Applied Animal Behaviour Science 238 (2021) 105301

Contents lists available at ScienceDirect

Applied Animal Behaviour Science

journal homepage: www.elsevier.com/locate/applanim

Maintaining long-term odor memory and detection performance in dogs

Lucia Lazarowski [a,*], Paul Waggoner [a], Bethany Hutchings [b], Craig Angle [a], Fay Porritt [b]

[a] *Canine Performance Sciences, 104 Greene Hall, Auburn University College of Veterinary Medicine, Auburn, AL, 36849, United States*
[b] *Defence Science and Technology Laboratory, Porton Down, Salisbury, Wiltshire, UK*

ARTICLE INFO

Keywords:
Detection dog
Odor memory
Generalization
Maintenance

ABSTRACT

Detection dogs represent an essential resource for security and defense sectors, yet logistical and security constraints limit the frequency in which relevant maintenance training can occur. Further, the extents and limits of canine odor memory are unknown. The aims of this study were to examine the duration of long-term odor memory in detection dogs, and to evaluate effects of single-target maintenance training on the recall of previously trained targets, as well as generalization to a set of related but untrained targets. Detection dogs ($n = 18$) were trained on a set of 10 target odors followed by an assessment of baseline performance in odor discrimination and operational search tests. During the subsequent 12 months, half of the dogs received minimal intermittent odor detection maintenance training (approximately 30 min total per month) with a single target odor from the trained set, while the other half received commensurate training not involving odor detection (i.e., obedience training). Detection of the targets not experienced over the 12-mo period was then re-assessed. Generalization to untrained variants of the trained targets, a critical aspect of optimal detection that may be affected by extensive training with specific odors, was also assessed at baseline and after the 12-mo period. Accuracy in the odor discrimination test significantly decreased from 99 % ($SE = .25$) at baseline to 72 % ($SE = 4.36$) after 12 months for dogs that received no odor detection maintenance training, and significantly increased from 94 % ($SE = 1.57$) to 99 % ($SE = .53$) for dogs that received single-odor maintenance training over the 12-mo period. However, the maintenance training did not appear to maintain operational search performance. Further, results from the generalization test were less straightforward, suggesting that single-target maintenance training may not be sufficient for optimal operational detection of untrained variants. Our results indicate that long-term odor memory is largely robust in dogs, and that odor discrimination accuracy can be effectively maintained and even improved through minimal maintenance training using a single target odor, but that the task of searching in operational environments may require more regular maintenance.

1. Introduction

Detection dogs are an essential resource for a wide range of security, defense, and other tasks ranging from the detection of drugs and explosives to pests and disease. Due to their extraordinary olfactory capabilities and tractable social sensitivity to humans, dogs are considered the most capable and flexible technology available for detection of explosives and other contraband (Furton and Myers, 2001). However, there remain significant gaps in defining the basic cognitive and behavioral parameters of the operational characteristics of detection dogs (Hall and Wynne, 2016; Troisi et al., 2019). As detection task requirements continue to become increasingly specialized and rigorous, bridging such gaps will be critical for enhancing their capabilities through enhanced training and employment strategies.

One challenge faced by detection canine teams is the need for routine maintenance training to sustain operational performance. However, training in operational scenarios can be limited by logistical, security, and safety constraints. For example, practical and security concerns may limit the use of certain threat materials in particularly sensitive settings or in remote locations, and there is currently a lack of evidence for the effectiveness of surrogate or "pseudo" training aids designed to mimic actual substances (Simon et al., 2020). This limitation in conducting operationally-relevant training often leads to a discrepancy between training contexts, typically dense in reinforcement, and operational contexts, where dogs may rarely encounter a target or may not be rewarded for alerting due to the inability to confirm the identity of the substance or other operational limitations. Previous studies have demonstrated that dogs readily learn such contextual cues, with

* Corresponding author at: 104 Greene Hall, Auburn University College of Veterinary Medicine, Auburn, AL, 36849, United States.
E-mail address: lucia.lazarowski@auburn.edu (L. Lazarowski).

https://doi.org/10.1016/j.applanim.2021.105301
Received 15 January 2021; Received in revised form 8 March 2021; Accepted 11 March 2021
Available online 18 March 2021
0168-1591/

detection of trained targets significantly reduced in areas associated with a low probability of encountering targets, despite demonstrating the ability to detect the targets in different contexts (Gazit et al., 2005; Porritt et al., 2015). It is therefore critical that dogs are given the opportunity to be rewarded for finding trained targets outside of training contexts.

One promising method for maintaining operational detection performance was demonstrated by Porritt et al. (2015), where the reinforcement for detection of a single innocuous odor in a search context maintained detection rates of other targets (not encountered in that setting) at similar levels of dogs that received practice with the primary targets (in the relevant context). This effect has also been supported by studies with rats, using both intra- and extra-modal stimuli as surrogate training targets (i.e., maintenance training with auditory targets maintained detection of visual targets and vice versa) (Thrailkill et al., 2018), which suggests that the effects of the surrogate maintenance target could not be due to simple stimulus generalization based on perceptual similarity of the surrogate to the targets. Rather, engaging in intermittent practice of the behavior chain enabled the maintenance of the response despite stark differences in the stimuli.

In Porritt et al. (2015) and Thrailkill et al. (2018), the maintenance of the response for the unpracticed target was obviously dependent upon remembering that stimulus across a period of time. In Porritt et al. (2015), dogs were only required to remember three target odors across approximately six weeks, and the rats in Thrailkill et al. (2018) were only required to remember two target stimuli across approximately three weeks. However, explosive detection dogs are required to detect an increasing number of different targets, and variants of those targets, without maintenance training on each specific odor for extended periods of time. Previous studies examining long-term odor memory in dogs have demonstrated durations of odor recognition for maximum durations tested of six weeks (Wright et al., 2017), 69 days (Lubow et al., 1973), four months (Johnston, 1999), and 233 days (Lo et al., 2019), but long-term durations of odor memory, search performance, and generalization, which are all critical components of detection dog ability, have not been examined. Therefore, the aims of this study were to 1) examine odor memory and search performance after a 12-mo period in the absence of any odor detection training, 2) evaluate the effectiveness of single-target training in maintaining the detection of other unpracticed targets over a 12-mo period, and 3) assess effects of single-target maintenance training on generalization to untrained variants, given that extensive practice with a single target can narrow the tendency to generalize to different targets (Moser et al., 2019). Two types of settings were used for training and testing. An odor discrimination procedure using a fixed-position sampling array was used for controlled testing of odor recognition, providing defined measures of detection accuracy. Additionally, dogs were tested in a simulated operational search setting to provide an operationally relevant assessment of performance.

2. Materials and methods

2.1. Subjects

Subjects were 18 Labrador retrievers (9 M/9 F) between 1–4 years of age (mean age: 1.9 years) from the Auburn University (AU) Canine Performance Sciences (CPS) detection dog program (see Lazarowski et al. (2018) for details on population). All dogs had prior initial detector dog training with the odor of smokeless powder (SP). Dogs were individually housed in the kennel complex at the AU College of Veterinary Medicine (AUCVM). All experimental and animal care activities were approved and monitored by the Auburn University Institutional Animal and Care Use Committee in accordance with the U.S. Animal Welfare Act. The AUCVM is an Association for Assessment and Accreditation of Laboratory Animal Care International (AAALAC) accredited facility.

2.2. Odor stimuli

Odor stimuli used throughout the study consisted of odorized cotton pads (Swisspers® 100 % Cotton Rounds Pads). Stimuli were created by storing the pads inside quart (1.14 L) size glass jars along with 20−100 grams of an odorant for at least 24 h in order for the pad to absorb the odor. New pads were used each session and disposed of after one use.

Table 1 lists the target odorants trained and tested throughout the study. Distractor (i.e., non-target) odorants consisted of common products such as food/food flavorings (e.g., orange extract, onion powder), cleaners/detergents (e.g., hand soap, shampoo), household items (e.g., rubber bands, leather items), and environmental substances (e.g., wood, mulch). Stimuli for distractor odors were created and stored in the same way as the target odorants using odorized cotton pads. Additionally, cotton pads stored in glass jars with no odorant served as “matched blanks”.

2.3. Experimental areas

2.3.1. Odor discrimination task

Training and testing in the odor discrimination procedure occurred in a dedicated indoor climate-controlled facility. An eight-position odor sampling configuration in the shape of a semi-circle was arranged within an enclosure created by 1.2-m high partitions (Fig. 1). Each odor sampling position consisted of a concrete block placed on top of a wooden box, such that the top of the concrete blocks were approximately the height of the dogs’ head. For presentation of odor stimuli, the odorized cotton pads were placed, using tweezers, inside of 8 × 1 cm round metal tins with perforated lids. The tins were then placed inside of larger metal cans (9 cm diameter), which were then placed inside the concrete blocks using metal tongs. A barrier panel along the inside of the semi-circle encouraged dogs to systematically search the positions in order from first to last. The area located just outside the enclosure served as a pre- and post-trial staging area where the dog and handler remained in between trials, and where the handler and experimenter remained during trials.

2.3.2. Operational searches

Training and testing for the operational search task occurred in an indoor athletics arena on the AU campus, a structure which has an open arena concourse, offices, classrooms, laboratories, and physical plant areas. Physical items such as rags, tools, and plastic containers were used as distractors in addition to the cotton pad distractors. These materials were generally brought in from outside of the search environment, and thus provided additional odors to that environment, and were hidden in places similar to where the targets were placed.

Table 1
Set of target odorants trained and tested (left column), and related untrained odors tested for generalization (right column).

Trained targets	Related untrained targets (generalization tests)
Ammonium nitrate (AN)	Tannerite®, AN & sugar (ANS)
2,4,6-trinitrotoluene (TNT)	Cast TNT
Untagged C4	Flex-X, tagged PW4
Safety fuse (SF)	Pyrodex®
PETN-based Detonating cord (DC)	Flex-X
Untagged PE4	
Methyl benzoate (MB)	
Hydrogen peroxide (HP)	
Hexamine fuel tablets (FT)	
Vanillin/DMNB/IPE*	

* Three dogs in each group were randomly assigned to one of three targets as the 10th odor trained, which served as the maintenance target for the Odor Maintenance group. DMNB refers to the taggant compound 2,3-dimethyl-2,3-dinitrobutane, and IPE refers to a proprietary DSTL inert plastic explosive training aid material formula.

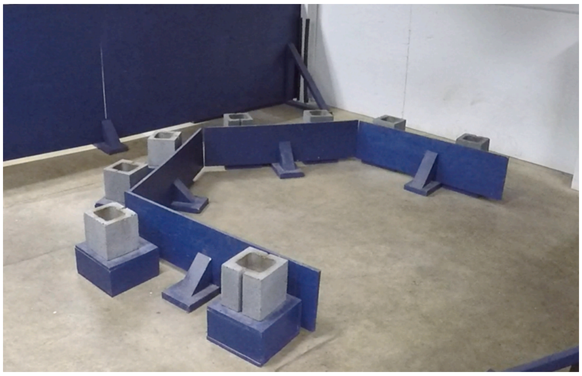

Fig. 1. Odor discrimination task setup. Dogs entered from an opening on the right side, and sampled the positions in a counter-clockwise direction.

2.4. Experimental design

The study was divided into initial training, baseline, maintenance, and test periods. Using a system similar to Porritt et al. (2015), dogs were allocated to experimental groups prior to the start of the study based on a ranking of dogs' performance. All of the dogs (n = 18) were divided evenly into three categories of low, medium, and high performance, relative to all dogs in the sample, based on the senior trainer's subjective opinion of their ability to perform operational detection tasks. From these performance categories, stratified random sampling was used to assign dogs to two experimental groups of nine dogs each (Odor Maintenance group: 5 M/4 F; No Odor Maintenance group; 4 M/5 F).

2.5. Training

2.5.1. Odor discrimination task

Dogs were trained to perform the following routine: 1) enter the enclosure upon command, 2) sample each position in order from first to last, 3) indicate (by sitting) at the position containing the target, 4) maintain the indication response until cued to exit the area, or, if no target was present, exit the arena after sampling the last position. A tone was used as an end-of-trial cue to exit the test area, regardless of whether the trial was correct or incorrect.

2.5.1.1. Pre-training. Initial training to perform the odor discrimination procedure was conducted using the odor of SP, which all dogs had equal prior experience detecting. To begin, an SP-odorized cotton pad was placed in the first position of the array. The handler encouraged the dog to investigate the first position, followed by a verbal prompt to sit. The cue was then given for the dog to exit the arena after which reinforcement was delivered in the form of a preferred toy and play with the handler. This procedure was continued with the can containing SP moved to a different position across subsequent trials and distractor odorants added to the array. Reinforcement was delivered for all correct target indications during initial training. If a dog indicated on a non-target odor, the cue was given to exit the arena but no reinforcement was delivered. Correct rejections on blank trials (i.e., sampling the entire array when no target was present without making a response to non-target positions) were reinforced with praise and handler interaction only. Missed targets resulted in absence of reinforcement.

Once dogs were sampling all positions reliably, a 70 % intermittent reinforcement schedule was implemented in order to prepare dogs for later generalization testing in which some correct responses were not reinforced. Criteria for advancing to the next phase of odor discrimination training required successfully completing a set of 12 consecutive trials (containing 10 target trials and two non-target/blank trials) with 8/10 correct responses on target trials and no false alarms (FA; responses to non-target positions). If there were more than two targets missed or a FA occurred, the sequence was reset until the criteria were met.

2.5.1.2. Training. Following pre-training with SP, dogs were trained to detect a set of 10 explosive odors (Table 1) in the odor discrimination procedure. The ten odors were trained serially, following a structured step-wise protocol whereby the location of potential target positions and the number and location of distractors increased across continuous steps in the protocol (Supplementary Table 1). Each step in the protocol was repeated until a predetermined criterion was met before moving to the next step. The terminal step of the protocol consisted of a 12-trial session, containing 10 target trials and two blank trials, with the position of targets and distractors and the sequence of target and non-target trials randomized. The criteria for having learned an odor in the odor discrimination procedure was correct identification of the target in 8/10 target trials, with no more than one FA across the 12 consecutive trials. If a dog did not meet the criteria in the minimum number of trials, the session continued with additional trials until the criteria was met. All correct target indications were reinforced during this phase. Except for initial prompting to sample when first introduced to the procedure, dogs performed the odor discrimination procedure off-leash with the handler and experimenter remaining behind the partition. Once criteria was met with a particular odor, the next odor was introduced until all 10 odors had been trained.

The 10th odor in the set was the odor to be used in subsequent maintenance training for the Odor Maintenance group. The maintenance odor was selected from three different targets, with three dogs in each group randomly assigned to one of the three. After meeting criteria with this odor, 70 % intermittent reinforcement was re-introduced. This intermittent reinforcement schedule was used during all subsequent odor discrimination maintenance training and testing.

2.5.2. Operational search

Following the baseline test in the odor discrimination task (below), dogs were trained in the operational search task. A subset of five odors randomly selected from the set of 10 (AN, TNT, SF, DC, PE4) plus the maintenance training odor were used for search training and testing. Search training began with relatively easy hides in the building environment (e.g., in a partially open cabinet drawer at dog-head height), and then extended to encompass various areas of the building and longer search durations (generally 5−15 min in length). For search training and testing, targets were always placed in separate rooms/areas (e.g., corridor, hallway), with additional blank rooms/areas interspersed throughout target areas. Each day, a different floor of the arena was used, so that the same areas were not re-used within the same week. The performance criterion in the search task for each odor was five consecutive correct indications with no FA. All operational search training and testing was conducted with the dog off leash.

2.6. Baseline testing

2.6.1. Odor discrimination task

Following odor discrimination training for all 10 odors, a baseline odor discrimination test was conducted in a separate session for each odor, in the same order in which they were trained. Each test session consisted of 10 trials containing the target odor and two trials with no target odor (blank trial) for a total of 12 trials, presented in random sequence. Target odors were placed in randomly selected positions with the exception of the first and last positions in the array to limit the common confound of dogs not sampling from the first position and an increased tendency to sit at the last possible opportunity for reinforcement (Johnen et al., 2017). However, false alarms that occurred at these positions were included in analyses of false alarm rate. At least three different distractors, one of which was a blank cotton round, were present on every trial, with empty cans in the remaining positions. On every trial at least one distractor not previously used in that session replaced one of the current distractors, and a remaining distractor was moved to a new position. Additionally, novel distractors not used at any point in the study were introduced in each of the tests.

All odor discrimination testing was conducted with the handler blind to the presence and locations of potential targets. On each trial, the experimenter arranged the trial stimuli and then placed an index card noting the trial condition (target or blank) and target position face down on a table behind the arena partition, and moved behind the partition out of sight of the dog. The handler then prompted the dog to enter the enclosure and observed the trial from behind a partition. If the dog responded at any position or reached the end of the array without responding, the handler gave the trial termination cue and flipped over the card. If correct, the handler delivered reinforcement according to the intermittent reinforcement schedule. If a FA occurred on a position prior to the dog encountering the target position, the trial was repeated at the end of the session until the dog had reached 10 target odor exposures (treated as an additional trial). If the response was incorrect, no reinforcement was delivered. Correct rejections of blank trials resulted in handler praise but no primary reinforcement. Dogs were free to go back and forth between positions, and correct rejections were recorded once the dog searched the last position for any positions in which the dog did not sit.

2.6.2. Operational search

Following completion of the search training, the operational search baseline test was conducted. Each odor was tested in a separate session in which nine samples of the target odor were hidden throughout the search areas. Blank and distractor cotton pads, as well as random items, were planted along the search path such that a minimum of one cotton pad distractor and one physical item distractor were encountered prior to each target.

At the start of each search test, the senior trainer informed the handler of the parameters of the search area and served as the test evaluator, remaining behind the handler and dog throughout the search. Searches were conducted off-leash with the handler guiding the dog to search the prescribed areas in a thorough manner. Handlers indicated when they observed the dog perform the trained indication response (sitting), after which the evaluator informed the handler as to whether or not the response was correct. If correct, the handler rewarded the dog and the response was scored as a hit. If incorrect, the handler prompted the dog to continue searching and a false alarm was recorded. To minimize the number of dogs searching the same areas consecutively and to maximize the number of searches that could be performed with a blind handler, two handlers alternated handling dogs and performed the searches in a rotation so that each dog started at a different location than the dog prior.

2.6.3. Generalization

Following completion of search testing for the trained targets, dogs' responses to six untrained variants of the trained odors were probed in generalization tests in both the odor discrimination and search settings (see Table 1 for list of generalization probe odors and corresponding trained target). Generalization testing was first conducted in the odor discrimination procedure. Each probe odor was presented once within a five-trial block consisting of four trials with a trained target and one trial with the probe odor, presented in a random sequence with the exception that the first trial contained a trained target and not a probe. The target odor used in the block of trials was one unrelated to the probe odor. To evaluate whether potential responses to the probe odors was controlled by perceptual similarity of the targets rather than the novelty of the odor, a new novel distractor was introduced on each of the target (non-probe) trials in a position prior to the target position such that dogs encountered a total of five novel distractors for each probe odor. Correct responses on probe trials were not reinforced. Responses to the non-probe targets were intermittently reinforced at a pre-determined schedule of 75 % reinforcement (i.e., one non-reinforced target per block).

Following generalization testing in the odor discrimination procedure, each probe odor was also presented in the operational search setting. On each search trial, dogs first encountered an unrelated trained target followed by the probe odor, with distractors placed throughout. If the dog missed the trained target odor, the handler was notified and assisted the dog to ensure it encountered and alerted to the target odor before moving on to the probe odor. Responses to probe odors were not reinforced.

2.7. Maintenance training

For the 12 months following the completion of all baseline testing, dogs in the Odor Maintenance group engaged in intermittent maintenance training using a single odor from the trained set (Odor #10; Table 1). Following Porritt et al. (2015), this odor was one of three non-explosive odors (vanillin, 2,3-dimethyl-2,3-dinitrobutane [DMNB], and IPE), with three dogs assigned to each of the three odors. The three maintenance odors were chosen based on dissimilarity to the trained targets, volatility range, and utility for operational use, and the use of three different maintenance targets ensured that any effects could not be attributed to characteristics of one particular odor (Porritt et al., 2015).

Maintenance training followed a schedule which cycled every four weeks. Week 1: each dog received four trials in the odor discrimination task consisting of three trials with their assigned maintenance target and one blank trial, presented in random sequence, with intermittent reinforcement of target indications; Week 2: no detection training; Week 3: each dog performed three searches for their assigned maintenance target in the search task, with all correct indications reinforced; Week 4: no detection training. Dogs received exercise in outdoor runs and an obstacle course area at least two days per week during training weeks (on non-training days), and approximately four days per week during weeks without maintenance training.

The No Odor Maintenance group was transferred to a partnering correctional institution dog program for the duration of the 12-mo period during which they received no odor detection training. Instead, to equate the groups with respect to general activity and participation in reward-based training, dogs were taught to perform a number of different obedience and assistance-type tasks such as sit, stay, going to and staying on a placemat, kenneling on command, and touch-stick targetting. One 4-hr training session was conducted each week (divided across the nine dogs). The remainder of the time, the dogs received daily exercise and socialization.

2.8. Recall testing

Twelve months following baseline testing, all tests were repeated using the same procedures used in baseline. Odor discrimination testing was conducted first with a subset of the 10 target odors (C4, FT, HP, MB). Order of testing each odor was randomized, but was consistent between groups. Prior to the test trials, both groups were given one blank trial (i.e., no target present) as a refresher on performing the sampling procedure. This was done to minimize possible issues related to performing the procedure rather than memory of the odor for the No Odor Maintenance group. Prompting to sample and move from position to position was provided as needed during this refresher trial.

Next, the baseline operational search task was repeated. In order to re-acclimate the No Odor Maintenance group to the operational search context, the dogs were taken on walks in areas similar to, but not the same as, the search testing area. Finally, dogs were again tested for generalization in the odor discrimination procedure, followed by search using the same procedures as the baseline generalization tests.

2.9. Statistical analysis

For odor discrimination testing, dogs' responses on each trial were recorded as hits (responses to target odor), FA (responses to non-target odors), misses (no response to target odor), and correct rejections (no response to each position not containing a target odor). Overall hit rate was then calculated as total hits out of total target exposures, averaged across the subset of four odors. FA rate was calculated as total FA out of total opportunities for a FA (FA + correct rejections). Because sensitivity (hit rate) and specificity (absence of FA) are equally important to a detection dog's overall accuracy, we also calculated positive predictive value (PPV) which is a measure that combines hits and false alarms, calculated as the proportion of a dog's responses that are correct (hits/hits + FA) (Simon et al., 2018). Effects of maintenance training on average hit rate, FA rate, and PPV were analyzed using Linear Mixed Models (LMMs) (lme4 package; Bates et al., 2015), with Maintenance (binary variable: yes or no), Time (baseline or 12-mo), and their interaction as fixed factors. Order of test session (1–4) was included as a fixed factor to assess potential learning effects across test sessions. Sex was included as a fixed factor and subject ID as a random factor to control for effects of repeated measures. For generalization testing, we used a Generalized Linear Mixed Model (GLMM) fit by maximum likelihood using the binomial family distribution with hit (yes or no) to each of the probes as the dependent variable.

The same measures were used for search testing, with the exception of FA rate due to the inability to quantify correct rejections needed for its calculation (i.e., the opportunities for FA in an operational search are infinite/indeterminate) (Helton, 2009). GLMMs were run for search testing using the binomial family distribution (lme4 package; Bates et al., 2015), with hit (yes or no) as the dependent variable, and the additional fixed factor of whether or not the handler was blind on the given trial (binomial: yes or no). LMMs were used to assess PPV in search testing. Analyses were performed in the R statistical program (RStudio Version 1.2.5033).

3. Results

3.1. Recall

3.1.1. Odor discrimination task

Average hit rate for the No Odor Maintenance group declined from 100 % ($SE = 0$) at baseline to 85.83 % (SE = 4.08) after the 12-mo period, and for the Odor Maintenance group declined from 96.11 % ($SE = 1.21$) to 92.22 % ($SE = 1.91$). There was a significant effect of Time, where average hit rate in the 12-mo test was significantly lower than the baseline test (LMM: $t = 4.27$, $p < 0.001$), and a significant interaction between Time and Maintenance ($t = -2.19$, $p = 0.03$) in which hit rate was significantly lower in the 12-mo test compared to baseline for the No Odor Maintenance group ($t = 3.48$, $p < 0.001$), but did not significantly change for the Odor Maintenance group ($t = 1.69$, $p = 0.09$). Hit rate in the 12-mo test for dogs in the Odor Maintenance group was 94 % ($SE = 3.64$) for dogs trained with vanillin, 91 % ($SE = 3.51$) for IPE, and 97 % ($SE = 1.85$) for DMNB (differences between maintenance odor were not formally analyzed due to the small number of dogs in each subgroup). Sex, order, and maintenance did not significantly affect hit rate ($ps > 0.06$).

Average FA rate for the No Odor Maintenance group increased from 0.05 % ($SE = 0.06$) at baseline to 7.62 % ($SE = 1.37$) at the 12-mo test, and decreased for the Odor Maintenance group from 1.36 % ($SE = 0.34$) to 0.15 % ($SE = 0.09$). There was a significant effect of time ($t = -7.56$, $p < 0.001$) and maintenance ($t = -5.96$, $p < 0.001$) and a significant interaction between the two, where FA rate significantly increased from baseline to 12-mo for the No Odor Maintenance group ($t = -5.51$, $p < 0.001$), and significantly decreased from baseline to 12 mo for the Odor Maintenance group ($t = 3.46$, $p = 0.001$). Sex and order did not significantly affect FA rate ($ps > 0.47$).

PPV (proportion of all responses that are correct) was calculated to obtain a more complete measure of accuracy taking into account both hits and false alarms. PPV was significantly affected by Time (LMM: $t = 8.56$, $p < 0.001$) and Maintenance ($t = 7.19$, $p < 0.001$) with a significant interaction between the two where PPV significantly decreased from 99.74 % ($SE = 0.25$) to 72.40 % ($SE = 4.36$) for the No Odor Maintenance group ($t = 6.49$, $p < .001$) but significantly increased from 93.53 % ($SE = 1.57$) to 99.07 % ($SE = 0.53$) for the Odor Maintenance group ($t = -3.36$, $p = 0.001$) (Fig. 2). There was no effect of sex or order on PPV ($ps > 0.46$).

3.1.2. Operational search

The initial GLMM revealed a significant effect of whether or not the handler was blind on search performance (GLMM: $z = -2.198$, $p = 0.023$), where number of hits was significantly lower on blind trials than non-blind trials. Therefore, subsequent analyses were run using only data from trials in which the handler was blind.

There was a significant three-way interaction between time, maintenance, and order on number of correct responses in search testing ($z = -1.98$, $p = 0.04$). Further analysis revealed a significant effect of order for the Odor Maintenance group at baseline, where hits increased across searches ($z = 2.28$, $p = 0.02$), but no such effect for the No Odor Maintenance group ($z = 0.89$, $p = 0.36$). Conversely, in the 12-mo test, hits increased across searches for the No Odor Maintenance group ($z = 2.41$, $p = 0.01$), but no such effect for the Odor Maintenance group ($z = 0.41$, $p = 0.684$).

There was also a significant two-way interaction between time and maintenance on hits ($z = 2.01$, $p = 0.04$), where the decrease in hits across time was significant for the Odor Maintenance group (84.70 % to 68.75 %) ($z = 2.45$, $p = 0.01$), but not for the No Odor Maintenance group (78.16%–66.66%) ($z = 1.799$, $p = 0.23$) (Fig. 3). There were no main effects of time ($z = 0.06$, $p = 0.95$), maintenance ($z = 0.05$, $p = 0.96$), order ($z = -0.22$, $p = 0.83$), or sex ($z = 0.01$, $p = 0.99$) on hits in search testing. Hit rate in the 12-mo test for dogs in the Odor Maintenance group trained with vanillin was 73 % ($SE = 6.12$), 76 % ($SE = 5.44$) for IPE, and 72 % ($SE = 4.18$) for DMNB (differences between maintenance odor were not formally analyzed due to the small number of dogs in each subgroup).

There was a significant effect of time on PPV, where PPV was significantly higher at baseline than at 12 months (LMM: $t = 4.05$, $p < .001$). There was no significant effect of maintenance ($z = 1.47$, $p = 0.16$), no interaction between the two ($z = -1.24$, $p = 0.23$), and no effect of sex on PPV ($t = -.48$, $p = .63$).

3.2. Generalization

3.2.1. Odor discrimination task

Generalization rate (total responses to probes out of total number of

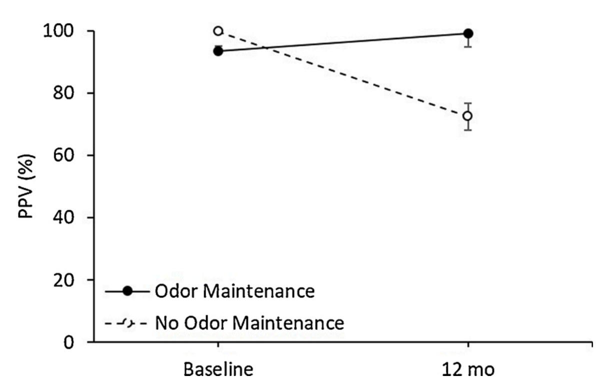

Fig. 2. Mean PPV (proportion of responses that are correct) in the baseline and 12-mo odor discrimination test for the Odor Maintenance (solid line, n = 9) and No Odor Maintenance (dashed line, n = 9) groups, depicting a significant interaction where PPV decreased for the No Odor Maintenance group and increased for the Odor Maintenance group ($p < .05$). Error bars indicate standard error of the mean.

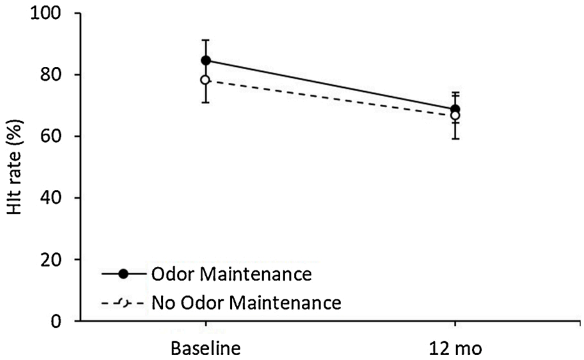

Fig. 3. Mean hit rate (total hits/total targets presented) in the baseline and 12-mo blind search test trials for the Odor Maintenance (solid line, n = 9) and No Odor Maintenance (dashed line, n = 9) groups, depicting a significant interaction where the decrease in hits across time was significant for the Odor Maintenance group not for the No Odor Maintenance group ($p < .05$). Error bars indicate standard error of the mean.

probes tested) in the odor discrimination test decreased from 96.29 % (0.09 % FA) at baseline to 27.78 % (3.79 % FA) in the 12-mo test for the No Odor Maintenance group, and from 88.89 % (.59 % FA) to 55 % (.87 % FA) for the Odor Maintenance group (Fig. 4). Generalization rate was significantly affected by time ($z = 3.68$, $p < 0.001$) and maintenance ($z = -2.51$, $p = 0.02$) and there was a significant interaction between the two ($z = 2.55$, $p = 0.01$), where generalization rate did not differ between groups at baseline ($z = 1.11$, $p = 0.26$), but was significantly higher in the 12-mo test for the Odor Maintenance group compared to the No Odor Maintenance group ($z = -2.71$, $p < 0.001$). Sex did not significantly affect generalization rate ($p = .74$).

3.2.2. Operational search

In search testing, generalization rate decreased from 88.89 % to 51.85 % for the No Odor Maintenance group and from 81.48%–72.22% for the Odor Maintenance group (Fig. 5). There was a significant interaction between time and maintenance, ($z = 2.19$, $p = 0.03$), where generalization significantly decreased over time for the No Odor Maintenance group ($z = 3.11$, $p = 0.002$), but not for the Odor Maintenance group ($z = 1.38$, $p = 0.16$). There were no significant effects of blind handler ($p = 0.78$), sex ($p = 0.48$), maintenance ($p = 0.06$), or time ($p = 0.23$) on generalization rate ($p < 0.06$). PPV was above 87 % for both groups at each time point indicating that responses to probes were controlled by the target odor rather than a lack of specificity.

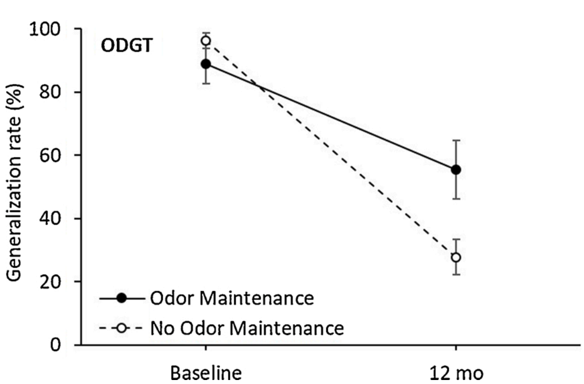

Fig. 4. Mean percentage of responses to the six probe odors in the baseline and 12-mo odor discrimination generalization test (ODGT) for the Odor Maintenance (solid line, n = 9) and No Odor Maintenance (dashed line, n = 9) groups, depicting a significant interaction where generalization rate was higher for the Odor Maintenance group compared to the No Odor Maintenance group only in the 12-mo test ($p < .05$). Error bars indicate standard error of the mean.

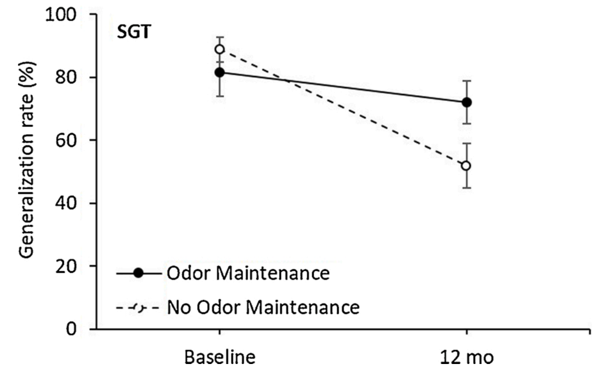

Fig. 5. Mean percentage of responses to the six probe odors in the baseline and 12-mo search generalization test (SGT) for the Odor Maintenance (solid line) and No Odor Maintenance (dashed line) groups, depicting a significant interaction where generalization significantly decreased over time for the No Odor Maintenance group, but not for the Odor Maintenance group ($p < .05$). Error bars indicate standard error of the mean.

4. Discussion

The present study examined the effects of single-odor maintenance training on detection dog long-term odor memory and search performance. Two groups of nine dogs were initially trained to alert to 10 target odors. Over the course of the next 12 months, one of the odors served as a maintenance training aid and was used in intermittent maintenance training for one of the two groups while the other group did not engage in any odor detection work. All dogs were then tested for accuracy in recalling the remaining nine odors in both controlled odor recognition trials and operational searches. Additionally, all of the dogs were tested for the generalization to untrained odors that had some presumed similarity to their trained target odors before and after the 12-mo period.

First, results from the No Odor Maintenance group demonstrating only a slight decline in odor discrimination accuracy after no odor detection training for 12 months serves as an examination of canine long-term odor memory. Our results extend previous findings to demonstrate relatively high recognition of odors (85 %) in the absence of any exposure after one year, indicating that dogs' long-term odor memory, as assessed in an odor discrimination procedure, is rather robust and resilient to the passage of time.

Second, regarding the primary objective of this study, odor memory was shown to be effectively maintained over a 12-mo period by participating in relatively infrequent minimal (approximately 30 min per month) maintenance training with a single odor. Recall in the odor discrimination test by this group showed only a slight decline of 4%, compared to dogs that received no odor maintenance training for which accuracy declined by 15 %. Moreover, maintenance training appeared to improve overall accuracy (combined hits and false alarms) whereas specificity declined (i.e., false alarms increased) for the group that did not engage in odor maintenance training. The very high PPV of the Odor Maintenance group after the 12-mo period (99 %) in comparison to that of the No Odor Maintenance group (72 %) suggests that the maintenance training not only effectively facilitated recall of the originally trained odors, but also enhanced dogs' ability to discriminate target odors from non-target odors, an equally important aspect of detection dog performance. Importantly, the effectiveness of an innocuous, unrelated odor in maintaining memory for a set of explosives suggests the utility of the maintenance odor is not constricted to the nature of the odor itself. From an operational perspective, this illustrates a very practical beneficial effect of intermittent training with a single odor in maintaining odor recognition accuracy. This finding is also consistent with, and extends the time horizon of the results of, Porritt et al. (2015) and Thrailkill et al. (2018) regarding the utility of single-target training in maintaining detection performance. Mechanisms of how the maintenance odor functions to maintain performance proposed by Thrailkill et al. (2018) include formation of a stimulus class based on equal outcomes (i.e., reinforcer) of the maintenance and other target odors, and stimulus control due to learning about shared and distinct features of the stimuli. However, the authors demonstrated that the effect of the maintenance odor is not due to simple stimulus generalization based on perceptual similarity, further exemplified by the effectiveness of very different maintenance odors in our study.

These results add to the evidence that the capacity for long-term memory for odors is very robust. The exceptionality of odor memory is likely due to distinct features of the olfactory system relative to the other senses, such as neuroanatomical pathways between the olfactory cortex and portions of the brain involved in emotion and memory processing, and direct projections to higher cortical regions of the brain bypassing sensory integration through the thalamus (Herz and Engen, 1996). In humans, studies have shown that odors are remembered far longer than visual or verbal stimuli and are highly resistant to forgetting and interference (Engen and Ross, 1973; Lawless and Engen, 1977; Lawless and Cain, 1975; Rabin and Cain, 1984; Roediger et al., 2017). Chemical sensing in humans can reasonably, through general behavioral observation and relative diminution of olfactory related neuroanatomical features, be considered to play a less important role in interfacing with the environment than for dogs. Macrosomatic animals for which chemical sensing plays a prominent role in guiding their behavior, such as rodents and canines, may possess particularly robust memory capacities and recall capabilities for odors that are considerably more robust than that of humans. While the current study has provided evidence that dogs are able to remember odors for periods up to one year, the extents and limits of long-term odor memory duration are still unknown, but is likely to endure beyond the duration tested here.

In contrast to recognition accuracy in the controlled odor discrimination test, the minimal maintenance training did not appear to effectively maintain operational search performance. This could partially be an artifact of the improvement across trials in the baseline test for the Odor Maintenance group. Thus, baseline performance may have been elevated for this group, resulting in a steeper decline relative to the other group that started out lower. In the future, this artifact could be avoided by allocating the groups after baseline testing based on performance to ensure groups are as equal as possible at the start. Similarly, the lack of difference between groups could have been due to the same apparent practice effects in the 12-mo test for the No Odor Maintenance group, for which performance improved across trials, potentially masking effects of the maintenance training. Regardless, declines of only 11 % and 15 % for the Odor Maintenance and No Odor Maintenance groups, respectively, combined with high accuracy in the odor discrimination test, are more indicative of search performance decline rather than forgetting of the odors. Nevertheless, further research is needed to examine the various factors related to maintenance training, such as the frequency and schedule of the training, and their effects on maintaining search performance.

Another potential reason for the lack of effect in the search task noted by the handlers was an apparent dependence on the handlers for guidance developed by dogs in the Odor Maintenance group. That is, handlers observed that during the 12-mo search testing, dogs in this group often paused and looked back at the handlers during searches. Recent research has shown that detection dogs look back to the handler during blind searches more often than in searches where the handler is not blind, indicative of dogs seeking guidance when faced with a challenge (DeChant et al., 2020). Search maintenance training was not blind, and because the goal was to ensure dogs were exposed to and alerted to each of the targets per search session, handlers may have used body position or other cues to help bracket the location of the targets, cues which were not present during the blind search tests. Therefore, it is important to consider the manner in which maintenance search training is conducted, with the goal that maintenance training matches test or operational scenarios as much as possible. The importance of blinding handlers to the location of targets was also highlighted by the difference in search accuracy between blind and non-blind trials, where dogs' performance was lower on searches that were run single-blind compared to searches where the handler was aware of the location of targets. While likely unintentional, this probably reflected handlers ensuring dogs were allowed sufficient time to search an area when it was known to contain a hide, and may have been less thorough in unknown areas where handlers may have incorrectly believed no target was present (DeChant et al., 2020).

While generalization remained higher for the Odor Maintenance group relative to the No Odor Maintenance group after the 12-mo period in the odor discrimination test, generalization declined significantly over time for both. In the operational search generalization test, responses to the variants decreased significantly over time for the No Odor Maintenance group, but not the Odor Maintenance group. Thus, the maintenance training appeared to sustain the propensity of dogs to perform their trained alert response to the untrained variants and inhibit the decline in generalization to a greater extent in the operational search task. This difference in generalization between the odor discrimination and search tests may reflect greater specificity in the discrete sampling odor discrimination context verses the more dynamic context of a search. However, though the decrease in generalization in the search test was not statistically significant, the roughly 10 % decline may be of operational significance due to the importance of generalization for optimal detection performance as detection dogs are likely to encounter targets that are similar but not identical to that which they were trained. Taken together, decreases in generalization following single-odor maintenance training could either reflect a narrowing effect of training with a single odor (Cleland et al., 2009), or simply a decay due to the passage of time. However, it should be noted that the robustness of the generalization testing results is limited due to the small number of observations, which was necessary to minimize effects of repeated exposure. These results warrant further investigation of methods for maintaining generalization.

5. Conclusions

The relatively small decline in recognition of odors in the absence of their exposure, or any detection training, over 12 months suggests that dogs have a remarkable ability to remember odors to which they have been trained to alert over long periods of time. Further, the use of a single maintenance training aid improved odor discrimination accuracy, though the same results were not observed in an operational search task. Generalization to untrained but similar odors, a critical aspect of the operational success of a detection dog, was reduced over the 12-mo period without contact with the related targets but to a lesser extent for dogs that received intermittent training with the single maintenance odor.

The current results have implications for prioritizing the components of maintenance training of operational detection dogs, suggesting that with proper initial training and validation of accurate odor recognition, dogs are capable of remembering odors for extended periods of time. This places the emphasis for expending resources on maintaining search performance rather than frequent odor recognition type trials to assure basic odor recognition, which appears to be resilient to the passage of time. The results also hint toward generalization being affected by the dwell time without exposure to similar target odors, and indicate the need for further examination of generalization due to the importance of this phenomenon in detection of variations of specific targets. This study has defined some broad parameters for odor memory over a 12-mo period to guide further detailed examination of procedures for optimizing odor detection performance by dogs.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgments

This study was supported by the Office of Naval Research grant # G00010016 and the U.K. Ministry of Defence contract # UKDPO/00009. The funding bodies were not involved in experimental design, data analysis, interpretation or article preparation. We thank CPS canine instructors Terry Fischer, Jenn Jankiewicz, and Laura Reeder for conducting canine training and testing activities, and Holli Thompson for managing the care and welfare of the dogs. We are very grateful to the participants in the Auburn dog development program and management at Lake City Correctional Facility for their ongoing collaboration that facilitated the execution of this study.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.applanim.2021.105301.

References

Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67 (1) https://doi.org/10.18637/jss.v067.i01.

Cleland, T.A., Narla, V.A., Boudadi, K., 2009. Multiple learning parameters differentially regulate olfactory generalization. Behav. Neurosci. 123 (1), 26–35. https://doi.org/10.1038/jid.2014.371.

DeChant, M.T., Ford, C., Hall, N.J., 2020. Effect of handler knowledge of the detection task on canine search behavior and performance. Front. Vet. Sci. 7. https://doi.org/10.3389/fvets.2020.00250.

Engen, T., Ross, B.M., 1973. Long-term memory of odors with and without verbal descriptions. J. Exp. Psychol. 100 (2), 221–227. https://doi.org/10.1037/h0035492.

Furton, K.G., Myers, L.J., 2001. The scientific foundation and efficacy of the use of canines as chemical detectors for explosives. Talanta 54 (3), 487–500. https://doi.org/10.1016/S0039-9140(00)00546-4.

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

Hall, N.J., Wynne, C.D.L., 2016. Canine olfactory learning and behavior. Canine Olfaction Science and Law, pp. 123–137.

Helton, W.S., 2009. Overview of scent detection work: issues and opportunities. In: Helton, W.S. (Ed.), Canine Ergonomics: The Science of Working Dogs. CRC Press, Boca Raton, FL, pp. 83–96.

Herz, R.S., Engen, T., 1996. Odor memory: review and analysis. Psychon. Bull. Rev. 3 (3), 300–313. https://doi.org/10.3758/BF03210754.

Johnen, D., Heuwieser, W., Fischer-Tenhagen, C., 2017. An approach to identify bias in scent detection dog testing. Appl. Anim. Behav. Sci. 189, 1–12. https://doi.org/10.1016/j.applanim.2017.01.001.

Johnston, J.M., 1999. Canine Detection Capabilities: Operational Implications of Recent R&D Findings.. Retrieved from. http://webcanine.com/wp-content/uploads/2007/04/k-9_det_capabilities.pdf.

Lawless, H.T., Cain, W.S., 1975. Recognition memory for odors. Chem. Senses 1 (3), 331–337. https://doi.org/10.1093/chemse/1.3.331.

Lawless, H., Engen, T., 1977. Associations to odors: interference, mnemonics, and verbal labeling. J. Exp. Psychol. Hum. Learn. 3 (1), 52–59. https://doi.org/10.1037/0278-7393.3.1.52.

Lazarowski, L., Haney, P., Brock, J., Fischer, T., Rogers, B., Angle, C., Katz, J.S., Waggoner, L.P., 2018. Investigation of the behavioral characteristics of dogs purpose-bred and prepared to perform vapor wake® detection of person-borne explosives. Front. Vet. Sci. 5, 50. https://doi.org/10.3389/fvets.2018.00050.

Lo, G.K.-H., Macpherson, K., MacDonald, H., Roberts, W.A., 2019. A comparative study of memory for olfactory discriminations: Dogs (Canis familiaris), rats (Rattus norvegicus), and humans (Homo sapiens). J. Comp. Psychol. 134 (2), 170–179. https://doi.org/10.1037/com0000205.

Lubow, R.E., Kahn, M., Frommer, R., 1973. Information processing of olfactory' stimuli by the dog I. The acquisition and retention of four odor-pair discriminations. Bull. Psychon. Soc. 1 (2), 143–145.

Moser, A.Y., Bizo, L., Brown, W., 2019. Olfactory generalization in detector dogs. Animals 9 (9), 702. https://doi.org/10.3390/ani9090702.

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. Appl. Anim. Behav. Sci. 166 (1), 112–122. https://doi.org/10.1016/j.applanim.2015.02.013.

Rabin, M.D., Cain, W.S., 1984. Odor recognition: familiarity, identifiability, and encoding consistency. J. Exp. Psychol. Learn. Mem. Cogn. 10 (2), 316–325. https://doi.org/10.1037//0278-7393.10.2.316.

Roediger, H., Zaromb, F.M., Lin, W., 2017. A typology of memory terms. In: Byrne, J.H. (Ed.), Learning and Memory: A Comprehensive Reference, 2nd ed. Academic Press, pp. 7–19.

Simon, A.G., Degreeff, L.E., Peranich, K., Furton, K.G., 2018. Canine Generalization to Molecularly Similar Odors and Odor Mixtures. NRL/MR/6181–18-9797.

Simon, A., Lazarowski, L., Singletary, M., Barrow, J., Van Arsdale, K., Angle, T., et al., 2020. A review of the types of training aids used for canine detection training. Front. Vet. Sci. 7 (June), 1–10. https://doi.org/10.3389/fvets.2020.00313.

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. Behav. Processes 157 (August), 161–170. https://doi.org/10.1016/J.BEPROC.2018.09.012.

Troisi, C.A., Mills, D.S., Wilkinson, A., Zulch, H.E., 2019. Behavioral and cognitive factors that affect the success of scent detection dogs. Comp. Cogn. Behav. Rev. 14, 51–76. https://doi.org/10.3819/ccbr.2019.140007.

Wright, H.F., Wilkinson, A., Croxton, R.S., Graham, D.K., Harding, R.C., Hodkinson, H.L., et al., 2017. Animals can assign novel odours to a known category. Sci. Rep. 7 (1), 3–8. https://doi.org/10.1038/s41598-017-09454-0.