Applied Animal Behaviour Science 247 (2022) 105568

Contents lists available at ScienceDirect

Applied Animal Behaviour Science

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

Effects of learning an increasing number of odors on olfactory learning, memory and generalization in detection dogs

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

[a] Canine Performance Sciences, College of Veterinary Medicine, Auburn University, AL, USA
[b] Defence Science and Technology Laboratory, Porton Down, Salisbury, Wiltshire, UK

ARTICLE INFO

Keywords:
Detection dogs
Odor memory
Odor discrimination
Generalization

ABSTRACT

Detection dogs continue to represent the most effective technology for the detection of explosives and other contraband. As the number of target odors and their variants that detection dogs are required to learn continues to increase, understanding the capacity for odor learning and the effects on performance of training on an increasing number of target odors on performance is critical. In the present study, we examined dogs' ability to learn an increasing number of odor discriminations and the effects of learning new odors on the recall of previously trained odors. We trained dogs to respond to up to 40 odors in an odor discrimination task over the span of 16 months. Odors were trained in sets of 10 every three months. Recall of previously trained subsets was assessed at intervals of < 1, 4, and 12 months since last exposure, and after learning 10, 20, 30, and 40 cumulative odors. We assessed the effects of training on these large numbers of target odors on generalization to untrained, but related, target odors. In addition to assessing performance in an odor discrimination task, we also evaluated the effects of learning many odor discriminations and the passage of time on performance in an operationally relevant search task. Dogs rapidly learned new odor discriminations across the 40 odors, with little to no decrease in the recall of previously trained odors or generalization to related odors. Further, dogs recalled odors not experienced within 12 months with 100% accuracy and no increase in false alarm rate. The results indicate that the limits of odor identification recall capabilities of dogs were not challenged by training on a cumulative total of 40 odor discriminations nor by up to a 12-month gap in exposure to those odors, therefore establishing the robust capabilities of dogs for learning and remembering many odor identifications.

1. Introduction

Detection dogs are widely considered the most capable and flexible technology for the detection of explosives and other contraband (Furton and Myers, 2001). However, unlike other detection technologies for which performance characteristics are understood through their fundamental physics, chemistry, and electronics, there is only a relatively nascent understanding of the underlying operational characteristics of detection dogs (Hayes et al., 2018). Indeed, there are significant knowledge gaps in defining the basic biological, cognitive, and behavioral parameters of detection dog operational performance (Hall and Wynne, 2016; Troisi et al., 2019). Bridging such gaps is critical for enhancing detection dog capabilities by improving preparation, training, and employment techniques.

The tasks of explosive detection dogs are increasingly specialized and sophisticated making their required performance repertoire more complex and challenging. The number of explosives and the different variants of those explosives that detection dogs need to identify continues to increase. Thus, a critical aspect of canine explosives detection is the ability to learn and remember an increasing number of target odors, but few studies have explored dogs' cognitive and behavioral capacities for cataloging, recalling, and responding to trained odors. Williams and Johnston (2002) sequentially trained dogs on an increasing number of odor discriminations, assessing rate of learning for each new odor and recall for previously trained odors as new odors were introduced. Fewer training trials were required to learn additional odors, with no decrease in recognition of previously learned odors or any increase in false alerts across the total of 10 odors trained in the study. The number of targets on which detection dogs are typically trained can be anticipated to continue increasing over time, as new threats or illicit substances are identified for surveillance. Whether there is a capacity for the number of odors dogs are capable of learning, at

* Correspondence to: College of Veterinary Medicine, Auburn University, 104 Greene Hall, AL 36849, USA.
E-mail address: waggolp@auburn.edu (P. Waggoner).

https://doi.org/10.1016/j.applanim.2022.105568
Received 2 November 2021; Received in revised form 12 January 2022; Accepted 24 January 2022
Available online 29 January 2022

which point the ability to learn and retain new odors is compromised (e.g., proactive interference), or whether learning new odors interferes with recall of previously learned odors (e.g., retroactive interference) has not been further explored in dogs.

In the practical application of detection dogs, it is also desirable that recognition of, and responses to, trained target odors generalizes to untrained variants of those target odors with similar properties (Moser et al., 2019). In most studies of odor recall, investigation has been limited to subjects' responding to trained odors and not responding to all other odors (cf. Wright and Smith, 2004 and Hall et al., 2016 for contrary examples). Extensive training with the same odors tends to produce increased discrimination of the trained odor, but also results in a narrowed generalization to chemically similar odorants (Moser et al., 2019). However, effects of training on an increasing number of different (i.e., largely chemically unrelated) odors on the narrowing or broadening of generalization gradients, and how this trend may change as time passes since exposure to trained target odors, has not been previously investigated.

Another consideration regarding odor memory with operational relevance to detection dog performance is how long trained target odors are maintained in long-term memory. Early work by Lubow et al. (1973) demonstrated that dogs trained on several odor discriminations exhibited high levels of retention for the discriminations after a period of up to 69 days. Other reports have indicated that dogs are capable of maintaining high response accuracies for maximum tested intervals of four months (Oxley and Waggoner, 2009). There is also evidence that intermittent training with any individual odor within a set maintains detection performance for the entire set of previously trained odors, even across extended periods of no exposure to, or training with, the other odors of that set (Porritt et al., 2015; Thrailkill et al., 2018). We recently extended these findings to demonstrate that dogs are capable of recalling a set of 10 odors after a period of 12 months with infrequent maintenance training using one odor from the set throughout that period (Lazarowski et al., 2021). However, effects of potential interference of learning intervening odors on long-term memory have not been explored in dogs (Hall and Wynne, 2016).

The objective of the present study was to investigate the performance characteristics of detector dogs in learning and remembering a large number of different odor discriminations over extended periods of time. Using a longitudinal, repeated measures experiment, we assessed the effects of training dogs to identify increasing numbers of odors on their ability to learn new odors, recall previously learned odors, and generalize to untrained related odors. Given evidence that there may be especially robust memory capacity for odors, we hypothesized that dogs would demonstrate high levels of accuracy in identifying previously trained target odors despite intervening training with increasing odors, as well as extended passages of time without exposure to those odors.

2. Materials and methods

2.1. Subjects

Nine experimentally naive adult Labrador Retrievers (4F/5M; mean age: 2.86 yrs) from Auburn University's Canine Performance Sciences detection dog program participated in this study. All dogs had prior initial detector dog training to identify and search for the odor of smokeless powder (SP) in buildings and outside areas. Dogs were housed in individual runs within the kennel complex at the Auburn University College of Veterinary Medicine (AUCVM). All animal use and care was 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 scented cotton pads, created by storing the pads in glass jars along with an odorant (see Lazarowski et al., 2021 for detailed description). Table 1 lists the target odorants trained throughout the study. Distractor (i.e., non-target) stimuli were created in the same way and consisted of common products such as food/food flavorings, cleaners/detergents, construction materials, rubber bands, plastics, leather items, and environmental substances such as mulch and wood. Additionally, cotton pads stored in glass jars with no odorant served as "matched blanks" relative to the target odorants. By the first testing event (see experimental design below), there were 45 distractors in use in continual rotation throughout trials. At each of the subsequent tests, at least five additional novel distractors were added to the rotation. Additionally, for the generalization tests, five novel distractors which had not been previously used were added to the rotation of total distractors. In sum, 72 distractor odors were used in the study.

2.3. Experimental settings

Two types of settings were used for canine detection training and testing (see Lazarowski et al., 2021). Briefly, an odor discrimination task using a discrete-trials fixed-sampling method was used to allow for controlled testing of odor recognition and defined measures of detection accuracy (Helton, 2009). Second, dogs were also trained and tested in real-world conditions (i.e., building searches) for operational relevance.

The setup for the odor discrimination task (ODT) activities has been previously described (Lazarowski et al., 2021). Briefly, eight stimulus positions were arranged in the shape of an arc inside an arena enclosed by 1.2-m high panel walls. Each position consisted of a concrete block on top a wooden box, such that the concrete blocks were approximately dog-head height. A scented cotton pad was placed inside a round metal tin with a perforated lid, which was then placed inside a metal can. The can was then placed inside a concrete block. During each trial, the handler and experimenter remained outside of the enclosure.

The setting for operational searches was the same used in Lazarowski et al. (2021), which consisted of a large indoor sports coliseum containing an open arena concourse, offices, classrooms, laboratories, and physical plant areas which were used as search areas. For searches, in addition to distractor odors absorbed onto cotton pads and blank cotton pads, items such as rags, tools, and plastic containers were used as distractors to provide additional odors to that environment hidden in places similar to where the cotton pads with adsorbed target and distractor odorants were placed.

2.4. Experimental design

The general experimental design was an alternating sequence of a training phase and a testing phase (Fig. 1). In each training phase, dogs were trained to detect a new set of 10 odors over the course of three months. Following each training phase, a baseline test was conducted for the set of odors most recently trained. Additionally, as the study progressed and more odors were trained, recall tests of odors trained in earlier training phases were conducted during the test phases. This sequence was iterated until a total of 40 odors were trained and tested.

Fig. 1 shows the sequence of training and testing events and the corresponding factors of interest (recall period and cumulative odors trained). Dogs were first trained to identify 10 odors (Set A) in the ODT setting, followed by a baseline test for Set A and a test of generalization to six untrained odors related to Set A odors. Next, 10 new (Set B) odors were trained, followed by a baseline test for Set B, as well as a recall test of four of the Set A odors (the remaining reserved for later recall tests of longer durations) (Fig. 1). Thus, this test assessed the ability to (1) identify newly trained odors after having learned 20 cumulative odors, and (2) recall odors from Set A after a period of approximately four

Table 1
Sets of target odorants of each segment of the study. Generalization test odors listed along with related training odor in parentheses.

ID	Set A	Set B	Set C	Set D	Generalization test odors
0 *	Smokeless powder (SP)	–	–	–	Pyrodex® (A0, A4)
1	Ammonium nitrate (AN)	Tea bags	Tea tree oil	Antibiotic cream	Tannerite® (A1)
2	Trinitrotoluene (TNT)	Nitrocellulose	Poppy seeds	Concrete ready-mix	Cast TNT (A2)
3	C4 (untagged)	Vinegar	Butter flavoring	Antacid tablets	Flex-X (A3, A5)
4	Safety fuse	Baking powder	Ranch flavoring	Beef flavoring cubes	AN & sugar (A1)
5	Detonating cord (PETN)	Grape drink mix	Butterscotch candy	Sesame oil	PW4 (tagged) (A3, A10)
6	PE4 (untagged)	Cloves	Band Aids® bandage	Crayon® shavings	
7	Methyl benzoate	Liquid smoke flavor	Potassium Chlorate	Sodium Chlorate	
8	Hydrogen peroxide	Honey	Aspirin	Fish oil capsules	
9	Hexamine (fuel tabs)	Bicycle inner tube	Semtex H	Parmesan cheese	
10^{+}	Vanillin/DMNB/IPE	Dried mushrooms	Nitromethane	Urea nitrate	

*Odor A0 was used in previous detection dog training with all dogs prior to this study, and subsequently used in initial pre-training for the current study.
$^{+}$ Each dog was assigned one of three odors as their A10 target (three dogs each). This was not of direct relevance to this study, but was done to match the use of these odors in a separate study (Lazarowski et al., 2021). DMNB refers to the taggant compound dimethyl di-nitro benzene and IPE refers to a proprietary Dstl inert plastic explosive training aid material formula.

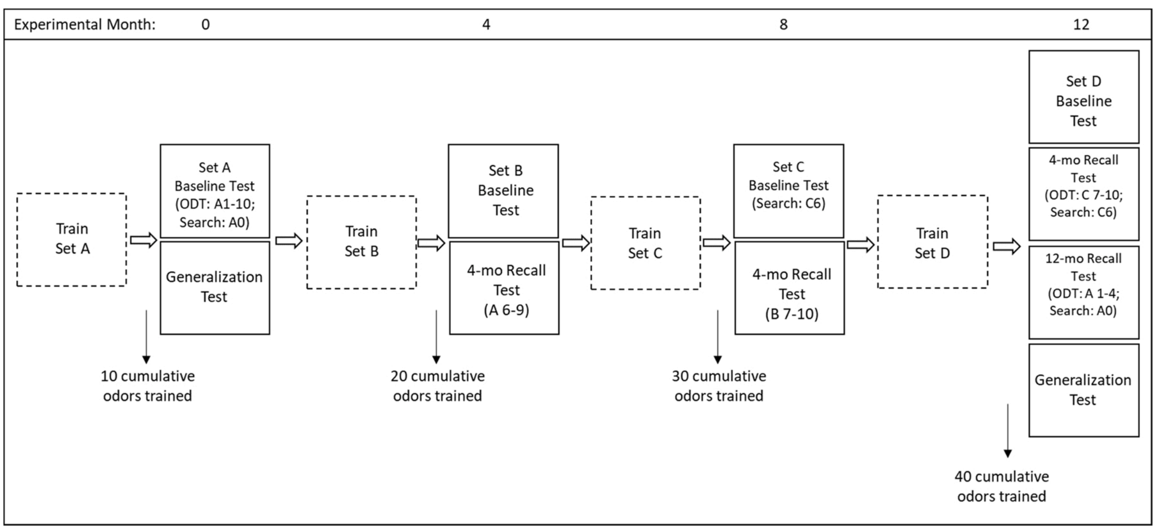

Fig. 1. Schematic diagram of training and testing events over the course of the study.

months without experiencing those odors. This pattern of training a new set followed by a baseline test of the newly trained odors and a recall test of previously trained odors was repeated every four months thereafter for Set C and then D (Fig. 1). A subset of the original set was tested at each recall test, reserving the remaining odors for later recall testing of longer periods. Rate of learning new odors, accuracy in identifying newly trained odors, and accuracy in recalling previously trained odors were measured at each phase. Finally, the generalization test was repeated after 12 months had passed and an additional 30 odor identifications had been learned.

Additionally, in order to examine effects on operational search performance, two odors (A0 and C6) were used in operational search training and testing. Odor A0 was tested at baseline after having learned A1–10 odors in ODT, and was tested again 12 months later after having learned 40 odors in ODT. Odor C6 was trained and then tested for a baseline measure in Month 8 and then tested again in Month 12. This provided a comparison, albeit limited, of operational search performance after having learned 10 and 30 odors and the passage of 4 and 12 months. The aim of this brief additional test was to observe any differences in ODT and search scenario outcomes.

2.5. Training

As in Lazarowski et al. (2021), dogs were trained to enter the arena, sample each position in order from first to last, and alert (by sitting) at the position containing the target until cued to exit the arena, 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. Reinforcement was delivered in the form of a preferred toy and play with the handler outside of the arena.

Dogs were initially trained to perform the ODT using the odor A0 (SP) (described in detail in Lazarowski et al., 2021). A schedule of 70% intermittent reinforcement was introduced once dogs were sampling all positions reliably during initial ODT training with SP. A predetermined randomized list of trials to be reinforced with delivery of toy was followed and the dogs received verbal praise for correct alerts for which the toy was not delivered. To advance to formal ODT testing, dogs were required to complete a set of 12 consecutive trials (containing 10 target trials and two non-target/blank trials) with at least 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. The 10 odors within each set (i.e., Sets A, B, C, D) were trained to criteria sequentially and the dogs had no intervening exposure to the previously trained odors prior to testing.

ODT training for each set of 10 (A,B,C,D) odors followed structured step-wise protocol whereby the number and location of potential target positions and the number of distractors increased across continuous steps in the protocol (Lazarowski et al., 2021). Each step was repeated until predetermined criteria were met before moving to the next step, allowing learning rate data to be recorded as number of trials required above the minimum mandated by the protocol for 37 substances (A4-D40); the initial three substances (A1-A3) were trained using a longer protocol that facilitated training the dogs to perform the task and this data was therefore excluded from learning rate analysis.

For odors A4-D40, a step wise protocol was followed; on the first step the target was placed in one of the first three sampling positions and the handler prompted the dog to sit when it encountered the odor. On the subsequent three steps, the target was placed, respectively, in positions 3 or 4, 5 or 6, and 7 or 8 with distractor odors in several of the intervening positions. To progress across each of these three steps, the dog had to perform three consecutive alerts to targets with no false alarms (FA). The final training step consisted minimally of 10 trials; eight with a target randomly placed amongst the eight positions and two with no target. The order of target and no-target trials was randomized and positions without a target contained either a distractor (at least three positions) or were blank. The criteria for completing this last step was the dog alerting to at least eight out of 10 targets with no more than one FA across 12 consecutive trials. The one handler prompted trial was not considered in the calculation of minimum number of trials to meet the learning criteria since it did not vary between dogs. Thus, the minimum number of trials to meet the learning criteria were [Step 2] 3; [Step 3] 3; [Step 4] 3 and; [Step 5] 10 resulting in a minimum of 19 trials for a dog to meet the learning criteria. If a dog did not meet the criteria in the minimum number of trials (19), the session continued with additional trials until the criteria was met. The number of unassisted (e.g., responses were made independently with no prompting by handler) trials above the required minimum was calculated and used for learning rate analyses.

Except for initial prompting to sample and sit at the target odor when first introduced to the procedure, dogs performed the ODT off-leash with the handler and experimenter remaining behind the partition outside of the arena. All correct target indications were reinforced during initial ODT training with 9 of the 10 Set A odors. Following completion of training for Set A, dogs were re-acclimated to the 70% intermittent reinforcement scheduled with the last odor trained for Set A. This intermittent reinforcement schedule was used during ODT testing and maintenance training (described below).

Because dogs had prior operational search training with A0 before this study, they were given minimal refresher training in the search procedure. Dogs were considered proficient in the operational search procedure once they had exhibited five consecutive correct indications to the target odor with no FA. All operational search training and testing was conducted with the dog off-leash.

2.6. *Testing*

2.6.1. *Recall testing*

Each time the learning criteria was met for all 10 odors in a set, each target odor was tested in the ODT in a separate session for each odor, with the order of odors tested randomized. ODT procedures were identical to that in Lazarowski et al. (2021). To summarize, each test of an odor set consisted of 10 target and two blank trials in random order. The target position was randomized on every trial with at least three different distractors, one of which was a “matched blank”, and empty cans in the remaining positions. On every trial at least one distractor was replaced with a distractor not yet used in that session. Additionally, novel distractors were introduced at each of the A, B, C, and D odor recall test points. Correct responses by the dog were reinforced according to the pre-determined schedule for intermittent reinforcement, and no reinforcement was given for incorrect responses.

All ODT testing was conducted with the dog not in the view of any personnel that were aware of the presence and location of targets. Experimenter 1 placed the targets and distractors according to a randomized pre-determined list and placed a card face down on a table behind the barrier that indicated whether a target was present and its location. Experimenter 2 was seated facing away several feet from the barrier holding the dog. Experimenter 1 exited the arena and traded location with experimenter 2. Experimenter 2 prompted the dog to enter the arena and observed the trial. Upon the dog alerting to a position or completing sampling of all positions, experimenter 2 turned over the card, signaled the dog to exit the arena and stated whether to deliver the toy and/or praise the dog, which was executed by experimenter 1 the position of which the dog was conditioned to go to from the arena. The dog would then remain in the control of experimenter 1 and experimenter 2 recorded the outcome of the trial and tended to the placement of odors and card for the subsequent trial. The two experimenters continued to revolve in this manner across the trials.

For operational search testing, odor A0 (which all dogs had experience in detecting prior to this study) was first tested after the initial Set A ODT baseline test. Search performance for odor C6 was tested after the dogs had been trained and tested on the Set C odors, at which point they had learned a total of 30 odors. Search performance for both A0 and C6 were subsequently tested at Month 12 after the dogs had learned a total of 40 odors. Operational search testing was identical to Lazarowski et al. (2021). At each test point, nine searches for A0 (after Set A ODT), N6 (after Set C ODT), and both A0 and N6 (after set D ODT at month 12) were conducted with each dog. Nine hides each of A0 and N6 were placed in 3 search areas so that no more than three dogs searched the same area so as to inhibit dog’s tracking of other dogs to the location of targets and alerting based on where other dogs had sat. One quarter of searches were conducted single-blind; remaining searches were not able to be conducted blind as the nine dogs rotated being run between the two handlers.

2.6.2. *Generalization testing*

Generalization testing examined dogs’ responses to six different explosives for which they had no prior experience but had chemical, and presumed perceptual, similarity to odors in Set A (Table 1). Generalization testing was conducted using the ODT procedure as described by Lazarowski et al. (2021), immediately following the Set A baseline test and again following the 12-mo post-baseline test. Briefly, each probe odor was tested in a five-trial block consisting of four trials with a previously trained target, and one trial with the probe target. A different novel distractor that the dogs had not experienced in training or testing was presented on each of the non-probe target trials, placed in a position ahead of the target such that the dog encountered a total of five novel distractors in each probe test. This was done to evaluate true generalization to perceptually similar odors versus responses to novel odors. Indications on probe trials were not reinforced, and one out of every four (randomly selected) correct indications of the trained targets was not reinforced in each five-trial block.

2.7. *Maintenance training*

In between testing and training cycles, dogs received minimal maintenance training with odor A10, which was not used in any recall testing. The schedule of such training was one ODT session with three target trials (2/3 reinforced) and one blank trial, and one operational search session with three targets (3/3 reinforced) for each dog. Only one ODT or search session occurred every other week with no training in the intervening week. Thus, even if the dogs expeditiously met their ODT training criteria in one week, they received at most three maintenance training sessions, given that testing took at least one week per four month cycle.

2.8. *Scoring and data analysis*

Learning rate was measured as numbers of trials over the predetermined minimum (19) to meet the criteria of having learned each odor (excluding odors 1–3). A repeated measures analysis of variance (ANOVA) was conducted to determine if the mean number of trials to learn each set of odors differed as a function of cumulative odors learned. A Greenhouse-Geisser correction was applied due to violation of the assumption of sphericity of equal variance assessed by Mauchly's sphericity test.

For testing, the primary behavioral measures were indication responses to targets (hits) and non-targets (FA). Change of behavior (COB), defined as a sudden change of ongoing search behavior characteristic of target odor detection (Minhinnick, 2016), was also recorded and used in analyses of search performance and in the generalization tests. Only two instances of COB were recorded in all of the ODT recall testing, therefore the COB measure was only included for search and generalization testing. Separate response rates were calculated using hits only (total hits out of total targets presented) and combining hits and COB (total hits and COB out of total targets presented). Reinforcement was never delivered for COB alone – only trained indication responses were reinforced. FA rate was calculated as the total number of responses to non-targets divided by the total number of trials for both ODT and search. Additionally, we calculated positive predictive value (PPV) as a measure of precision in the search and generalization tests, calculated as: PPV = hits/ hits + false positives, where a PPV closer to 100% indicates a higher probability that a response is correct and not a FA (Simon et al., 2018). Paired samples *t*-tests were performed to compare response rates between the same odors tested at different times. Hit, COB, and FA rates were also compared between blind and non-blind search trials. Statistical analyses were conducted using SAS® 9.4 for Windows®.

3. Results

3.1. Learning rate

Dogs required more trials and exhibited greater variability in learning the first odor in the first set trained (A1) (M = 60.33, SE = 18.95), with one dog a clear outlier requiring over 200 trials over the required minimum to meet the learning criteria for the first odor. The average number of trials to meet the learning criteria rapidly decreased across the remaining Set A odors, from 13 (SE = 2.45) for the second odor to 5.33 (SE = 0.84) for the last odor.

After learning to identify the first 10 odors (Set A), dogs exhibited a generally very high (i.e., low numbers of trials) and consistent rate of learning the subsequent 30 odors (Fig. 2). On average, dogs required 5.4 trials over the required minimum (range: 3.8 – 13.4) across odors in sets B-D. The average number of trials to learn new odors did not significantly differ across sets A-D, $F(3, 24) = 4.88$, $p = .053$ (odors A1–3 were excluded from analysis due to planned procedural difference in training the first three odors designed to facilitate the dogs learning the

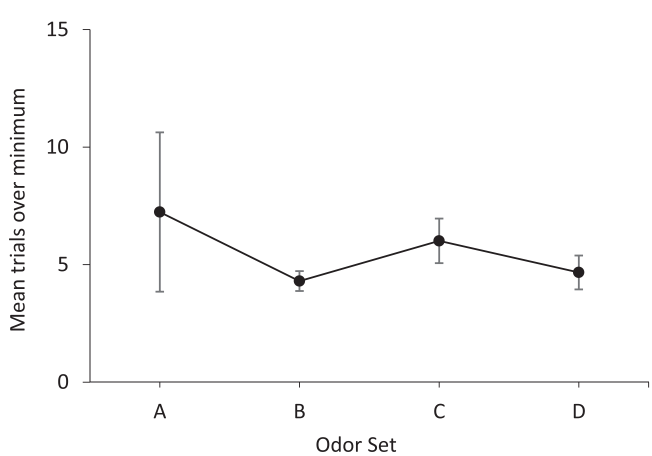

Fig. 2. Average (+ SE) number of trials across 9 dogs over the required minimum (19) before reaching the learning criteria for each set of odors, excluding odors A1–3 from the average for Set A.

procedure).

3.2. Recall

Four-month recall was assessed across three different tests in the ODT, with either 20, 30, or 40 cumulative odors learned. Fig. 3 shows a comparison of the baseline hit rate to the 4-month recall hit rate for the A, B, and C subsets tested. For the A odors, hit rate increased significantly from 97.2% (SE = .88) at baseline, at which point 10 odors had been learned, to 100% (SE = 0) at the 4-month recall test, at which point 20 odors had been learned, $t(8) = 3.16$, $p = 0.013$ Hit rate did not significantly differ between the baseline and 4-month recall test for the B odors (tested after 20 and 30 cumulative odors trained), $t(8) = 2.0$, $p = .0805$, or the C odors (tested after 30 and 40 cumulative odors trained), $t(8) = -.43$, $p = 0.6811$ FA rates were negligible in each 4-month recall test (A baseline: M = 3.4%, SE = 1.09; A 4-month: M = 0.46%, SE = .30; B baseline = 0.46%, SE = 0.46 B 4-month: 0.23%, SE = 0.23 C baseline = 0.93%, SE = 0.70 C 4-month = 1.38%, SE = 0.88.

One odor (C6) was reserved for 4-month recall testing in the operational search setting. Average hit rate did not change from the first (M = 65.43%, SE = 6.26) to the second test (M = 65.43%, SE = 8.77), $t(8) = 1.18$, $p = .27$ (Fig. 4). Average FA rate was 2.46% (SE = 2.33) at the first test and 3.7% (SE = 1.85) at the 4-month test. Average PPV was 96.29% (SE = 2.45) at the baseline test and 95.17% (SE = 2.43) after four months. There was no difference between single-blind and non-blind searches for hit (excluding and including COB, $p = .731$) or FA rate ($p = .1872$).

For assessing 12-mo recall, a subset of four Set A odors was reserved after the baseline test and not tested again until 12 months later in the ODT. Hit rate significantly increased from 97.5% (SE =.93) at baseline (after learning 10 odors) to 100% (SE = 0) at the 12-mo recall test (after learning 40 odors), $t(8) = 2.68$, $p = 0.0278$ FA rates were negligible in the baseline (0.23%) and the 12-month test (0.91%). A0 was reserved for testing 12-mo recall in operational search. Hit rate for this odor decreased significantly from the baseline test (M = 93.8%, SE = 2.69) to the 12-month test (M = 58.02%, SE = 9.76), $t(8) = -4.35$, $p = .0025$ (Fig. 5). However, when COB was included in the response rate, performance did not significantly differ between baseline (M = 98.7%, SE = 1.23) and 12-mo recall (M = 95%, 2.69), $t(8) = -1.15$, $p = .28$. Average PPV was 100% (SE = 0) at the baseline test and 97.77% (SE = 2.22) at the 12-mo test. There was no difference between single-blind and non-blind searches for hit (excluding and including COB, p = .347 and p = .368, respectively) or FA rate (p = .521).

3.3. Generalization

At baseline, all dogs responded to all six of the probe odors except one dog that responded to five of the six, exhibiting a COB to the one odor to which he did not make an indication response. There were no FA to distractors or novel distractors (100% PPV). Average response rate to probe odors decreased from 98.15% (SE = 1.85) at baseline to 88.89% (SE = 5.55) after 12 months; however, this difference was not significant, $t(8) = -1.75$, $p = .14$. At the individual level, only three of the nine dogs showed a decrease in responding; one of these dogs showed a COB to one of the probes in the 12-mo test. Average PPV was 86.57% in the 12-mo generalization test; there were a total of four FA committed by the same two dogs, all to novel distractors.

4. Discussion

The goal of this study was to assess whether detection dogs' ability to learn and retain new odors is impaired by increasing the cumulative number of odors trained, and to examine the duration of long-term memory for previously learned odors when a large number of additional odors have been trained in the interim. Across 12 months, the rate of learning new sets of 10 odors trained every three months and

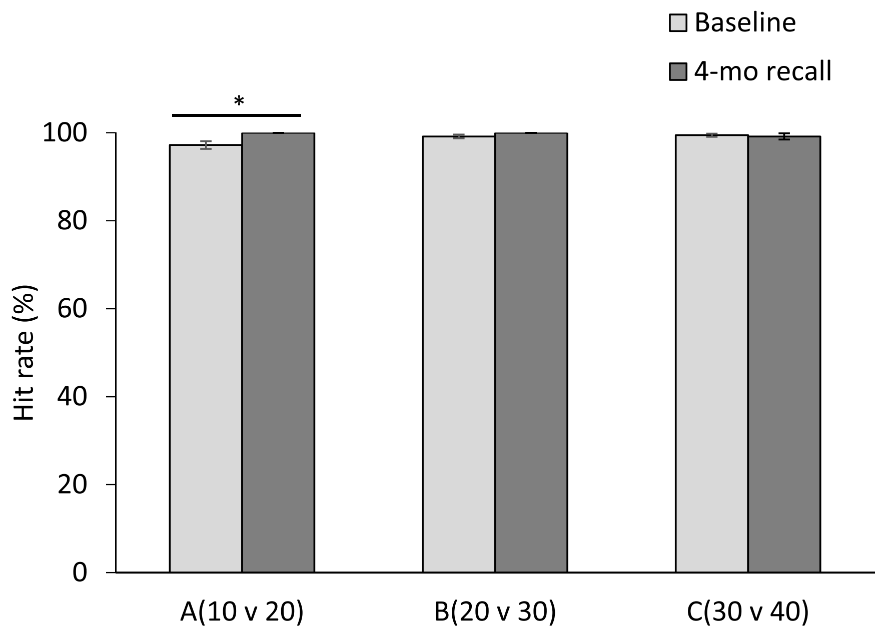

Fig. 3. Mean (+SE) baseline and 4-month recall for subsets of odors from sets A, B, and C. Number of cumulative odors trained at each test shown in parentheses. Asterisk indicates a significant difference (p <[0.05. There were 10 odor trials for each of 4 odors from each Set A, B, and C for each of 9 dogs at both baseline and 4-month tests.

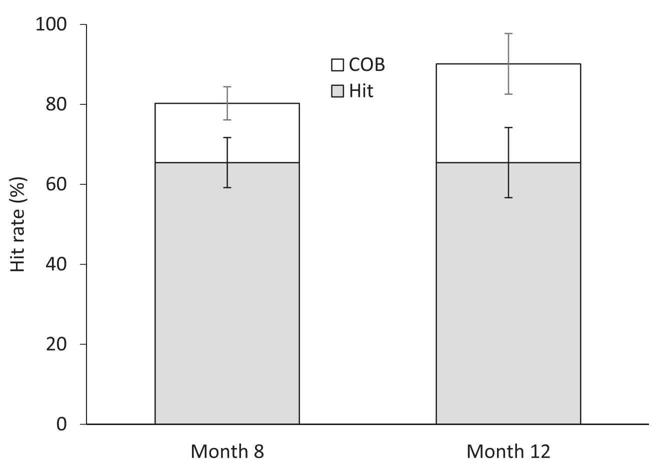

Fig. 4. Mean (+SE) hit rate for odor C6 at baseline and after four months in the search setting.

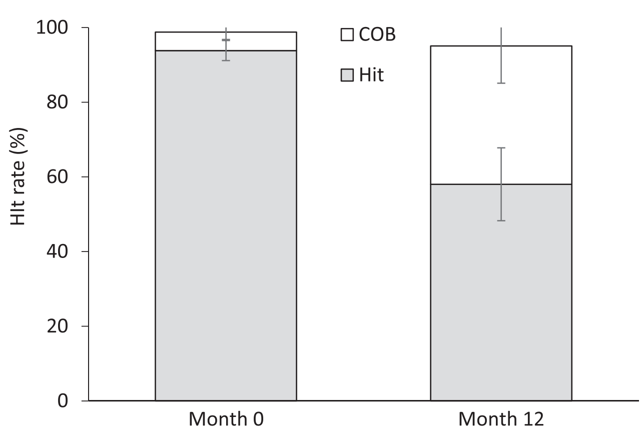

Fig. 5. Mean (+SE) hit rate across nine searches for nine dogs for the odor tested 12 months apart in the operational search setting.

accuracy in recalling previously learned odors was assessed. Additionally, we assessed whether extensive odor discrimination training and extended passages of time influenced dogs' tendency to generalize to untrained variants of the trained odors.

After an initial trend of decreasing trials needed to learn the first odors, dogs exhibited a rapid rate of learning new odors with minimal variability for up to the maximum of 40 odors trained across the 12 months. The results for learning rate across the 40 odors suggest that, once dogs are competent in performing the procedure for learning odor discriminations, they are capable of readily acquiring a large number of odor discriminations. These results are consistent with those of Williams and Johnston (2002), where dogs required fewer trials to meet the criterion as new odors were trained, for up to the maximum of 10 odors trained. Our results extend these findings to a much larger training set, and showed no trend toward increasing variability between dogs or decreases in learning rate for up to 40 odors, suggesting that 40 odors was not approaching any limit on the odor learning capacity of dogs.

Effects of learning an increasing number of odor discriminations as well as the passage of time was assessed by comparing response rates on baseline tests for subsets of odors to response rates for the same odors after periods of four and twelve months, after having learned an additional 20, 30, or 40 odors, respectively. Across the four tests of 4-month recall, dogs successfully recalled odors with very high accuracy. There was a small but significant improvement in 4-month recall for the first set trained, which may have been an artifact of the baseline accuracy for the A odors being relatively low as this was the first set trained. This result is consistent with Porritt et al. (2015), in which dog detection accuracy improved from baseline measures following initial training across six weeks as their execution of a repetitive search task, which included reinforced responses to targets, improved. Similarly, learning rates were lower and more variable across dogs for the A odors, suggesting that the dogs' performance was not yet stable. Thus, performance may have increased as dogs gained more experience with the task. Taken together, performance in both the ODT and search tests suggests that recall in dogs regularly engaged in an odor detection task is not impaired after a period of four months having learned between 10 and 40 intervening odors.

Dogs recalled the four odors not seen for 12 months, and with intervening training of an additional 30 odors (40 cumulative), with

100% accuracy. This was slightly but significantly higher than their accuracy when tested directly after these odors were trained (97.5%), at which point the dogs had been trained on only 10 odors. Similar to the increase seen for the Set A 4-month recall, this increase may be due to accuracies for the A odors in the initial test being lower than for the subsequent odor sets, which were trained and tested after the dogs had more experience with the ODT procedure. These increases in recall accuracy, albeit small (2.8% and 2.5%), are important given that they indicate there was no decrease in accuracy after four and 12 months with no exposure to the target odors during these periods. In our recent examination of long-term odor memory in a different sample of dogs, we found that recall of a set of 10 odors in the ODT decreased significantly over a period of 12 months of no odor detection maintenance training, but remained high for dogs that received intermittent maintenance training with a single surrogate odor throughout the 12 months (Lazarowski et al., 2021). The current results extend these findings to suggest that in addition to single-odor maintenance training, intervening training with a large variety of odors can also effectively maintain odor memory.

In contrast to the 12-month ODT test, accuracy decreased substantially over 12 months in the operational search test. However, there was no reduction in performance when COB, indicative of target odor recognition, was included in the response rate, suggesting that dogs recognized the odors but failed to perform the final trained response. That is, the previously conditioned association to the odor appears to have been retained, but the function of the odor as a discriminative cue for the operant alert response in the search context had partially decayed over time. The lack of difference between performance, including COB, on blind versus non-blind trials suggests that the enhanced accuracy when considering COB was not due to the rather subjective nature of this metric. The difference in search performance as compared to ODT performance may be expected given the relatively infrequent search training compared to the intensity of ODT experience across the time course of study. These findings are similar to Lazarowski et al. (2021) in which infrequent maintenance training was effective in sustaining recall in the ODT, but not performance in the operational search test. Taken together, the results suggest that the memory for the odor persists up to 12 months in the absence of exposure to that particular odor, but that the mechanics of performing the trained response to alert to the odor in an operational search scenario deteriorated in the absence of maintenance training with the odor.

Interpretation of the results from operational searches should be tempered by consideration of the limited amount of search data and experience of dogs in performing searches. The purpose of including searches was to take advantage of the experimental context to examine the effects of learning many odor discriminations and elapsed time on recall of odors in an operational context to parallel that of the effects on odor discrimination, which was the principal topic of investigation. Due to logistical constraints, the examination of search recall was limited to just two target odors and only partial single-blind observations. The decrease in performance for 12-month search recall may have been an artifact of the relatively low frequency of search maintenance training (three searches per dog per month). Nonetheless, when hits and COB are combined, there was no significant decrease observed in search odor recognition after 12 months. Although a somewhat limited probe, these results suggest that dogs that receive regular odor discrimination and search maintenance training, but have no contact with the target odors for up to 12 months after learning them, may not have reduced recognition of them in operational-type searches.

Finally, learning up to 40 unrelated odors and a lapse of 12 months since last exposure did not result in an appreciable decline of generalization to variants of the first set of odors trained for the majority of the dogs. These results are in contrast to Lazarowski et al. (2021) in which generalization decreased substantially over a 12-month period with and without single-odor maintenance training. Results from the current study suggest that extensive training on a large number of odors did not appear to narrow the generalization gradient, as is often the case following extensive training with the same stimulus (Moser et al., 2019). The slight decline in generalization at 12 months may be related to the time since exposure to the related odors or, especially given the large amount of odor discrimination training experienced by the dogs, prior unreinforced exposure to the generalization odors may have made it more likely that they reject these odors upon subsequent presentation. However, these results should be considered in the context of this being a very limited (i.e., small number of observations) probe of the effects of learning many odor discriminations on generalization, warranting further examination.

5. Conclusions

Our results demonstrate that learning an increasing number of target odors (up to 40) did not result in proactive interference, which occurs when previous learning disrupts the ability to form new memory associations, nor did it result in retroactive interference, which occurs when learning new stimuli reduces the memory for those learned previously (Koster et al., 2002). Overall false response rates across the study were negligible, indicating that learning to respond to numerous odors did not increase dogs' tendency to respond to odors in general (in other words, did not reduce their specificity).

We also show that up to 12 months since last contact with particular odors did not reduce dogs' recall of those odors in a discrete trials setting, nor their recognition in operational searches despite some degradation in the alert response. Although dogs had no contact with the odors between tests, they were nearly continuously engaged in odor detection activity throughout the 12 months of the study because additional odors were trained every three months. Further investigation of the frequency and form of ongoing training necessary for efficient and effective maintenance of odor detection performance is needed.

The results indicate that the limits of odor identification recall capabilities of dogs were not challenged by having learned up to a total of 40 odors or by up to a 12-month gap in experience with those odors, demonstrating the robust capabilities of dogs for learning and remembering many odor identifications. Combined with the limited examination of operational search performance in this study, it appears that detection dogs have a significant capacity for detecting many different target odors and remembering those odors over several months without explicit training on those odors. Further research and technology development is warranted to optimize the design of training and maintenance systems to fully utilize the odor memory capabilities of working dogs. Given its importance to operational detection tasks, additional attention to how maintenance training affects generalization of odor learning should be considered in such future research and development efforts. At a more fundamental level, the results of this study provide further evidence of a special odor memory facility of animals for which odor detection is essential to their navigation of the environment, which is deserving of further scientific inquiry.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was supported by the U.K. Ministry of Defense contract # UKDPO/00009 and the U.S. Navy, Office of Naval Research grant # G00010016. The funding bodies were not involved in experimental design, data analysis, interpretation or article preparation. We thank Auburn University Canine Performance Sciences (CPS) canine instructors Terry Fischer, Jenn Jankiewicz, and Laura Reeder for executing canine activities, and Holli Thompson for managing the care and welfare of the dogs. We also thank Philippe Gaillard and Rachel Stuckwisch of the Statistical Consulting Service of the Auburn University, Department of Mathematics and Statistics for assisting with data analysis.

References

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.

Hall, N.J., Collada, A., Smith, D.W., Wynne, C.D.L., 2016. Performance of domestic dogs on an olfactory discrimination of a homologous series of alcohols. Appl. Anim. Behav. Sci. 178, 1–6. https://doi.org/10.1016/j.applanim.2016.03.016.

Hall, N.J., Wynne, C.D. L., 2016. Canine olfactory learning and behavior. In: Jezierski, T., Ensminger, J., Papet, L.E. (Eds.), Canine Olfaction. Science and Law, pp. 123–137.

Hayes, J.E., McGreevy, P.D., Forbes, S.L., Laing, G., Stuetz, R.M., 2018. Critical review of dog detection and the influences of physiology, training, and analytical methodologies. Talanta 185 (April), 499–512. https://doi.org/10.1016/j.talanta.2018.04.010.

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

Koster, E.P., Degel, J., Piper, D., 2002. Proactive and retroactive interference in implicit odor memory. Chem. Senses 27 (3), 191–206. https://doi.org/10.1093/chemse/27.3.191.

Lazarowski, L., Waggoner, P., Hutchings, B., Angle, C., Porritt, F., 2021. Maintaining long-term odor memory and detection performance in dogs. Appl. Anim. Behav. Sci. 238 (March), 105301 https://doi.org/10.1016/j.applanim.2021.105301.

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.

Minhinnick, S., 2016. targets, alerts, distractors, reinforcement, and other issues. Statistical reliability confounders and improvement in advanced dog training: patterns, routines. Science and Law, pp. 123–137.

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

Oxley, J., Waggoner, L.P., 2009. Detection of Explosives by Dogs. In: Aspects of Explosives Detection. Elsevier, Amsterdam, pp. 27–40.

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.

Simon, A.G., Degreeff, L.E., Peranich, K., Furton, K.G., 2018. Canine generalization to molecularly similar odors and odor mixtures. (No. NRL/MR/6181-18-9797). Naval Research Lab, Washington, DC.

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. Process. 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.

Williams, M., Johnston, J.M., 2002. Training and maintaining the performance of dogs (Canis familiaris) on an increasing number of odor discriminations in a controlled setting. Appl. Anim. Behav. Sci. 78 (1), 55–65. https://doi.org/10.1016/S0168-1591(02)00081-3.

Wright, G.A., Smith, B.H., 2004. Variation in complex olfactory stimuli and its influence on odour recognition. Proc. R. Soc. B: Biol. Sci. 271 (1535), 147–152. https://doi.org/10.1098/rspb.2003.2590.