor prevalence was high allowed us to better evaluate the effect of low target odor prevalence during the testing period.

Testing period

The testing period started immediately after the baseline period. During testing, the target odor prevalence in the operational room was decreased to 10% (e.g., only four of the 40 trials had a target odor) and the target prevalence in the training room remained at 90%. The odor trials in the operational room were randomized by the computer program and, on average, an odor trial occurred every 10 trials, but not exactly after 10 trials. The reduction of target odor only in the operational room was to mimic a real-life scenario where a detection dog will have a high odor prevalence during frequent training sessions but a low odor prevalence in an operational scenario (e.g., when patrolling or searching an area). The dogs received one daily session in each room for five consecutive days under these conditions to evaluate whether the reduction in odor prevalence from 90% to 10% produced a substantial decrement in the dogs' performance and search-related behaviors in the operational room but not in the training room.

The testing period for both the training and operational rooms were video coded using Behavioral Observation Research Interactive Software (BORIS; Friard & Gamba, 2016). A total of 180 videos were coded (five videos in each room for each dog) to evaluate the effects that reducing the target odor frequency in the operational room had on the dogs' behavior (Table 1). Ten percent of the videos were double coded (18 videos total) by two observers to determine interobserver agreement. The videos were coded using the behaviors defined in Table 1 in addition to tracking the trial, name for each dog, session, and room. The BORIS video-coded behaviors were merged with the olfactometer-collected accuracy data to evaluate the relation between the dogs' behavior and performance. The interclass correlation coefficient (ICC) for continuous and Cohen's kappa for categorical variables are listed in Table 1 for interobserver agreement.

Recovery period

To evaluate whether the dogs' performance in the operational room could be recovered to baseline levels, after the fifth day of testing, the target odor prevalence in both rooms was set to 90% for four consecutive days. This was to evaluate whether the observed performance decrement was able to be recovered by just increasing the target odor frequency to training levels.

Control test

A control test was conducted at the end of the testing period and before starting the recovery period in the training room (where the dogs were expected to do well) to ensure that the dogs were alerting to the SP and not to any other olfactometer cue (e.g., sound of the odor valve). For this, the dogs received a 10-trial session at a 90% target odor prevalence where the odor jars were disconnected from the olfactometer. Apart from this change, the control session was conducted as normal. Thus, during nine of the 10 trials of the control session, one of the olfactometers activated the odor valve and the other two olfactometers distractor valves. During the blank trial all olfactometers activated a distractor odor valve. Because no odor vials were connected to the valves, if the dogs were alerting exclusively to the odor stimulus of SP, we expected them to not be able to respond to the correct port at a rate greater than chance. A correct response during the control test was recorded if the dog alerted to the olfactometer port that activated the target odor valve even when no odor was present. An incorrect response during the control test was recorded if the dog performed an all clear or if it alerted to an olfactometer port that activated a distractor odor valves (also without odor jars). Because no odors were present, we expected the dogs to perform mostly all-clear responses during the control test and not be able to identify the correct port. Performance above chance during the control test would indicate that the dogs were able to identify the correct port using unintentional nonolfactory stimuli from the apparatus (e.g. "click" or valve noises from the olfactometer).

Data collection

As mentioned above, using IRs, the olfactometer recorded search latency, the number of times a dog

TABLE 2 Measurements used to evaluate performance and search vigilance decrement

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 alert, %	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 s after the trial started	(Number of timeouts / 40 trials) × 100
Hit rate, %	Dog alerted to the odor correctly when present.	(Number of correct responses to odor / number of odor trials) × 100
Correct rejection, %	Dogs did 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 where the dog sampled all ports) × 100
Search–hit ratio, %	Dog alerted to the target odor after sampling the odor port	(Number of correct responses after the dog sampled the odor port / number of odor trials where the dog sampled the odor port) × 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 set as 45 s.	($\sum$ Latency / 40)
Entries	Number of times a dog sampled all ports in a trial.	($\sum$ Entries / 40)
Sniff time, s	Cumulative time a dog spent sniffing all three ports in a trial excluding the 4 s to make an alert.	($\sum$ Sniff time / 40)

sampled a port, the amount of time the dog sniffed each port, and each dog's response during each trial. The performance for each session was calculated by averaging the dogs' responses for all 40 trials within a session for each room. Table 2 shows the definition of the measurements used to evaluate the dogs' performance and how they were calculated.

Behavior data were collected for each trial within a session during the testing period. A search score was calculated for each trial that was coded from 0 to 4, qualitatively describing a range of search from "*no search*" to "*complete thorough search*." If a dog engaged in multiple searches in a trial (i.e., revisited the panel multiple times within a trial), the highest search score was retained for that trial, indicating the maximum level of search engaged for the individual trial. Tail position was retained as a categorial variable. Grooming behaviors were combined (licking and scratching) to make one duration-based behavior for grooming, which was the sum duration of grooming behaviors for each trial. Behaviors that were directed toward the environment/handler were also combined to make one behavioral category. The duration of each behavior was summed to create one combined duration of behaviors directed toward the environment for each trial. In total, 5,065 trials with accuracy data and behavioral data were successfully video coded and merged for analysis. Some video coding was not possible due to hard drive and server failure. The data from both the operational and training rooms were combined to optimize the search for behavioral differences between high-performance and reduced-performance sessions.

Data analysis

All data analyses and visualization were conducted using SAS 9.4 statistical software and R studio. The average performance of each dog within a session was calculated and used for the statistical analyses.

A linear mixed model was used to evaluate the effect of testing session (1–5), room (operational vs. training), period (baseline, testing, and recovery), and the Room × Period interaction. A random effect of dog was included. A statistically significant difference was declared when $p < .05$. If the main effects or interaction showed a statistically significant effect, the Tukey–Kramer test was used for multiple comparisons. We used this model to evaluate the effect of these factors on all the performance measurements described in Table 2.

A separate linear mixed model was conducted to evaluate the effect of session within each period. This model included the fixed effect of room, session, and their interaction with the random effect of dog. A statistically significant difference was declared when $p < .05$, and the Tukey–Kramer test was used for multiple comparisons.

For video-coded data, frequency for event behaviors and duration for state behaviors were calculated for each dog, trial, and session. This data set was then merged with the olfactometer trial data to include latency and performance for each individual trial (i.e., correct or incorrect). Behaviors were then filtered to remove behaviors that occurred too infrequently for analysis (defined as occurring on <1% of trials) in addition to behaviors with poor interobserver agreement, defined as Cohen's kappa < .60. Table 1 shows the behaviors that were excluded and the behaviors that remained. The remaining behaviors ranged from moderate (Cohen's kappa = .60) to excellent reliability (ICC > .90; Viera & Garrett, 2005).

To predict whether observed behavior was associated with performance, a generalized linear mixed-effect model (binomial link) was fit in which trial outcome (correct or incorrect) was predicted by latency (time from start of trial until the first nose insert), search score, grooming behavior, environment-directed behavior, tail position, and a random effect for each dog. To maximize the prediction of behavioral characteristics for high and low performance regardless of room, variables related to room (operational or training) were not fit. An analysis of variance from the car package in R studio was used for a chi-square analysis for each fixed effect. Post hoc tests for tail position were conducted using Tukey-adjusted p values. An additional model was fit in which we included only the trials where a target odor was present to detect specifically odor misses. The purpose of this model was to evaluate the effect of each behavior on the hit rate.

RESULTS

Within 25 days of training, most participants had an overall accuracy ≥85% in both rooms under double-blind testing. Only one dog (Axel) needed two extra weeks of training to meet training criterion. After training, his performance did not differ from the performance of other dogs. Figure 3 shows the individual performance of each dog during the different phases of the experiment. One dog (Pumpkin) received only four baseline sessions in the operational room due to problems with the olfactometer, and another dog (Dallas) received only three sessions in the operational room during recovery due to the handler's unavailability for the final day. A visual examination of individual performance shows that all dogs had similar performance in the operational and training room during baseline. Charm, Raven, Wishbone, and Dozer had one session in the operational room where performance was below our training criterion during baseline, but overall, there was no significant difference in performance during baseline in both rooms. All dogs showed a significant decrement in performance during the first session of the testing period in the operational room. Charm was the only dog that also had low

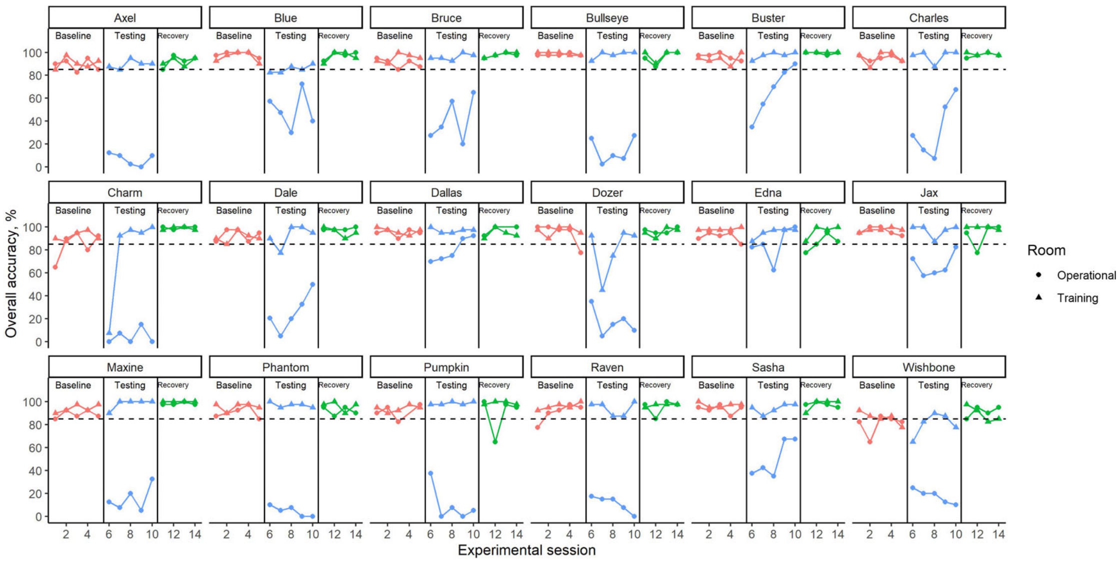

FIGURE 3 Individual dog overall accuracy across the different testing periods and room. The dashed line indicates 85% overall accuracy, our training criterion. Most dogs showed highly similar performance decrement during the testing period. Six dogs (Bruce, Buster, Charles, Dale, Jax, and Sasha) showed performance improvement in the operational room by the end of the testing period and one dog (Edna) did not show a significant performance decrement during the testing period in the operational room. This indicates that with time, some dogs can improve performance in a low-prevalence room, suggesting that some dogs are more resilient to these conditions than others are.

performance in the training room during the first session of the testing period. The performance of nine dogs (Axel, Bullseye, Charm, Dozer, Maxine, Phantom, Pumpkin, Raven, and Wishbone) in the operational room during the testing period did not show any significant improvement during the five sessions and was below 40% correct responses. However, the data from the remaining nine dogs show that their performance tended to improve with sessions. Moreover, by the end of the fifth testing session in the operational room, Buster's, Dallas's, Edna's, and Jax's performance was at or above training criterion. This might suggest that some dogs may be more resistant to low target odor prevalence with time. All dogs recovered performance in the operational room immediately during the first recovery session.

The room (operational or training) by period (baseline, testing, recovery) interaction was significant for all the variables measured ($p < .001$), except for the sniff time, false all clears, and the search–hit ratio (Table 3). For these measurements, the main effects of room and period were also not statistically significant ($p > .05$). Table 3 and Figures 3 and 4 show the average overall accuracy and all the performance measurements during the different testing periods.

The overall accuracy during baseline in the training and operational room did not differ statistically (training: 94.66 ± 0.48% vs. operational: 92.11 ± 0.73%). Search-related behaviors measured by the olfactometer (latency, port entries, sniff time, and timeouts) and the other performance variables measured were also not statistically different from each other during baseline in the operational or training room (Table 3; Figure 5), indicating that performance and search-related behaviors were similar in both rooms during baseline.

The performance of 17 of the 18 dogs tested decreased below training criterion immediately after the first session where the target odor prevalence was reduced to 10% in the operational room (Figure 3). Reducing the target odor prevalence in the operational room during the testing period resulted in a statistically significant increase in timeouts (operational: 53.31 ± 3.53% vs. training: 2.69 ± 1.17%), increased search latency (operational: 25.78 ± 1.39 s vs. training: 8.06 ± 0.52 s), and a decrement in the number of port entries (operational: 1.98 ± 0.15 vs. training: 2.79 ± 0.06) relative to the training room where the odor prevalence was unaltered. This significant decrement in search-related behaviors resulted in a statistically significant reduction in the overall accuracy (operational: 33.37 ± 3.08% vs. training: 92.53 ± 1.30%) and correct rejections (operational: 31.89 ± 3.08% vs. training: 82.50 ± 2.72%), an increase in the number of false alerts (operational: 13.26 ± 1.32% vs. training: 3.00 ± 0.51%), and thus a statistically significant decrement in the hit rate (operational: 46.67 ± 3.95% vs. training: 93.64 ± 1.26%).

Interestingly, the proportion of false all clears (operational: 1.89 ± 1.55% vs. training: 2.11 ± 0.34%), search–hit ratio (operational: 98.11 ± 1.55% vs. training: 97.79 ± 0.35%), and the sniff time (operational: 0.90 ± 0.08% vs. training: 0.84 ± 0.04%) in the operational and training room were not statistically different from each other during testing. Thus, the dogs' probability of alerting if they thoroughly investigated the port with the target odor was unchanged throughout, indicating that the decrement in performance measured was due

TABLE 3 Mean ± standard error of the mean (*SE*) of dogs' ($n = 18$) performance during the different testing periods and room

	Baseline		Testing		Recovery		*Room*		Period		Interaction	
Variable	**Training ($n = 18$)**	**Operational ($n = 18$)**	**Training ($n = 18$)**	**Operational ($n = 18$)**	**Training ($n = 18$)**	**Operational ($n = 18$)**	***F***	***p***	***F***	***p***	***F***	***p***
Overall accuracy, %	94.66 ± 0.48^{a}	92.11 ± 0.73^{a}	92.53 ± 1.30^{a}	33.37 ± 3.08^{b}	96.74 ± 0.50^{a}	95.21 ± 0.75^{a}	102.96	<.001	89.58	<.001	84.69	<.001
False alert, %	2.70 ± 0.38^{b}	3.96 ± 0.47^{b}	3.00 ± 0.51^{b}	13.26 ± 1.32^{a}	1.42 ± 0.39^{b}	1.19 ± 0.27^{b}	32.47	<.001	23.54	<.001	25.03	<.001
Timeout, %	0.99 ± 0.21^{b}	1.71 ± 0.46^{b}	2.69 ± 1.17^{b}	53.31 ± 3.53^{a}	0.48 ± 0.19^{b}	1.81 ± 0.60^{b}	60.26	<.001	56.24	<.001	53.77	<.001
Hit rate, %	96.30 ± 0.38^{a}	93.85 ± 0.67^{a}	93.64 ± 1.26^{a}	46.67 ± 3.95^{b}	97.27 ± 0.48^{a}	95.38 ± 0.74^{a}	56.31	<.001	54.45	<.001	43.55	<.001
Correct rejection, %	79.82 ± 3.11^{ab}	76.40 ± 2.89^{b}	82.50 ± 2.72^{a}	31.89 ± 3.08^{c}	92.01 ± 2.26^{a}	93.73 ± 1.62^{a}	38.38	<.001	48.76	<.001	35.61	<.001
False all clear, %	1.87 ± 0.25	2.57 ± 0.37	2.11 ± 0.34	1.89 ± 1.55	1.53 ± 0.34	2.06 ± 0.38	0.45	.50	0.14	.87	0.23	.79
Search–hit ratio, %	98.11 ± 0.26	97.27 ± 0.37	97.79 ± 0.35	98.11 ± 1.55	98.39 ± 0.34	97.75 ± 0.38	0.57	.45	0.11	.89	0.37	.69
Latency, s	8.28 ± 0.32^{b}	8.66 ± 0.35^{b}	8.06 ± 0.52^{b}	25.78 ± 1.39^{a}	7.14 ± 0.31^{b}	7.78 ± 0.38^{b}	59.33	<.001	49.40	<.001	50.21	<.001
Port entries	2.73 ± 0.04^{a}	2.80 ± 0.07^{a}	2.79 ± 0.06^{a}	1.98 ± 0.15^{b}	2.83 ± 0.04^{a}	2.78 ± 0.04^{a}	8.27	.006	7.68	.002	9.10	<.001
Sniff time, s	0.79 ± 0.04	0.73 ± 0.02	0.84 ± 0.04	0.90 ± 0.08	0.72 ± 0.02	0.77 ± 0.04	0.10	.75	1.71	.19	0.57	.56

Note. $^{a–c}$ Values with different superscripts in a row are statistically different from each other.

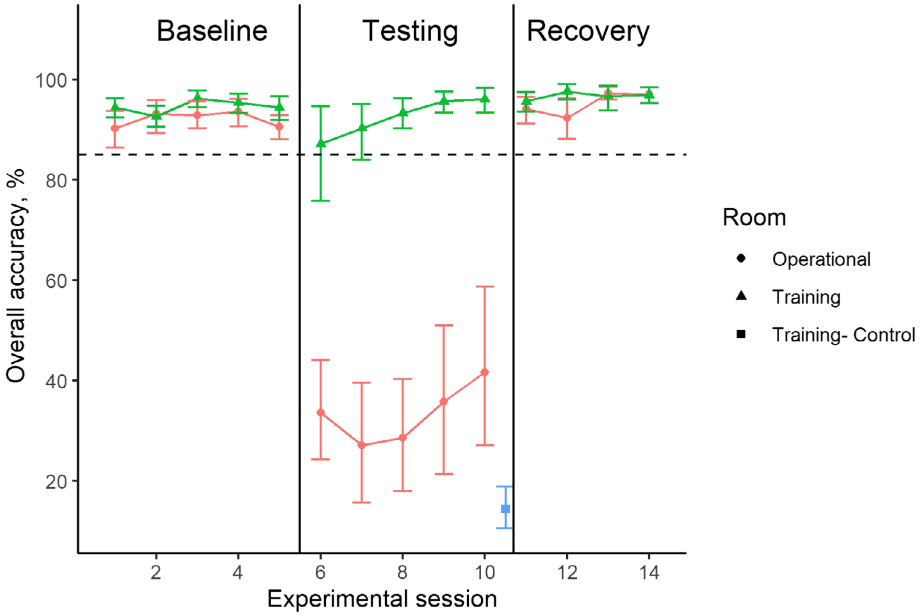

FIGURE 4 Dogs' ($n = 18$) overall accuracy during each session within the different experimental periods, rooms, and control test. The points show the mean, and the error bars show the 95% confidence interval of the mean for each experimental session within room. The dashed line indicates 85%, which was our training criterion. The horizontal lines indicate the change in experimental period (baseline, testing, and recovery). There was a significant effect of room and period interaction. No difference was observed in the dogs' overall accuracy during baseline between rooms. Decreasing the target odor prevalence in the operational room produced a significant reduction in the overall accuracy during the testing period. The overall accuracy returned to baseline levels immediately during the first recovery session when the target odor prevalence was increased to 90%. This indicates that the performance decrements observed in the operational room during the testing period were mainly due to changes in the odor prevalence rate. A significant decrement in performance was observed during the control test (training–control). This indicates that the dogs' performance was dependent on the presence of an odor stimulus within the olfactometer.

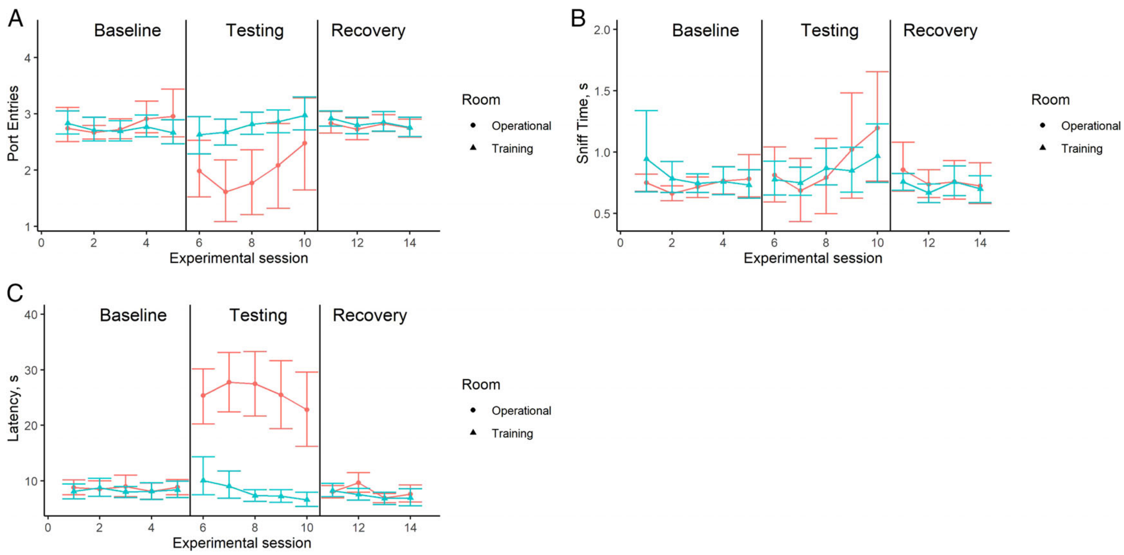

FIGURE 5 Dogs ($n = 18$) search-related behaviors during the different testing phase and room. Panel A: Mean ± 95% confidence intervals of the numbers of cumulative port entries the dogs did in a trial. The horizontal lines indicate the change in experimental period (baseline, testing, and recovery). The number of port entries was similar for both rooms during baseline and recovery. A slight reduction in the number of ports sampled was observed in the operational room during the testing phase. Panel B: Mean ± 95% confidence interval of the cumulative sniff time (s) during a trial. The horizontal lines indicate the change in experimental period (baseline, testing, and recovery). The effect of room, period, and their interaction on sniff time were not statistically significant. Panel C: Mean ± 95% confidence intervals of dogs' search latency. The horizontal lines indicate the change in experimental period (baseline, testing, and recovery). Search latency was not statistically different in the operational or training room during the baseline or recovery period. Latency significantly increased during the testing period in the operational room but not in the training room. This suggests that low target prevalence rates reduced search behavior because the dogs required more time to engage in searches.

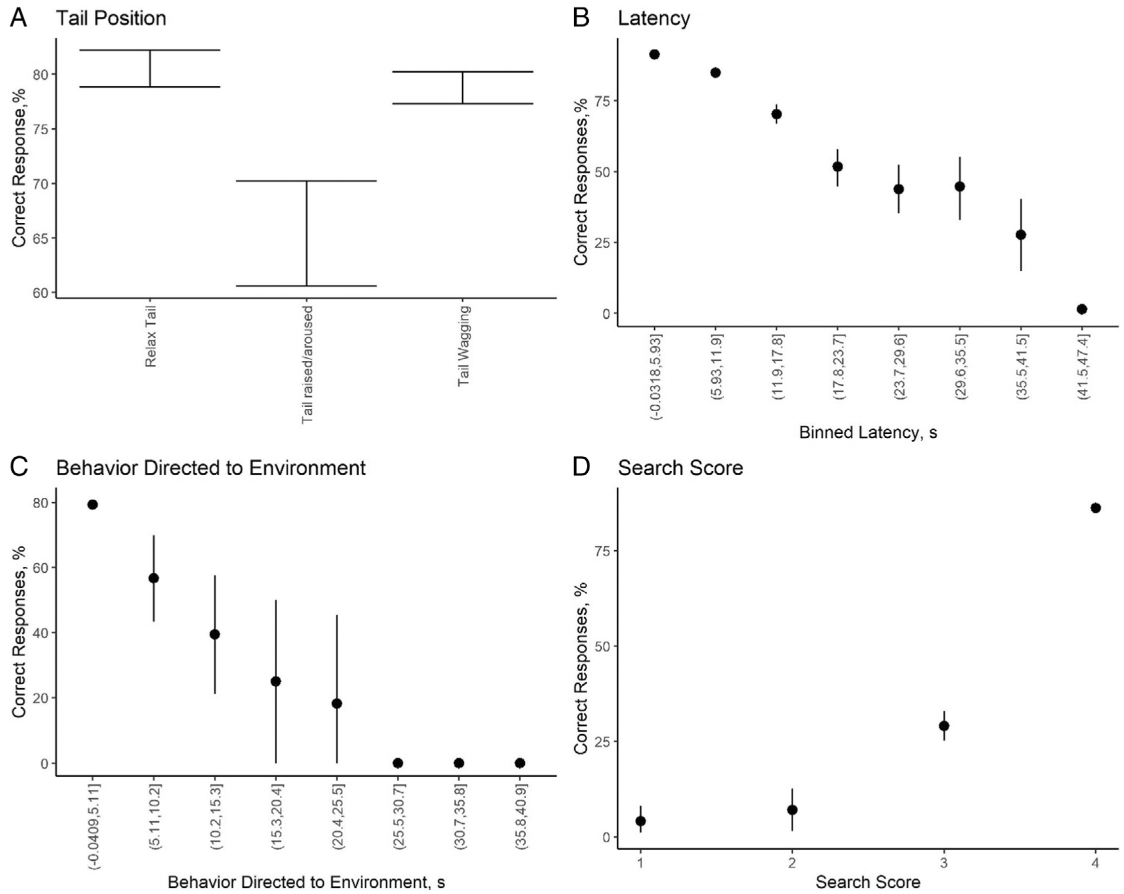

FIGURE 6 Behavioral predictors of overall accuracy when the dog was correct. Means ± bootstrap-estimated 95% confidence intervals are shown for overall accuracy (correct vs. incorrect). The continuous predictor variables are binned to improve visualization for the binomial data and continuous predictors. Binned latency and behavior directed to the environment are represented in seconds. Panel A: The percentage of correct responses during trials where dogs showed a particular tail position. An aroused tail was correlated with lower percentages of correct responses relative to a raised and wagging tail. Panel B: The percentage of correct responses in relation to search latency. The graph shows how shorter latency to search predicted higher accuracy. Panel C: Shows how the percentages of correct responses improved as the dogs showed less environmental directed behaviors. Panel D: Illustrates how a higher search score predicted more correct responses.

to not investigating the target port. Additionally, although the dogs' performance decreased in the operational room during the test phase, performance in the training room was consistently above training criterion and it was not statistically different from their performance during baseline (Table 3; Figure 5) for any of the variables measured.

Performance in the operational room was recovered to baseline levels immediately after the target odor prevalence was increased to 90% during the recovery period. The dogs showed significant improvement in the operational room during recovery relative to the testing period, with improvements in overall accuracy (testing: 33.37 ± 3.08% vs. recovery: 96.74 ± 0.50%), hit rate (testing: 46.67 ± 3.95% vs. recovery: 97.27 ± 0.48%), correct rejection (testing: 31.89 ± 3.08% vs. recovery: 92.01 ± 2.26%), and port entries (testing: 1.98 ± 0.15 vs. recovery: 2.83 ± 0.04). During the recovery, dogs also showed improvements in search latency (testing: 25.78 ± 1.39 s vs. recovery: 7.40 ± 0.31 s), proportion of timeouts (testing: 53.31 ± 3.53% vs. recovery: 1.81 ± 0.60%), and false alerts (testing: 13.26 ± 1.32% vs. recovery: 1.42 ± 0.39%) in the operational room compared with the testing period.

When the effect of session was evaluated within each testing period separately, we found that neither session nor the Room × Session interaction was statistically significant for any of the performance or behavior measurements during the different periods (baseline, testing, or recovery). Even though these effects were not statistically significant, a visual inspection of the individual dog data (Figure 3) shows that some of the dogs showed a slight improvement in the overall accuracy in the operational room at the end of the testing period (Figure 3.). This suggests that some dogs are more resilient to the low prevalence rate and that, with time and training, they can improve performance in a low target prevalence room. Nevertheless, most of the dogs were not able to improve performance in the operational room within five sessions.

During the control test, where the odor jars were disconnected, all dogs showed a substantial drop in

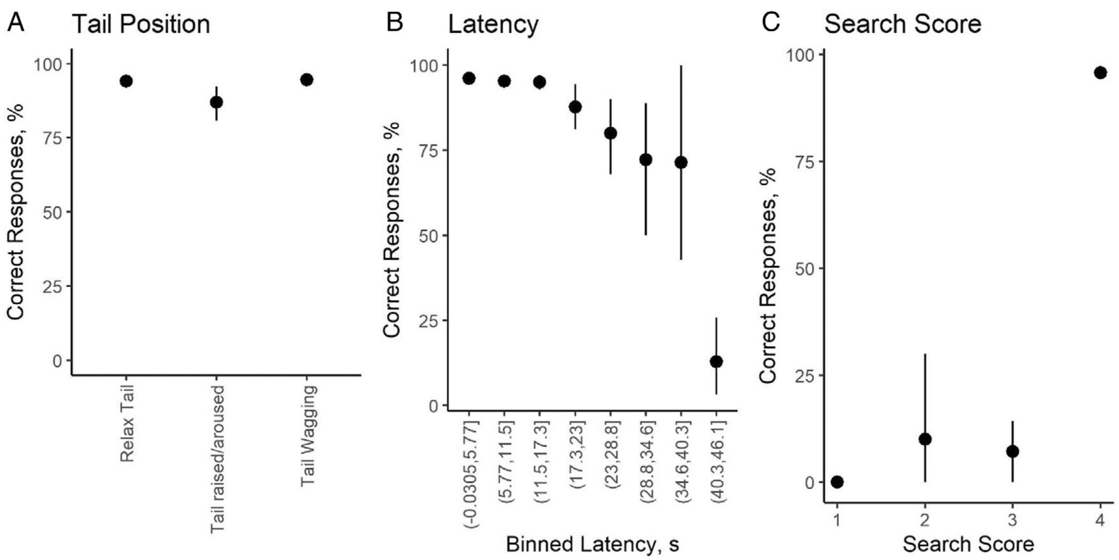

FIGURE 7 Behavioral predictors of correct responses during trials containing the target odor. Means ± bootstrap-estimated 95% confidence intervals of correct responses. The continuous predictor variables are binned to improve visualization for the binomial data and continuous predictors. Panel A: The effect of tail position on the percentage of correct responses. Panel B: Shows how shorter latency is associated with higher percentage of correct responses. Similarly, Panel C: Shows how a higher search score is also associated with higher percentage of correct responses.

performance. The average accuracy of the dogs during the control test was 11.66 ± 2.40%. Performance in the control test was well below their performance during the baseline, testing, and recovery phases in the training room. The control test results confirmed that the dogs' performance was mediated by the presence of the target odor and was not influenced by unintentional cues from the olfactometer.

For the video-coded data, the logistic mixed-effect model indicated that several behavioral predictors were associated with overall performance accuracy (Figures 6 and 7). Increased trial accuracy was associated with tail position ($\chi^2 = 6.59$, $p = .04$), higher search score ($\chi^2 = 305.47$, $p < .0001$), reduced latency ($\chi^2 = 320$, $p < .0001$), and less environmentally directed behavior ($\chi^2 = 6.60$, $p = 0.01$), but performance was unrelated to the duration of grooming behaviors ($\chi^2 = 0.83$ $p = 0.37$). The post hoc test for tail position indicated that a "relaxed tail" was associated with higher accuracy compared with a stiff tail ($z = 2.52$, $p = 0.03$), but there was not a difference between a relaxed tail and wagging tail ($z = 1.42$, $p = 0.33$) or between a wagging tail and stiff tail ($z = 1.75$, $p = 0.19$).

Next, behaviors that were associated with accuracy only in trials where the target odor was present (i.e., an error would therefore be a miss of the target odor) were evaluated. Similar to overall accuracy, lower search latency ($\chi^2 = 40.20$, $p < .01$), higher search score ($\chi^2 = 79.61$, $p < .0001$), and tail position ($\chi^2 = 6.94$, $p = 0.03$) were associated with higher probability of alerting to the target odor correctly. Environmentally directed behavior ($\chi^2 = 2.86$, $p = .09$) and grooming ($\chi^2 = 0.01$, $p = .92$) were not associated with a correct response during a trial where the target odor was present. The post hoc test for tail position again indicated that a relaxed tail compared with a stiff tail was associated with higher probability of a hit ($z = 2.57$, $p = .02$). In addition, a wagging tail was nearly associated with higher hits than a stiff tail ($z = 2.26$, $p = .06$), but there was no difference between a relaxed and wagging tail ($z = 0.96$, $p = .60$).

DISCUSSION

Performance

The effect of reducing the frequency of the target odor on dogs' search-related behaviors and performance was evaluated. The data indicated that reducing the target odor frequency in the operational room during testing resulted in an increase in the proportion of timeouts and search latency, as well as a reduction in the number of port entries. These negative changes in search-related behaviors measured by the olfactometer were responsible for a significant decrement in detection performance in the operational room relative to the training room during testing. Figure 8 shows a graphical representation of the effects that reducing the target odor frequency had on search behavior and detection performance. During the testing phase, in the operational room, the dogs did not search all ports (timeouts) after the handlers' command in 53% of the trials that contained the target odor. The proportion of timeouts in the operational room during the testing period was 18 times greater than that in the training room during the same period.

We also observed that the dogs took longer to initiate search (search latency) and made fewer port entries (e.g., more port entries result in a higher search score).

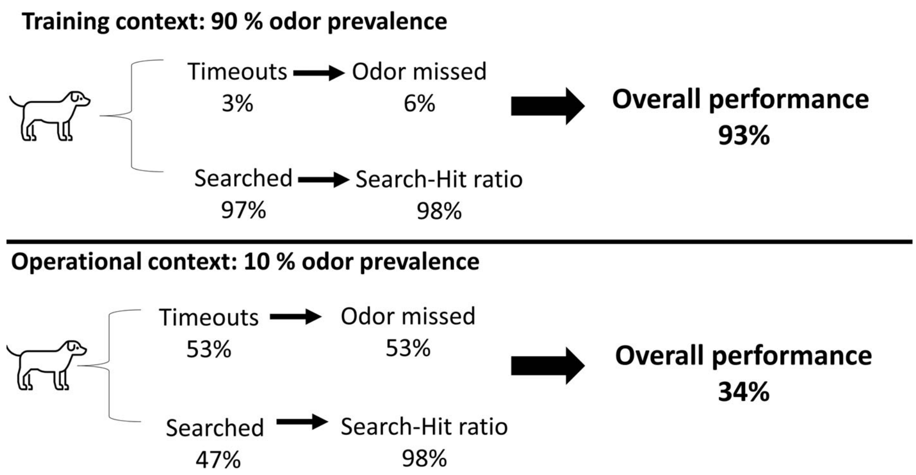

FIGURE 8 An illustration of the main effect a low target odor frequency room has on search behavior and detection performance

Thus, based on the changes observed on these behavioral measurements, we can conclude that the main effect of reducing the target odor prevalence in the operational room was a significant decrement in search-related behaviors. Based on our results we infer that this decrement in search-related behaviors was then responsable for the reduction in performance observed. This result indicates that our manipulation of target odor frequency in one context satisfactorily induced extinction effects in search behavior and an overall performance decrement, creating a new laboratory model for evaluating these effects for detection dogs.

Interestingly, there was no difference in the search–hit ratio between the training and operational rooms during testing. This indicates that when the dog searched the port containing the target odor, the probability of a correct response was high (98%) in both rooms (Figure 8). This measurement confirms that the decrement in hit rate observed in the operational room was due to the extinction of search behavior and not that the odor–response relationship was extinguished or disrupted because even in the operational room, if a dog searched the port containing the target odor, they alerted to the odor at a high rate. This confirms that low target odor frequency induces extinction of search behaviors (first link of the behavior chain; Figure 1) rather than extinction of the odor–response relationship (second link of the behavior chain; Figure 1).

The fact that dogs always maintained high performance in the training room shows that the decrement in search behavior was due to the low target frequency in the operational room specifically. This agrees with previous research on this topic. For instance, Gazit et al. (2005) found that search behavior and detection performance was unaltered in a path with high target odor frequency (Path A) when a decrement was observed in a similar path after experiencing no target odors (Path B). Our results together with Gazit et al. (2005) confirm that dogs can discriminate between similar places and are highly sensitive to target odor frequency.

The data showed that the effects of reducing the target odor frequency are immediate. For instance, we found that search behavior and performance decreased during the first session. As in our study, Gazit et al. (2005) also reported that search behavior decreased during the first session in Path B (path with no target odor) and that it remained low until the last session. Porritt et al. (2015) also found a statistically significant reduction in what they described as vigilance quickly when dogs searched an area with no target odor. This shows that the search and performance decrement in places with low target odor prevalence occurs rapidly. In our study, this rapid decrement in search behavior most likely occurred because the dogs were previously trained with high target odor frequency during training and baseline (90%). Therefore, the rapid transition from 90% target odor prevalence to 10% led to extinction-like effects on the search behavior. Thus, based on our results, we speculate that a possible way to mitigate the negative effects of low target odor frequency could be to gradually train dogs under an intermittent schedule of reinforcement to condition them to low target odor frequencies.

Search-related behaviors and detection accuracy in the operational room were recovered quickly by increasing the target odor frequency to training levels. This recovery in performance was not observed by Gazit et al. (2005). This difference between studies could have been because the target odor frequency used by Gazit et al. (2005) to recover performance was still significantly lower than in the control path. For instance, in their study, Path A (the control path) contained five targets within a session, whereas Path B (path with no explosives) contained one target odor every four sessions. This increment from zero to one target every four sessions most likely was not enough to recover search behavior. The fact that we were able to recover performance by simply increasing the target odor frequency indicates that the extinction of search behavior in a low target odor prevalence place is not permanent and that search behavior and performance can be recovered quickly.

Furthermore, it also indicates that the low target odor frequency was the main factor producing extinction of the search behavior. This suggests that the extinction of search behavior in detection dogs can be overcome by planting or placing a target odor in operational scenarios (operational finds) where the probability of finding a target is low. However, because of logistical and safety reasons, providing explosive detection dogs with operational finds may not always be feasible. Porrit et al. (2015) showed that the use of an innocuous odor to provide operational finds to dogs maintained performance in comparison with dogs that had zero finds for a period of 6 weeks and that the use of a nonexplosive target odor did not affect the dogs' performance on the explosive. Another way to provide dogs with operational finds is to use nondetonable training aids. The purpose of these training aids is to present dogs with a realistic representation of the odor of the actual explosive safely (Aviles-Rosa, Fernandez, et al., 2022). Future studies on the development of these training aids or alternative methods for providing dogs with operational finds are needed to further evaluate whether it is feasible to mitigate the search decrement observed when dogs are exposed to a low target odor frequency.

The results of the behavioral coding suggest that there are observable behavioral differences before the dogs make an incorrect response. Primarily, these differences are related to search behavior, which is what is primarily reduced when the target odor prevalence decreases. In this case, increased latency to search and poorer search behavior (missing ports or only scanning ports) led to a substantial decrease in performance. Increased latency likely reflects the extinction of search behavior. Increased delay to start searching was one of the strongest predictors of poorer performance, suggesting that this may be a very useful and simple behavioral change that handlers could use as a measurement to indicate reduced engagement in search and increased likelihood of an incorrect response. For instance, a dog might not be in a vigilance state if the handler gives the search command and notices that the dog does not engage in search behavior quickly.

The "search score" also appeared to be a very important predictor of accuracy and probability of missing a target. The results indicated that there is a substantial decrement if the dogs do not engage in a thorough search and that simple "drive by" or "bumping" of ports was not sufficient for accurate detection. This agrees with Porrit et al. (2015) where they reported that dogs that performed a less thorough search were more likely to miss a target odor. Dogs need to fully sample the port to make an accurate response. This also has interesting implications for handlers, suggesting that engaging in less detailed search patterns (e.g., simple pass by) could be an indication that the dog is not in a vigilance state and thus the probability of an incorrect response or a miss of a target is higher. Altogether the data showed that rapid engagement in search behavior and thorough searches were the two strongest behavioral predictors of the dog being in a vigilance state and thus ensured optimal performance.

Tail position showed a relatively weak association with overall accuracy and hit rate. The dogs performed poorer in trials where they had an upright tail than in trials where their tails were relaxed but not significantly lower than in trials where their tail was wagging. This suggests that perhaps an upright/aroused tail may indicate some change in the valence of the task or environment that is associated with increased probability of an incorrect response. An upright tail could be an indicator of stress in dogs. An increase in stress during the testing period in the operational room could have been the result of testing with a target odor frequency near extinction levels. It is important to note that the interobserver agreement for tail position was lower than the other behaviors. This variability in scoring suggests it was more difficult to code and results should be interpreted cautiously. The results of the behavior coding also suggest that engagement in other behavior directed to the environment can indicate decline in accuracy.

Altogether, the coded behaviors highlight that robust search behavior is critical to performance and variables that decline search behavior (such as reduced odor prevalence) may degrade search, leading to consequential declines in performance. Having handlers attend to deterioration in search latency or the quality or thoroughness of the search (decline from thorough to simply passing) might be a good method for ensuring that their dog is in a vigilance state. Monitoring dogs' behavior may allow handlers to anticipate when their dogs' performance is likely to decline and engage in responses to mitigate this decline, such as providing the dog a "find" in the current environment.

Limitations

Although the results of this study show a clear effect of target odor frequency on search behavior and performance, there are some limitations when trying to extrapolate these findings to detection dogs in general: (1) In this study we used mixed-breed dogs and not dogs from working lines. Thus, it could be that reducing the target odor frequency reduces search behaviors in mixed-breed dogs and not on dogs from working lines because working dogs are bred specifically for this behavior. However, it is very unlikely that our findings are not representative because previous studies that evaluated the same phenomena using working dogs found similar results (Gazit et al., 2005; Porritt et al., 2015). (2) An additional limitation is that dogs performed off-leash searches and were able to choose to search or not because the handler did not intervene in the search. This is different from the way many detection dog teams perform searches. Thus, further studies need to be conducted to determine whether this phenomenon also occurs in directed searches (the handler guides the dog to search an area). Nevertheless, in a previous study, Porrit et al. (2015) used free and directed searches to evaluate dog search vigilance and found similar results. (3) A third limitation is that we conducted the study over a short period where dogs previously naïve to odor detection work were trained for the purposes of the study. It is possible that with more training time and experience, the dogs' performance in the operational room would have rapidly recovered during the test as they learned the frequency of the target odor. Thus, our behavioral results may highlight a more dramatic effect of odor prevalence on search behavior and performance because these dogs had less experience and time training on the task. This suggests that additional research with operational dogs would be a critical next step. (4) Stark contrast between the two rooms could have also magnified the effect of odor prevalence. For instance, due to the very low target odor frequency in the operational room, the dogs could have distinguished the differences within the first couple of trials within a session and thus the reduction in search behavior could have occurred faster. Thus, using a prevalence rate closer to 10% in the training room could mitigate the negative effects observed in this experiment because it will be harder for the dogs to discern between rooms. This has to be further evaluated, as it could be a potential mitigation strategy for preventing the reduction in search-related behaviors observed herein.

CONCLUSION

We developed a laboratory model that mimicked how detection dogs are trained with high target odor frequency schedules in a training context but often deployed in an operational context where the target odor prevalence is often low. Within this model, we were able to replicate the reduction in search behavior and performance observed in previous studies. Altogether the data suggest that dogs can have excellent search behavior and detection performance when trained under high target odor prevalence but can simultaneously have poor search behavior and performance in areas with low target odor frequency. These findings have significant implications to detection dog teams because the data suggest that their performance during training might not be representative of their operational performance due to the negative effect that low target odor prevalence can have on search behavior. The laboratory model developed provides a new methodology to further investigate and remediate search behavior decrement in scenarios where targets are infrequent.

AUTHOR CONTRIBUTIONS

NJH and PAPT were the lead investigators for the project. They were responsible for the conceptualization of the study. EOAR and MD trained the dogs to use the olfactometer and collected the performance data. The performance data were analyzed by EOAR and NJH. Behavior coding was conducted by MD and analyzed by NJH. The manuscript was written by EOAR and NJH. All authors agreed with the final version of the manuscript.

ACKNOWLEDGMENTS

We acknowledge all the undergraduate and graduate students that help in the data collection and in the video coding, especially Lauren S. Fernandez. A special thanks to our dedicated husbandry crew that took excellent care of our dogs daily and ensure their wellbeing. We most importantly thank the dogs who made this work possible.

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 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.

DATA AVAILABILITY STATEMENT

The raw data can be found in the supplementary materials (Data S1).

ETHICAL STANDARDS

All procedures and animal handling were approved by the Texas Tech University Institutional Animal Care and Use committee (protocol # 19093-10).

ORCID

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

REFERENCES

Adams, J. A. (1987). Criticisms of vigilance research: A discussion. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *29*(6), 737–740. https://doi.org/10.1177/001872088702900612

Aviles-Rosa, E. O., Fernandez, L. S., Collins-Pisano, C., Prada-Tiedemann, P. A., & Hall, N. J. (2022). The use of an intermittent schedule of reinforcement to evaluate detection dogs' generalization from smokeless-powder. *Animal Cognition*, *25*, 1609–1620. https://doi.org/10.1007/s10071-022-01648-y

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., McGuinness, G., & Hall, N. J. (2021). Case study: An evaluation of detection dog generalization to a large quantity of an unknown explosive in the field. *Animals*, *11*(5), 1341. https://doi.org/10.3390/ani11051341

Aviles-Rosa, E. O., Nita, M., Feuerbacher, E., & Hall, N. J. (2022). An evaluation of spotted lanternfly (*Lycorma delicatula*) detection dog training and performance. *Applied Animal Behaviour Science*, *258*, Article 105816. https://doi.org/10.1016/j.applanim.2022.105816

Ballard, J. C. (1996). Computerized assessment of sustained attention: A review of factors affecting vigilance performance. *Journal of Clinical Neuropsychology*, *18*(6), 843–863.

DeChant, M. J. (2021). Training and experience factors affecting detection dog performance [Unpublished doctoral dissertation]. Texas Tech University, Lubbock.

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

Freeman, F. G., Mikulka, P. J., Scerbo, M. W., & Scott, L. (2004). An evaluation of an adaptive automation system using a cognitive vigilance task. *Biological Psychology*, *67*(3), 283–297. https://doi.org/10.1016/j.biopsycho.2004.01.002

Friard, O., & Gamba, M. (2016). BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*, 1325–1330.

Furton, K. G., Hong, Y.-C., Hsu, Y.-L., Luo, T., Rose, S., & Walton, J. (2002). Identification of odor signature chemicals in cocaine using solid-phase microextraction-gas chromatography and detector-dog response to isolated compounds spiked on U.S. paper currency. *Journal of Chromatographic Science*, *40*(3), 147–155. https://doi.org/10.1093/chromsci/40.3.147

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

Hancock, P. A. (2013). In search of vigilance: The problem of iatrogenically created psychological phenomena. *American Psychologist*, *68*(2), 97–109. https://doi.org/10.1037/a0030214

Hanigan, L., & Smith, N. (2014, September 1–4). *Determining the feasibility of training a dog to detect Hieracium species* [Paper presentation]. Proceedings of the 19th Australasian Weeds Conference. Hobart, Australia.

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experimental Brain Research*, *212*(3), 429–437. https://doi.org/10.1007/s00221-011-2749-1

Jantorno, G. M., Xavier, C. H., & de Melo, C. B. (2020). Narcotic detection dogs: An overview of high-performance animals. *Ciência Rural*, *50*(10), Article e20191010. https://doi.org/10.1590/0103-8478cr20191010

Jinn, J., Connor, E. G., & Jacobs, L. F. (2020). How ambient environment influences olfactory orientation in search and rescue dogs. *Chemical Senses*, *45*(8), 625–634. https://doi.org/10.1093/chemse/bjaa060

Lazarowski, L., & Dorman, D. C. (2014). Explosives detection by military working dogs: Olfactory generalization from components to mixtures. *Applied Animal Behaviour Science*, *151*, 84–93. https://doi.org/10.1016/j.applanim.2013.11.010

Lazarowski, L., Simon, A., Krichbaum, S., Angle, C., Singletary, M., Waggoner, P., Van Arsdale, K., & Barrow, J. A. (2021). Generalization across acetone peroxide homemade explosives by detection dogs. *Frontiers in Analytical Science*, *1*, Article 797520. https://doi.org/10.3389/frans.2021.797520

Mackworht, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, *1*(1), 6–21.

Matthews, G., Davies, D. R., & Holley, P. J. (1993). Cognitive predictors of vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *35*(1), 3–24. https://doi.org/10.1177/001872089303500101

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

See, J. E., Howe, S. R., Warm, J. S., & Dember, W. N. (1995). Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin*, *117*, 230–249.

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

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine*, *37*(5), 360–363.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Aviles-Rosa, E. O., DeChant, M. T., Prada-Tiedemann, P. A., & Hall, N. J. (2023). A laboratory model of canine search vigilance decrement, I. *Journal of the Experimental Analysis of Behavior*, 1–17. https://doi.org/10.1002/jeab.832