Can Concentration-varied Secondary Target Training Improve Generalisation Across Primary Target Concentrations in Scent-detection Dogs?

A thesis

Submitted in partial fulfilment

of the requirements for the degree

of

Master of Applied Psychology in Behaviour Analysis

at

The University of Waikato

by

Ishan Bhasin

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

March 13, 2026

Abstract

Scent-detection dogs are often expected to identify target odours across a range of concentrations, despite typically being trained using a single baseline concentration. This raises the possibility that concentration changes may disrupt stimulus control and reduce accurate responding in applied settings. Previous research suggests that dogs may show limited generalisation when target quantity or concentration differs substantially from training, indicating that concentration may be a functionally important dimension of olfactory stimulus control. The present study examined whether dogs trained to detect target odours at a single concentration would spontaneously generalise responding to higher concentrations of those same odours. Three dogs were trained using an automated olfactometer to discriminate two target odours, cinnamaldehyde and hexanoic acid, from non-target odours. After discrimination training, novel non-target testing, and intermittent reinforcement training, two dogs completed non-reinforced probe trials involving higher-concentration variants of the trained target odours. Probe responding differed systematically across odours. For both dogs, indications to higher-concentration hexanoic acid probes were more frequent than indications to non-targets, but less frequent than indications to the trained target, consistent with partial generalisation. In contrast, indications to the higher-concentration cinnamaldehyde probe overlapped with indications to non-targets, suggesting little reliable transfer of stimulus control. These findings suggest that training at a single concentration may not be sufficient to support robust generalisation across concentration changes and highlight the importance of treating concentration as a relevant training dimension in scent-detection training.

Acknowledgments

I would first sincerely thank my supervisor, Associate Professor Tim Edwards, for his immense guidance, feedback, and genuine support throughout the process of this thesis.

His advice in all parts of the project, from the technical side of the experiments to the writing process, made this work much more manageable than it would have been. I am also very grateful to Rob Bakker for his assistance in the dog lab and for his support with many technical aspects of the experimental work.

Both Tim and Rob have helped me a lot and have been incredibly supportive throughout this research. I would also like to thank the dog owners, Rebecca, Jill, and Kate, for their time, commitment, and willingness to bring their dogs in for this project. They were very kind and supportive throughout the whole process. I also want to acknowledge the three dogs, Atlas, Shiner, and Ellie. Working with them was the most rewarding part of this thesis for me. This thesis would not have been possible without them. Finally, I would like to thank my mother, Shilpa Bhasin, and my brother, Nikhil Bhasin, for supporting me in countless ways and making it possible for me to complete this thesis.

Table of Contents

List of Figures

List of Tables

Introduction

Scent-detection dogs (SDDs) are an effective means for detecting contraband, including explosives, for security and other relevant purposes, due to their demonstrated detection capabilities (Furton & Myers, 2001). Detection training for scent detection dogs is established through operant conditioning, where responses to target odours are reinforced with preferred items, while responses to non-target odours are typically not reinforced and placed on extinction (Edwards et al., 2022). When discrimination training is appropriately designed and maintained, scent-detection dogs may demonstrate reliable alerting performance to trained target odours in operational contexts (e.g., Porritt et al., 2015).

Although detection dogs reliably alert to trained target odours under familiar training conditions, performance may deteriorate in operational contexts when encountered targets fall outside the range of training experience, particularly when they differ substantially in quantity or concentration (Aviles-Rosa et al., 2021). During training, detection dogs are typically exposed to a limited set of target odours presented under controlled conditions, with consistent contexts and relatively high target availability. In contrast, operational searches may involve unfamiliar environments, low target density, novel concealment methods, and substantial variation in target properties such as composition and packaging (Lazarowski et al., 2020). Since real-world scenarios rarely match training conditions exactly, reliable detection requires that the trained response be maintained across variations in odour concentration (DeChant & Hall, 2021b), location (Gazit et al., 2004), and surrounding stimuli (Lazarowski et al., 2020).

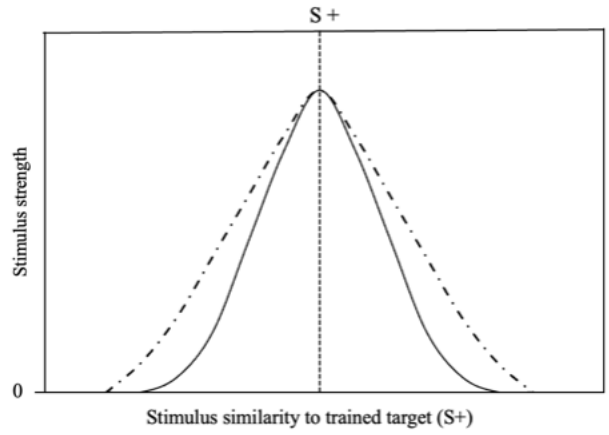
Dogs' effectiveness as detectors depends on their olfactory sensitivity and the degree to which stimulus control is appropriately balanced between discrimination and generalisation across trained odours (Moser et al., 2019). Olfactory generalisation refers to the extent to which

responding occurs to odours that share features with a trained target, whereas olfactory discrimination refers to the degree to which responding is restricted to the trained odour relative to similar non-targets; in practice, both processes vary along a generalisation gradient: increased discrimination is shown by a narrow curve, whereas increased generalisation is indicated by a broad curve (Caldicott et al., 2024; Moser et al., 2019; see Figure 1). It is essential that SDDs maintain a task-appropriate balance between generalisation and discrimination. Stokes and Baer (1977) proposed that generalisation often needs to be explicitly programmed through the arrangement of training conditions, rather than assumed to emerge automatically following acquisition.

Variability in generalisation performance in detection dogs has been documented across multiple target odour stimulus features, including chemical composition, mixture structure, and training history, dimensions that can differ substantially between training and operational environments (Moser et al., 2019). Among these features, variation in odour concentration represents a particularly consequential challenge for operational detection, as field targets may be encountered at quantities far outside the range experienced during training.

Figure 1

Schematic Stimulus Generalisation Curves Illustrating Narrow (High Discrimination) and Broad (Greater Generalisation) Response Profiles Relative to a Trained Target (S+)



Empirical evidence from operational contexts illustrates the consequences of insufficiently programmed generalisation along concentration differences. A case study by Aviles-Rosa et al. (2021) highlights a crucial issue: SDDs often fail to generalise their response to the target component when it is of a larger quantity. The study involved experiments simulating a real-life incident where two certified SDDs did not detect a bag containing 13 kg of a putative ammonium nitrate fuel oil (ANFO) mixture. They hypothesised the following reasons as the plausible causes: the context was different; the ANFO substance was different from the trained ANFO samples, hence a lack of generalisation; the dogs did not generalise to the quantity which was significantly larger than the trained amount. The study concluded after extensive research that context might not be the reason for the failure as the dogs detected the ANFO confiscated explosives when it was of the same quantity. However, four out of five dogs failed to generalise to the 13 kg ANFO explosives in the laboratory setting as was seen in the operational example. They highlighted the need to promote generalisation in training to reduce such

difficulties in work-based environments. Realistically, however, training with large quantities of explosives to promote generalisation is typically not feasible due to the high cost, significant safety risks, and strict storage regulations associated with such materials.

Collectively, failures of generalisation across odour concentration are not limited to increases in intensity, but are also observed when odour concentrations decrease, indicating that these effects cannot be explained as simple strength-related performance changes but instead reflect alterations in stimulus control. Concentration-related limitations in generalisation are not confined to cases where target quantity increases beyond training levels. DeChant et al. (2021a) examined whether dogs trained at a single baseline concentration of isoamyl acetate would generalize responding to lower concentrations of the same odour. Dogs responded reliably across concentrations proximal to the training stimulus; however, at a ten-fold reduction in vapour concentration relative to training, generated through controlled air dilution and validated by PID measurements, responding declined to clean-air levels. Additionally, this decline was unlikely to reflect a sensory detection limit, as estimated vapour concentrations remained well above previously reported canine detection thresholds for isoamyl acetate. This suggests that the observed decline in responding more likely reflected a failure of stimulus generalisation rather than an inability to detect the odour.

Similar failures of generalisation across concentration-dependent changes in stimulus control are not unique to detection dogs but are evident across olfactory systems in multiple species, wherein the odour ceased to function as the same stimulus. In humans, variations in odour concentration of 100-fold or more can be interpreted as having a novel odour quality that is as different as a new chemical (Gross-Isseroff & Lancet, 1988). Further, research shows that rodents generalise from a trained odour to concentrations up to ten-fold higher or lower than the

training concentration, whereas responsiveness declines markedly when concentration differences exceed this range (Cleland et al., 2002).

Differences in target concentration between training and operational contexts therefore represent a training–work mismatch comparable to another well-documented operational challenge: low target density. In practice, SDDs are typically exposed to targets far more frequently during training than during real-world deployments. Consequently, operational environments may provide insufficient reinforcement to sustain trained search behaviour, resulting in reduced persistence and performance despite intact discrimination abilities (Porritt et al., 2015). Gazit et al. (2004) documented the impact of performance degradation in a low target density environment on explosive detection canines, demonstrating that the dogs were able to quickly distinguish between two comparable search regions, one of which included target odours and the other did not. In contrast to target-rich and new areas, dogs' search speeds were slower in target-free areas. This decline in performance may be attributed to extinction processes, which occur when reinforcement is no longer provided for a previously reinforced behaviour, resulting in a decrease in that behaviour over time (Cooper et al., 2020).

An obvious solution to this is planting the target contraband in the operational setting to mitigate the performance decline and maintain the search behaviour. However, this may involve placing contrabands where they are prohibited and is logistically not feasible in most cases.

To work around these practical constraints, training-based alternatives that maintain reinforcement without introducing the primary target into operational settings have been proposed. Porritt et al. (2015) conducted a controlled experimental evaluation of a training strategy which involves training SDDs with an innocuous secondary target in training in addition to their primary target. They demonstrated that incorporating a co-trained secondary target into

the work environment following initial training maintained detection performance for the primary, low-frequency targets. Notably, dogs that received secondary target exposure during work searches showed significantly higher detection rates for the primary contraband targets than dogs that received no target exposure in the work environment. This strategy prevented extinction of the trained search behaviour by reinforcing search persistence, even in the absence of the primary target, thereby sustaining the dogs' motivation and vigilance in operational contexts. Thrailkill et al. (2018) reported reinforcing indications of a secondary target- preserved detection responding to a primary target even across sensory modalities in rats. These findings are consistent with the interpretation that training with secondary targets support persistence of detection behaviour beyond stimulus-specific control. Hence, this is a feasible alternative to maintain performance in SDDs without involving the need for planting the primary target in the work context.

Although originally developed to address extinction effects, this method may also influence generalisation. Since, training dogs with very high concentrations of the primary target is often impractical or unsafe, introducing a secondary target across a range of concentrations may serve as an alternative way to incorporate variability along the concentration dimension without directly increasing the concentration of the primary target. This increased exposure to reinforced detection responding under systematically different stimulus concentrations may reduce overly narrow stimulus control and increase tolerance to concentration change, even in other targets. We aimed to determine whether training across a wide range of concentrations for one stimulus transfers to improved detection of another stimulus across varied concentrations, when that second stimulus has only been trained at a single concentration. In other words, we

tested whether widening the generalisation gradient for a secondary target might transfer into a broader gradient for the primary target.

Research on stimulus control shows that training history with one stimulus can influence generalisation to other stimuli, even when those stimuli were never explicitly trained. For example, discrimination training arranged to sharpen control over one auditory cue has been shown to produce a narrower generalisation gradient for a different auditory stimulus that was not trained as a discriminative signal (Jenkins & Harrison, 1960). Work on transfer and generalisation indicates that the amount of variability, the balance across training conditions, and the reinforcement history influence how stimulus control extends to new stimuli that share functional properties with the training set (Howard, 2000). In the present context, different odours may become functionally related because they occasion the same trained detection response and contact the same reinforcement contingencies, such that training-induced changes in stimulus control for one odour may influence responding to another. In contrast to prior work showing that sharpening stimulus control for one stimulus can influence generalisation to another, the present study examines whether training that broadens stimulus control for a secondary target can also broaden generalisation for a primary target along the same stimulus feature, across differences in concentration.

The objective of this study was to examine whether dogs spontaneously generalised to novel, untrained changes in concentration of a trained target odour when probed, and if not, whether subsequent training with a secondary target presented across varied concentrations influences generalisation to the higher concentration of the primary target. The present work represents Phase 1 of the study, where dogs were trained to detect a primary odour and a

secondary odour and then assessed using a probe test with higher concentrations to evaluate spontaneous generalisation.

Method

Participants

Dogs were recruited via posters, social media adverts, and prior recruitments. Recruitment involved a screening procedure, which included completion of a screening form to assess dogs' eligibility and general characteristics. Dogs who met criteria according to the forms were brought to the Scent Detection Research Group facility on the University of Waikato campus for an initial assessment of eligibility. During this session, dogs were observed while crated and during brief separation from their owners to assess tolerance of these conditions. Dogs that showed signs of distress in either context were excluded from participation. Initial eligibility criteria included reliable engagement with food delivered via the automated feeding apparatus, tolerance of a 2–3 min separation from owners without sustained vocalisation or escape-directed behaviour, and tolerance of leash walking and brief manual handling without resistance or aggressive behaviour. Owners received detailed information about the study and were given the opportunity to ask questions before providing written informed consent for their dog's participation.

Dogs completed initial training sessions without their owners present. Dogs were trained to eat from an automated feeder, which dispensed food when activated by the experimenter. Dogs progressed to subsequent phases of the study, if they reliably approached the feeder and consumed the food promptly. All experimental procedures were approved by the University of Waikato Animal Ethics Committee (Protocol No. 1226).

A total of twelve dogs were recruited for the study. Three were excluded during preliminary assessment due to difficulty tolerating separation from their owners, and a further six were withdrawn following limited progression during initial training, which included low

engagement with food reinforcement, frequent requirement for experimenter prompting to complete sessions, and low accuracy across multiple sessions of the initial study phases. Three dogs progressed to the later experimental stages, and two of the three dogs completed probe testing. Table 1 includes demographic information for the dogs included in the study. None of the dogs had prior experience with scent detection training or with the experimental apparatus used in this study.

Table 1

Subject Information

Subjects	Breed	Age (years)	Sex
Atlas	Border Collie x Husky	2	NM
Ellie	Springer Spaniel	6	NF
Shiner	Blue Heeler	3	NF

Note. NM = neutered male; NF = neutered female.

Setting

All experimental sessions were conducted at the Scent Detection Research Group facility at the University of Waikato in Hamilton, New Zealand. Experimental testing took place in an experimental room, while dogs were housed individually in a separate kennel room when not actively participating in sessions. Kennels were equipped with a mat and blanket, and water was available *ad libitum*. Dogs were tested between Tuesday and Friday. On testing days, dogs were dropped off by their owners at approximately 12:30–1:00 pm and collected at the conclusion of sessions between 4:00 and 4:30 pm. Dogs were walked outside the facility by the experimenter approximately every two hours.

Project Stages

The study employed a single-case experimental design involving phased operant discrimination training and embedded probe trials to assess stimulus generalisation. This experiment was conducted in several stages (see Table 2). Following eligibility assessment, dogs were trained on a primary target odour using an operant discrimination procedure. Once stable discrimination performance was achieved, dogs were trained on a second target odour using the same procedure. After both target odours were trained independently, dogs progressed to sessions in which both target odours and multiple non-target odours were presented within the same session.

Subsequent stages assessed dogs' discrimination performance when novel non-target odours were introduced, to confirm that responding was controlled by the trained target odours rather than by familiarity and rejection of specific distractors. Once discrimination performance was stable, reinforcement for target responses was gradually reduced using an intermittent reinforcement schedule.

The final stage of the experiment involved non-reinforced probe trials to assess generalisation of responding across untrained changes in odour concentration.

Table 2

Project Phases

Phase	Task	Target odours	Non-target odours	Probe stimuli
1	Preliminary training	A	X	-
2	Primary target discrimination	A	X	-
3	Secondary target discrimination	B	Y	-
4	Mixed target discrimination	A, B	X, Y	-
5	Novel non-target introduction	A, B	X, Y, L, D, K, S...	-
6	Intermittent reinforcement	A, B	X, Y, L, D, K, S...	-
7	Concentration generalisation probes	A, B	X, Y	A+, B+, B++

Note. Odour labels (A, B, X, Y) correspond to stimulus assignments described in Table 3.

Apparatus

Dogs were trained and tested using an automated olfactometer, a computer-controlled device that delivers controlled odour stimuli to a designated sampling port under continuous airflow. The apparatus was housed within a 12-inch cubic unit, allowing precise regulation of odour presentation and removal. An overview of the automated olfactometer apparatus, including external and internal components, is presented in Figure 2.

The apparatus contained six individual odour vials; each connected to a computer-controlled valve manifold. On each trial, the control software activated the valve corresponding

to the designated odour channel. When opened, clean carrier air flowed through the selected vial at a regulated rate of 0.2 L/min, carrying odour vapour from the vial. The odour-laden air was then directed into a polytetrafluoroethylene (PTFE) manifold, where it merged with a continuous stream of clean carrier air delivered at approximately 4 L/min. This mixing within the PTFE manifold diluted the concentrated odour before it was transported to the sampling port.

The diluted airstream was delivered to a cylindrical nose port located on the front panel of the apparatus. An infrared beam positioned across the entrance of the port recorded nose entries and sampling duration. A lever adjacent to the port registered active rejection responses. A continuous extraction system simultaneously drew air from the sampling port and vented it outside the testing room, thereby minimising residual odour accumulation between trials. Reinforcement for indication responses was delivered automatically via an automated feeder (see Figure 3), and all trial events, response timing, and durations were recorded automatically by custom control software.

Figure 2

Automated Olfactometer Apparatus

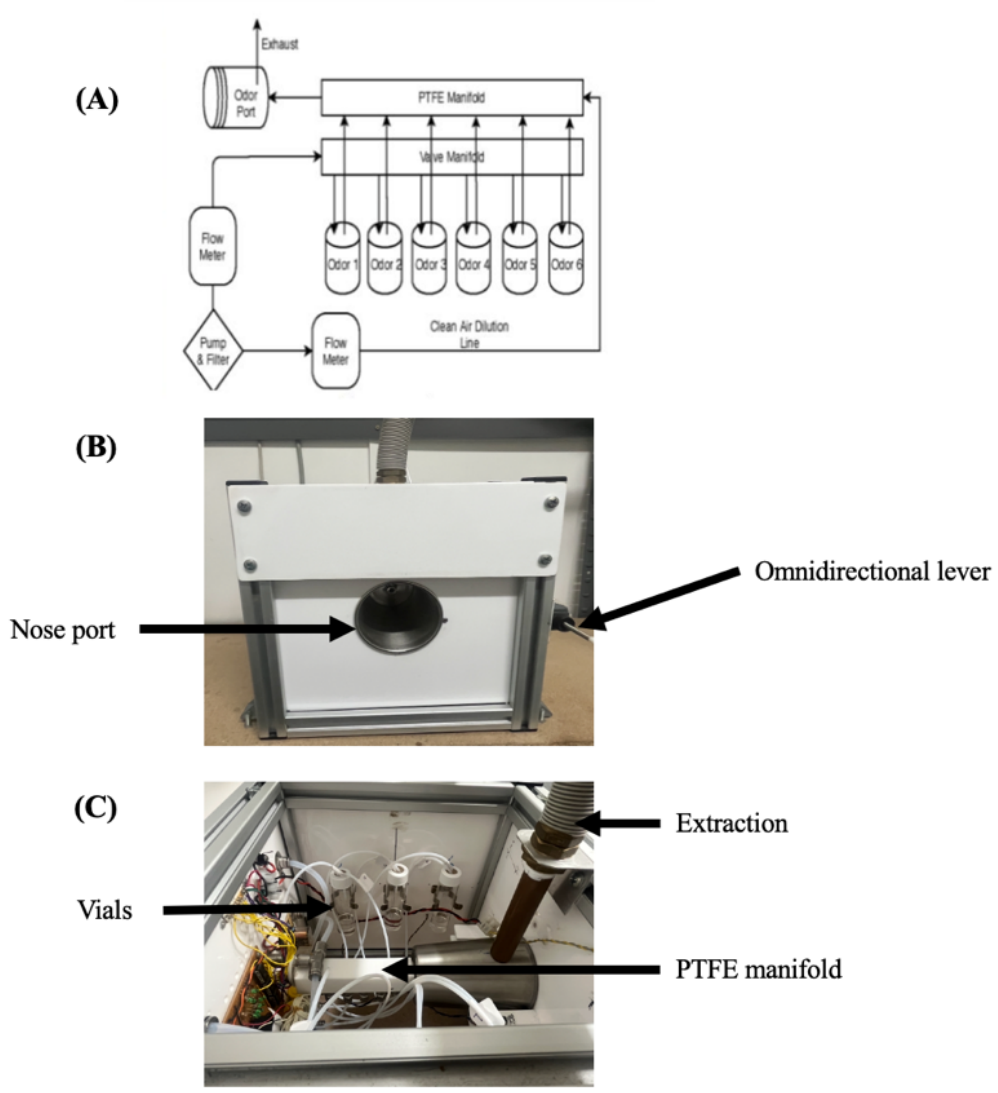


Figure 3

Automatic Feeder

Odour Stimuli

Two target odours were used: cinnamaldehyde and hexanoic acid. Across dogs, assignment of these odours to the primary (Target A) and secondary (Target B) conditions was counterbalanced. Two non-target odours were also used, which were benzyl acetate and linalool (denoted X and Y, respectively). The functional roles and corresponding stimulus labels are summarised in Table 3. All odours selected for this study were commonly used in scent detection research and were approved for use by the University of Waikato Animal Ethics Committee.

Table 3

Assignment of Odour Labels and Roles in the Discrimination and Probe Phases

Label	Odour	Role	Dilution (v/v)
A	Cinnamaldehyde	Primary target	1:2000
B	Hexanoic acid	Secondary target	1:2000
X	Benzyl acetate	Non-target	1:2000
Y	Linalool	Non-target	1:2000
A+	Cinnamaldehyde	Probe	undiluted
B+	Hexanoic acid	Probe	1:100
B++	Hexanoic acid	Probe	1:50

Note. Assignment of cinnamaldehyde and hexanoic acid to Target A and Target B roles was counterbalanced across dogs.

Additional non-target odours were introduced during later stages of the experiment to assess generalisation of non-target discrimination. These novel non-target odours included hexanol, ethyl acetate, propanol, amyl acetate, cyclohexanone, and *n*-hexyl acetate. All odour compounds were sourced from Sigma-Aldrich. Odours were prepared by diluting liquid odour compounds in mineral oil and storing them in sealed vials. For training samples, each odour was prepared at 1:2000 (v/v) dilution achieved by adding 10 µL of the odour component to 20 mL of mineral oil in sealed 40 mL glass vials. This dilution was selected to establish a training baseline that allowed systematic evaluation of generalisation across substantial nominal increases in odorant concentration in solution during probe trials and to ensure that odour presentation was detectably above baseline airflow during apparatus setup. These odours were used across initial training and discrimination stages. Air flow and delivery parameters were kept constant across conditions, ensuring that relative differences between odours were determined by liquid preparation.

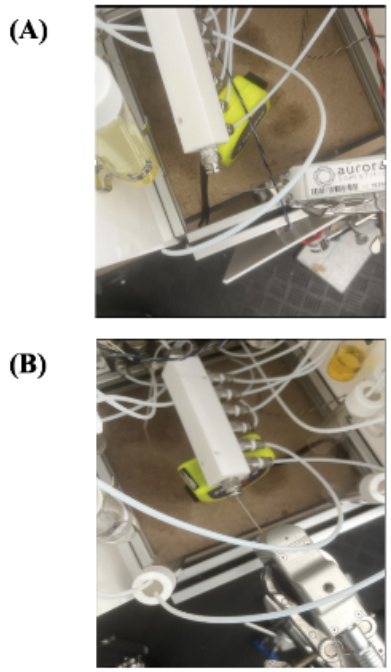
Probe trials aimed to deliver odour stimuli at approximately 20× baseline concentration at the sampling port, as assessed using a photoionisation detector (PID). Although an

approximately 20-fold increase could be verified for hexanoic acid during initial PID testing, this increase could not be reliably replicated subsequently. For cinnamaldehyde, a 20-fold increase could not be verified at the sampling port, and therefore an undiluted liquid sample was used with final odour delivery determined by carrier airflow dilution within the apparatus. PID testing yielded inconsistent quantitative readings, likely reflecting environmental and apparatus-related variability (see Appendix C). As a result, PID data were used for relative validation of sample preparation only and not for calculation of exact vapour-phase concentrations.

PID measurements were obtained using a 200B miniPID Fast Response Olfaction Sensor (Aurora Scientific, Canada). For initial trials, the PID probe was positioned approximately 1 cm from the manifold sampling port in a parallel orientation relative to airflow (see Figure 4A). In subsequent trials, the probe was positioned directly in front of the sampling port to increase signal detection sensitivity (see Figure 4B). During measurement, carrier airflow was switched off and odour flow was delivered at 0.8 L/min. Baseline air readings were recorded prior to each sample presentation, and subsequent odour samples were measured after a 10 s interval between presentations.

Figure 4

Positioning of the PID Probe During Testing



The B+ hexanoic acid probe was prepared at 1:100 (v/v) dilution by adding 200 µL of the odour component to 19.8 mL of mineral oil. The A + cinnamaldehyde probe was presented as an undiluted liquid sample of around 10 mL in sealed 40 mL glass vials. After initial probe testing, an additional probe test was conducted using a stronger dilution of hexanoic acid (B++). This odour was prepared at a 1:50 (v/v) dilution by adding 400 µL of hexanoic acid to 19.6 mL of mineral oil. All odour samples were prepared and stored under consistent conditions. Samples were refrigerated prior to use and equilibrated to room temperature before testing. During training phases, odour samples were replaced on a weekly basis. During probe testing sessions, fresh samples were prepared for all target, non-target, and probe stimuli to align with PID validation procedures and minimise variability related to sample age.

Procedure

Habituation and Feeder Training. The initial training procedure required habituation to the laboratory environment, experimental apparatus and the automated feeder. This included testing whether dogs approached the automatic feeder independently when it dispensed kibbles. Dogs were eligible to progress for shaping once they consistently engaged with the feeder without prompting.

Response Shaping and Sampling Behaviour. Shaping procedures were used to establish sampling behaviour at the odour port. Differential reinforcement of successive approximations was used to train dogs to orient toward, approach, and insert their nose into the sampling port, with nose entry detected by an infrared beam. Prompting was used sparingly during early shaping and was systematically faded as performance stabilised. A minimum duration threshold was defined as the length of time the infrared beam within the sampling port was required to remain interrupted for an indication response to be registered by the apparatus control software.

During initial shaping, the indication threshold was set at 500 ms to establish consistent beam interruption responding and was gradually increased via differential reinforcement of longer-duration nose holds, reaching approximately 1000–1500 ms.

For subsequent discrimination training, indication thresholds were further increased to promote sustained sampling and reduce impulsive responding, thereby strengthening stimulus control over indication behaviour. A minimum observation threshold of approximately 500 ms was implemented to ensure that dogs sampled the odour prior to emitting a rejection response.

A rejection response was defined as activation of the lever following interruption of the infrared beam within the sampling port that exceeded the minimum observation threshold but did not meet the indication threshold. Lever pressing was shaped using differential reinforcement,

and responses falling between the minimum observation threshold and indication threshold were classified as rejections. A lever press did not count as a rejection unless the minimum observation threshold had been met and the trial would remain active.

Odour Discrimination Training (A vs. X). Following acquisition of the response requirements, dogs were trained to discriminate between a primary target odour (A) and a non-target odour (X). Assignment of odours to the A and B target conditions was counterbalanced across dogs. Correct indication responses on target trials were reinforced, whereas responses on non-target trials were not reinforced. Correct rejections produced no programmed consequences, consistent with naturalistic detection contingencies and with applied detection procedures in which reinforcement is contingent on correct indications rather than target absence (Edwards et al., 2022). Each session consisted of 20 trials, comprising 8 target trials and 12 non-target trials. During initial acquisition, trial order was structured such that target and non-target trials alternated. Once indication and rejection responses were recorded independently for complete sessions and without prompting, trial order was randomised while maintaining the same target-to-non-target ratio. Indication duration thresholds were increased sequentially from 1500 ms in 500 ms increments contingent on criterion-level performance. Thresholds were progressively increased until the correct rejection rate stabilized above or at 80% reaching approximately 6000–8000 ms. Although final threshold requirements were adjusted individually based on hit rate and correct rejection accuracy, this range is consistent with prior automated olfactometer research indicating that stable detection performance is often observed at higher indication duration requirements (Edwards et al., 2022). Progression through this phase and progression through this and all subsequent phases was determined using performance-based criteria

, requiring both hit rate and correct rejection rate to be at or above 75% for three consecutive sessions.

Odour Discrimination Training (B vs. Y). Once discrimination for target odour A against non-target odour X was obtained according to performance criteria, a second target odour (B) and non-target (Y) were introduced. Training procedures and protocols mirrored those used previously including reinforcement contingencies, trial sequencing and progression criteria. Thresholds were initially maintained at 1500 ms during early sessions and were subsequently increased sequentially contingent on stable performance at the previously achieved threshold. Progression to the next phase required both hit rate and correct rejection rate to reach at least 75% for three consecutive sessions.

Odour Discrimination Training (A + B vs. X + Y). Once discrimination performance met criteria, dogs progressed to sessions in which both targets (A + B) and both non-targets (X + Y) were presented within the same session. Each session consisted of 20 trials: 8 target trials and 12 non-target trials. Of the target trials, four involved odour A and four involved target odour B. The non-target trials were divided equally, with six trials consisting of non-target odour X and six consisting of the non-target odour Y. The indication threshold was maintained at the previously established working level for each dog, as determined in earlier experimental phases. Correct indications on either target odour were reinforced equivalently, and performance was assessed separately for each stimulus to confirm maintenance of discrimination across multiple stimulus classes. The performance criterion for this phase required a hit rate of at least 75% for Target A and B, evaluated separately, and a correct rejection rate of at least 75% for the non-targets calculated separately for three consecutive sessions.

Novel Non-Target Generalisation. To assess stimulus control and generalisation of non-target rejection, novel non-target odours were subsequently introduced. Dogs were required to reject unfamiliar non-target odour combinations while continuing to correctly indicate target odours. Novel non-target combinations were introduced sequentially once dogs met the standard progression criteria. When a novel non-target combination was introduced, dogs were required to achieve ≥75% hit rate and ≥75% correct rejection rate across two consecutive sessions. If criterion performance was achieved during the initial exposure session and maintained during the subsequent session, this was considered indicative of generalisation of discrimination performance to the novel non-target combination, and the dog progressed to the next phase of the study. If criterion performance was not met on first exposure, training continued with the same combination until both hit rate and correct rejection rate reached ≥75% across two consecutive sessions. At that point, a new novel non-target combination was introduced, and the evaluation process was repeated.

Intermittent Reinforcement. Once discrimination performance was stable, reinforcement for correct target indications was gradually thinned using an intermittent reinforcement schedule. Sessions consisted of 20 trials: 8 target trials and 12 non-target trials. During the initial phase of reinforcement thinning, two of the eight target trials per session were unreinforced, following a variable-ratio (VR) 1.33 reinforcement schedule. Criteria to move to the next phase required a hit rate of at least 75% and a correct rejection rate of at least 75% across two consecutive sessions. In the next phase, the same session structure and trial proportions were maintained; however, four of the eight target trials per session were unreinforced, following a variable-ratio (VR) 2 reinforcement schedule. Criteria to move to the next phase required maintenance of a hit rate of at least 75% and a correct rejection rate of at least 75% across two consecutive sessions. This

phase was implemented to maintain responding under reduced reinforcement density prior to probe testing. Dogs advanced to probe testing once performance criteria under intermittent reinforcement were met.

Probe testing. The final phase involved non-reinforced probe trials embedded within regular training sessions to assess spontaneous generalisation to untrained changes in target odour concentration. Probe trials presented the trained target odours (A, B) at increased concentrations relative to training conditions (A+, B+). Probe trials were embedded within regular training sessions and were not reinforced. Each probe session comprised 20 randomised trials, including eight target trials (A and B), ten non-target trials (X and Y), and two probe trials (A+ or B+). Target trials were reinforced on a VR 1.33 schedule, with six of eight target trials reinforced per session, while indications on non-target and probe trials produced no programmed consequences. Each dog completed approximately 10–12 probe trials per high-concentration target odour, achieved across multiple sessions. The probe phase concluded once all dogs had completed this exposure, allowing assessment of spontaneous generalisation to increased target odour concentration in the absence of direct training.

Data Analysis

Performance was analysed using trial-by-trial responses recorded automatically by the apparatus. Each trial was classified based on the dog's response to the presented odour stimulus. A hit was defined as a sustained nose hold at the sampling port that exceeded the positive indication threshold during a target trial. A correct rejection was defined as sampling the port for at least the minimum observation threshold followed by pressing of the adjacent lever, without exceeding the positive indication threshold, during a non-target trial.

1. Hit rate, defined as the percentage of target trials on which a correct indication response occurred:

$$\text{Hit rate} = \frac{\text{Number of hits}}{\text{Total target trials}} \times 100$$

2. Correct rejection rate (CRR), defined as the percentage of non-target trials on which a correct rejection response occurred:

$$\text{CRR} = \frac{\text{Number of correct rejections}}{\text{Total non-target trials}} \times 100$$

Binomial 95% confidence intervals were calculated for probe indication probabilities using the Wilson score method implemented in the Epitools epidemiological calculators (EpiTools, n.d.).

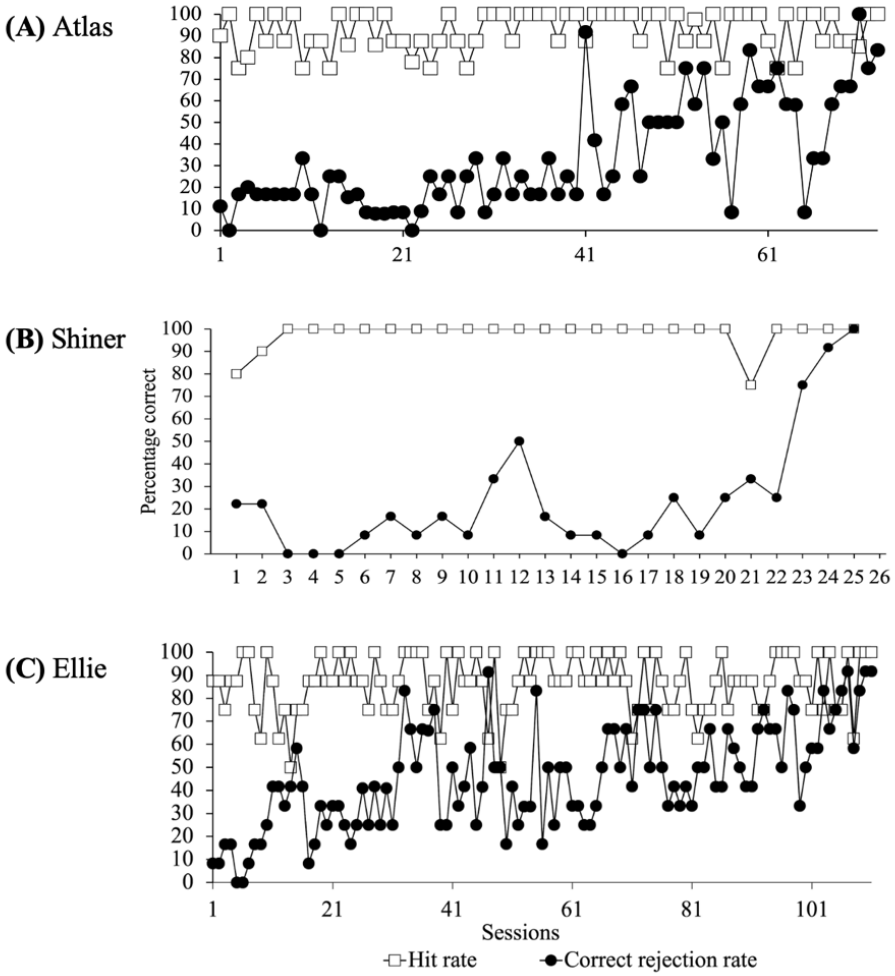
Results

Primary Target Discrimination

All three dogs participated in the primary target discrimination training phase upon completing initial preliminary training. Ellie met the discrimination criteria for target odour A (≥75% hit rate and ≥75% correct rejection rate for three consecutive sessions) after a total of 111 sessions. Shiner met the same criteria for discrimination of odour A after 26 sessions. Atlas met the same criteria for odour A after 73 sessions. Odour assignments were counterbalanced; cinnamaldehyde was designated as odour A for Ellie and Shiner, and hexanoic acid was designated as odour A for Atlas. Indication thresholds for Ellie stabilised at approximately 11,500 ms, for Shiner at approximately 3,500 ms, and for Atlas at approximately 7,000 ms. For all three dogs, acquisition of correct rejection responses lagged behind acquisition of target indication responses. Indication performance across sessions for primary target discrimination is shown in Figure 5.

Figure 5

Hit Rate and Correct Rejection Rate Across Sessions During Primary Target Discrimination Training (Odour A) for Atlas, Shiner, and Ellie



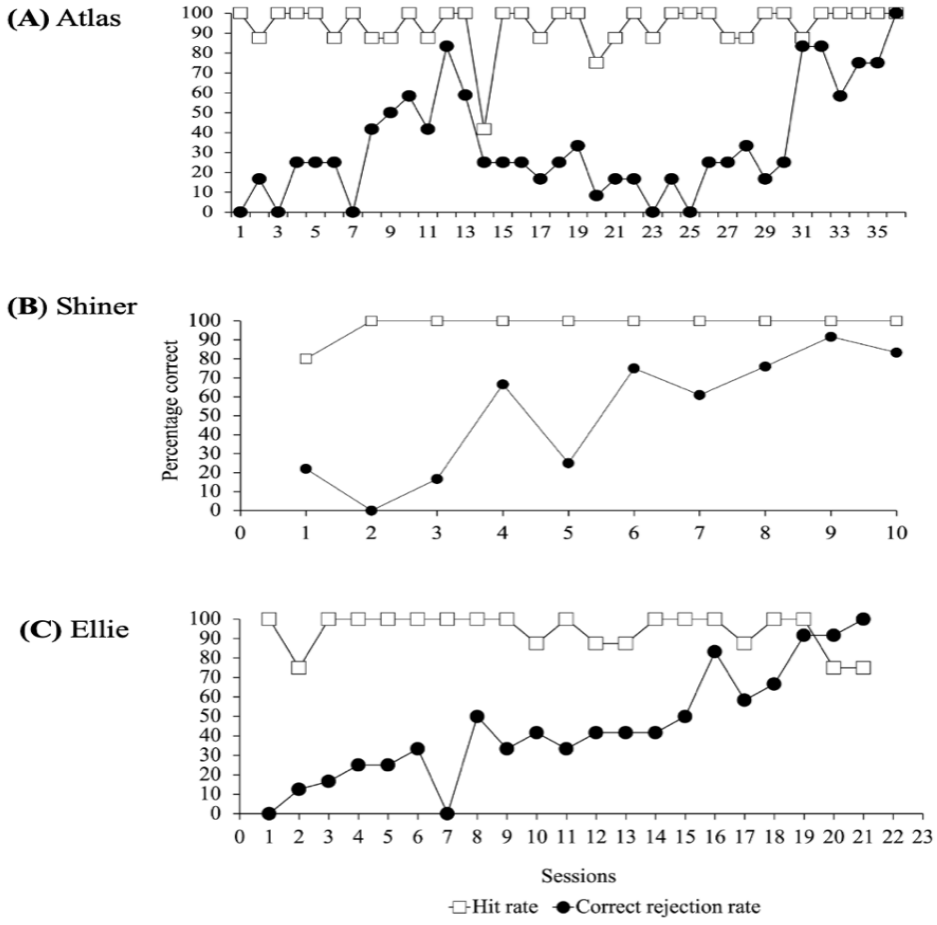
Secondary Target Discrimination

All three dogs participated in secondary target discrimination training upon meeting criteria for primary target training. Ellie met the criteria for this phase (≥75% hit rate and ≥75% correct rejection rate across three consecutive sessions) after 23 sessions. Atlas met the same criteria after 36 sessions. Shiner met the criteria after 10 sessions. Working indication thresholds for all dogs were maintained at the previously achieved levels, as stable discrimination

performance was observed at these thresholds. Across dogs, acquisition of discrimination for target odour B required fewer training sessions than acquisition of discrimination for target odour A. Performance across sessions for secondary target discrimination is shown in Figure 6.

Figure 6

Hit Rate and Correct Rejection Rate Across Sessions During Secondary Target Discrimination Training (Odour B) for Atlas, Shiner, and Ellie



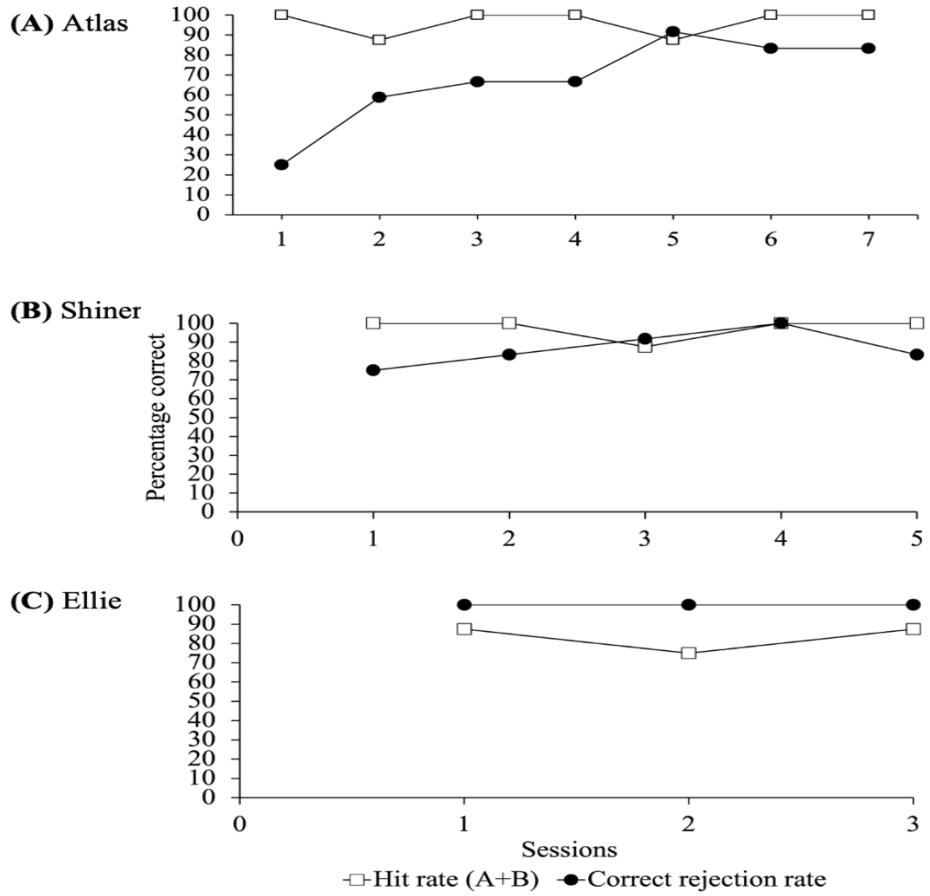
Mixed Target Discrimination

Ellie, Shiner, and Atlas all progressed to this phase upon meeting criteria for secondary target discrimination training. Ellie completed this phase and met the criteria (≥75% hit rate for both target odours within the same sessions and ≥75% correct rejection rate across three consecutive sessions) after three sessions. Shiner met the same criteria after five sessions. Atlas met the criteria after seven sessions. Acquisition of the mixed target discrimination phase

required fewer sessions across all dogs compared to earlier discrimination phases. Performance across sessions during mixed target discrimination training is shown in Figure 7.

Figure 7

Hit Rate (A + B) and Correct Rejection Rate Across Sessions During Mixed Target Discrimination Training for Atlas



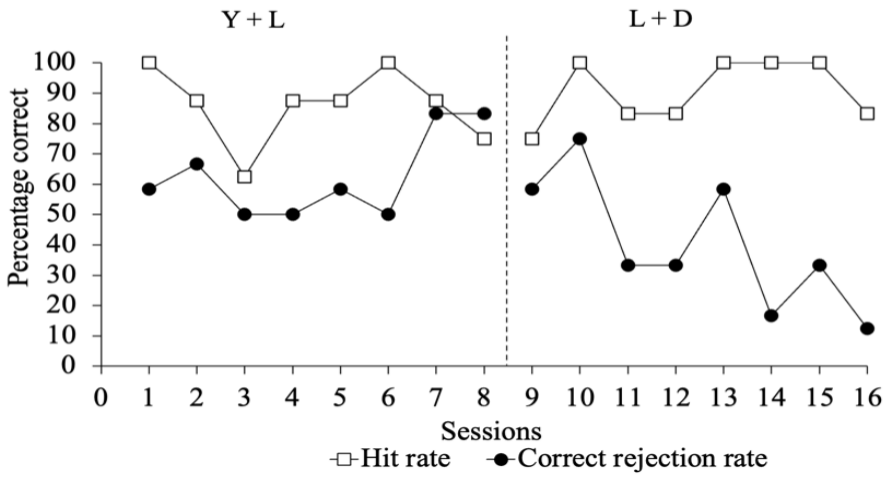
Note. Hit rate values represent a combined hit rate across target odours A and B. Progression criteria were evaluated separately for each target odour, as specified in Methods.

Novel Non-Target Generalisation

All three dogs participated in this phase of the study. The criteria for progression from this phase required that a novel combination of non-target odours not previously encountered be rejected with a correct rejection rate of at least 75% on first introduction, while maintaining a hit rate of at least 75% for target odours. This criteria assessed whether rejection responding generalized to unfamiliar non-target combinations rather than reflecting learning of specific distractor sets. Shiner rejected the first novel non-target combination (Y + L) at criteria levels and progressed to the subsequent phase. Ellie was exposed sequentially to the non-target combinations Y + L, L + D, and D + K, meeting the progression criteria for each. When a further novel non-target combination (K + F) was introduced, Ellie rejected this combination on first exposure and maintained criteria-level performance across subsequent trials. Atlas did not progress from this phase; although he met the rejection criteria for the initial non-target combination (Y + L), performance declined following the introduction of the subsequent combination (L + D), and criteria-level rejection was not maintained at this stage. Hit rate and correct rejection rate across sessions during novel non-target discrimination are shown in Figures 8–10.

Figure 8

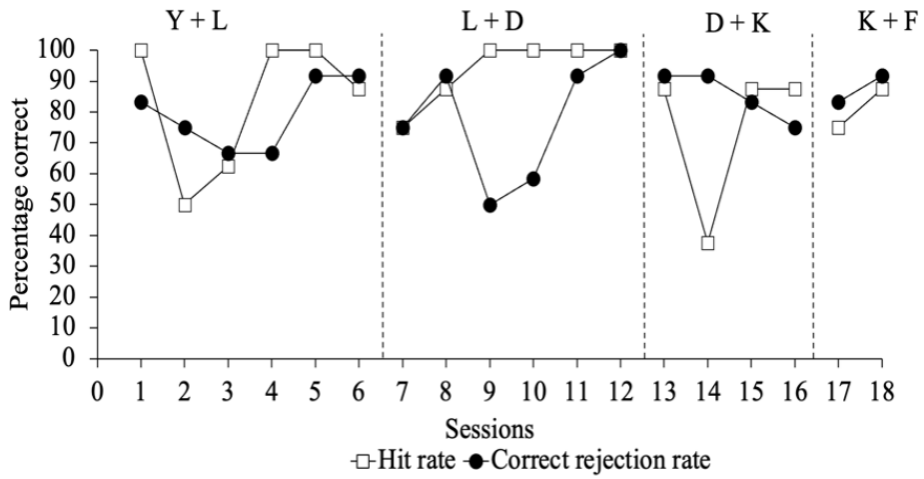
Hit Rate and Correct Rejection Rate Across Sessions During Novel Non-Target Discrimination for Atlas



Note. Phase labels indicate the non-target odour combinations presented during each block of sessions. Dashed vertical lines denote transitions between blocks in which different non-target odour combinations were used. See Table 2 for details of the odour combinations.

Figure 9

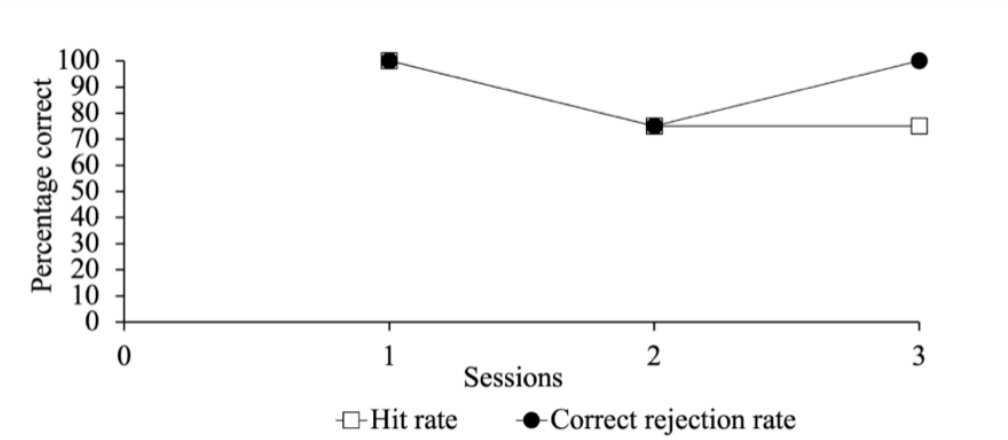
Hit Rate and Correct Rejection Rate Across Sessions During Novel Non-Target Discrimination for Ellie



Note. Phase labels indicate the non-target odour combinations presented during each block of sessions. Dashed vertical lines denote transitions between blocks in which different non-target odour combinations were used. See Table 2 for details of the odour combinations.

Figure 10

Hit Rate and Correct Rejection Rate Across Sessions During Novel Non-Target Discrimination for Shiner



Note. Phase labels above the graph indicate the non-target odour combinations presented during each block of sessions. See Table 2 for details of the odour combinations.

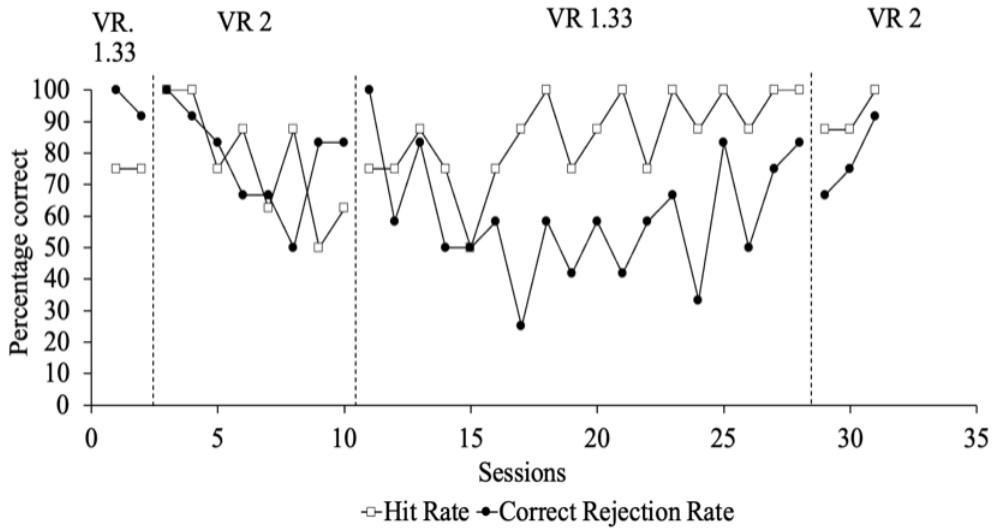
Intermittent Reinforcement Training

Only Ellie and Shiner progressed to the intermittent reinforcement phase. Following re-establishment of stable performance under full reinforcement, Ellie met the progression criteria under a variable-ratio (VR) 1.33 schedule after two sessions (≥75% hit rate and ≥75% correct rejection rate). Reinforcement was then thinned to a VR 2 schedule. Under this condition, performance declined and responding ceased, and Ellie did not meet criteria. Training was therefore returned to the VR 1.33 schedule, under which criteria-level performance was re-established after 18 sessions. Ellie subsequently progressed again to the VR 2 schedule and met criteria for advancement to probe testing after three sessions. A similar pattern was observed for Shiner. After meeting criteria under the VR 1.33 schedule, performance declined following the

transition to the VR 2 schedule, and criteria-level responding was not maintained. Training was returned to the VR 1.33 schedule, and criteria-level performance was re-established. Shiner then progressed to the VR 2 schedule and met criteria to advance to probe testing after two sessions. Performance across sessions during intermittent reinforcement training for Ellie and Shiner is shown in Figures 11–12.

Figure 11

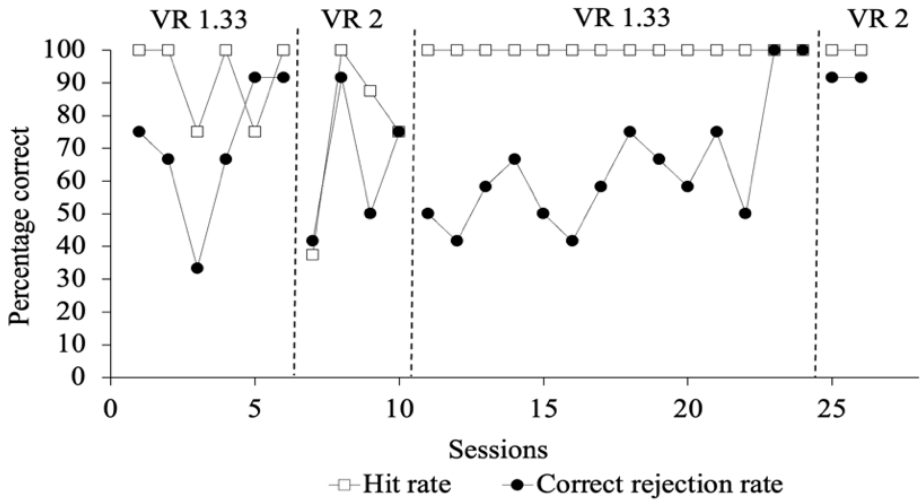
Hit Rate and Correct Rejection Rate Across Sessions During Intermittent Reinforcement Training for Ellie



Note. VR = variable-ratio schedule of reinforcement; numbers indicate the average number of responses required for reinforcement. Phase labels indicate the schedule in effect; dashed lines denote schedule changes.

Figure 12

Hit Rate and Correct Rejection Rate Across Sessions During Intermittent Reinforcement Training for Shiner



Note. VR = variable-ratio schedule of reinforcement; numbers indicate the average number of responses required for reinforcement. Phase labels indicate the schedule in effect; dashed lines denote schedule changes.

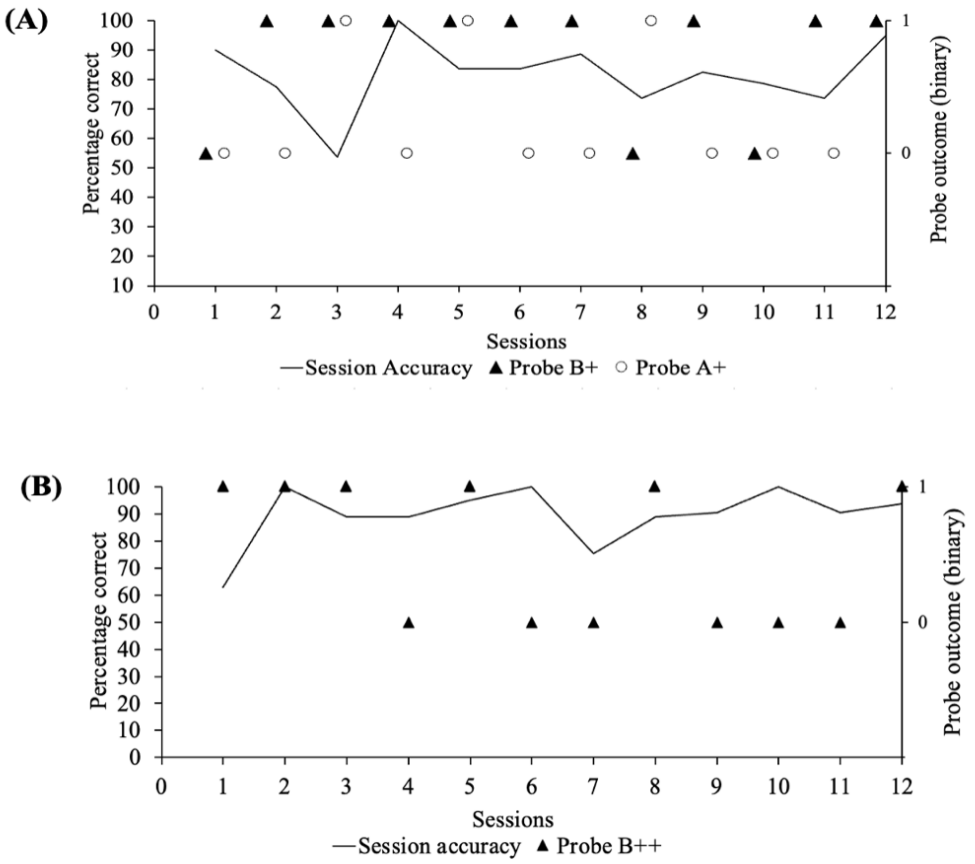
Probe Testing: Ellie

During the probe testing phase, Ellie completed non-reinforced probe trials assessing responding to increased target odour concentrations. Session-by-session probe outcomes alongside concurrent discrimination accuracy are shown in Figure 13. Summary probabilities of indication for probe trials, false alarms, and trained target trials with 95% binomial confidence intervals are shown in Figures 14 and 15. The false-alarm rate reflects the probability of an indication on non-target trials (i.e., when no target odour was present) and provides a baseline level of responding against which probe responses can be compared. Ellie showed an elevated probability of indication for the B+ probe relative to the false-alarm rate, whereas responding to

the A+ probe overlapped with false-alarm responding. At the higher concentration probe (B++), indication probability again exceeded the false-alarm rate while remaining below the trained hit rate. Session accuracy remained high throughout probe testing.

Figure 13

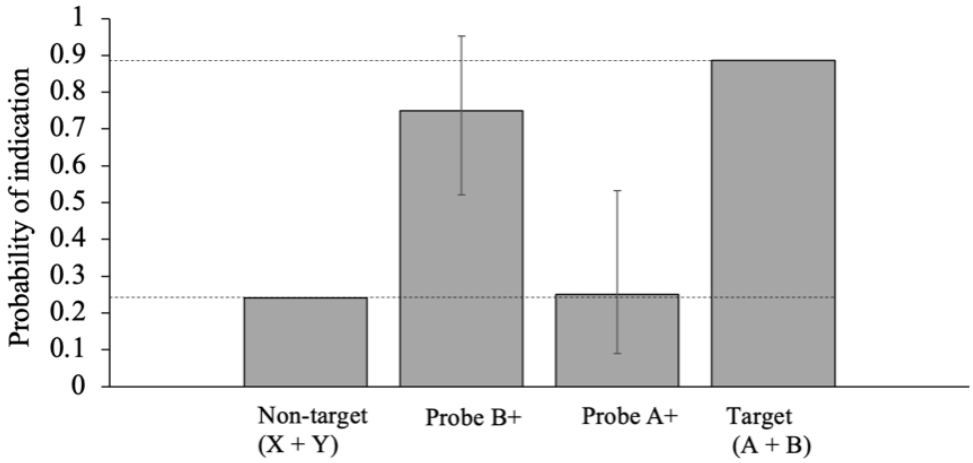
Session-by-Session Probe Outcomes Alongside Concurrent Discrimination Accuracy for Ellie



Note. Panel A shows session-by-session discrimination accuracy alongside outcomes of probe trials for the increased odour concentrations B+ and A+. Panel B shows session-by-session discrimination accuracy alongside outcomes of probe trials for the higher concentration B++. Session accuracy represents the percentage of correct responses on discrimination trials within each session. Probe outcomes are plotted on the secondary axis as binary responses (1 = indication; 0 = no indication).

Figure 14

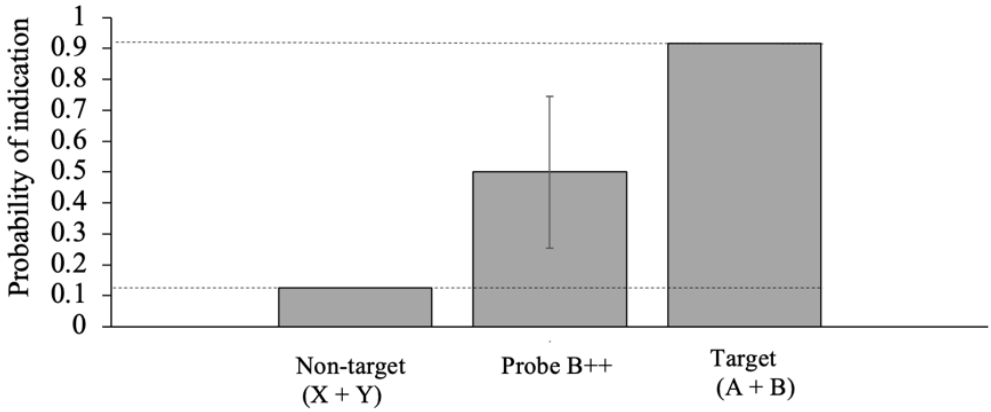
Probability of Indication for Probe Trials (A+ and B+), False Alarms, and Trained Target Trials for Ellie with Binomial Confidence Intervals



Note. A and B denote the target odours used during training (A = cinnamaldehyde; B = hexanoic acid).

Figure 15

Probability of Indication for Probe Trials (B++), False Alarms, and Trained Target Trials for Ellie with Binomial Confidence Intervals



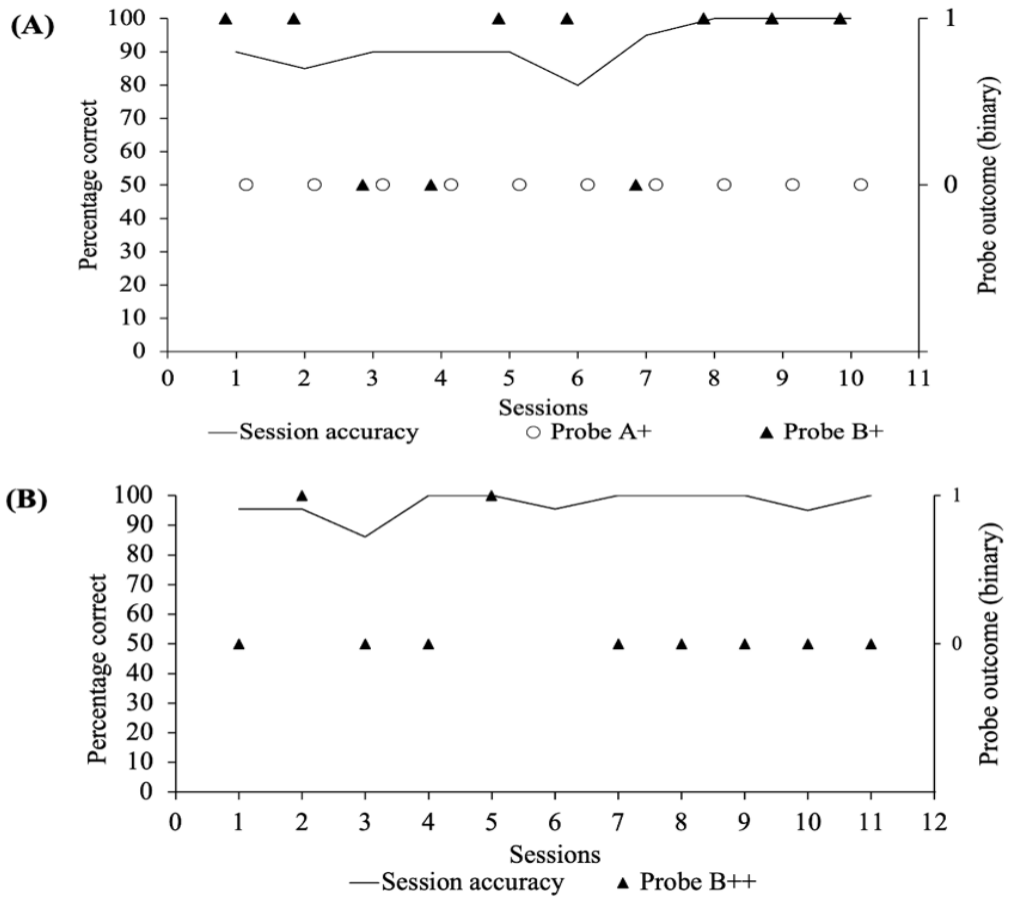
Note. A and B denote the target odours used during training (A = cinnamaldehyde; B = hexanoic acid).

Probe Testing: Shiner

Probe testing was conducted for Shiner as well to assess responding to increased target odour concentrations during non-reinforced trials, with concurrent measures of discrimination accuracy maintained throughout. Session-by-session probe outcomes alongside concurrent discrimination accuracy are shown in Figure 16. Summary probabilities of indication for probe trials, false alarms, and trained target trials with 95% binomial confidence intervals are shown in Figures 17 and 18. Shiner showed elevated probabilities of indication for both the B+ and B++ probe conditions relative to the false-alarm rate. In contrast, responding to the A+ probe overlapped with false-alarm responding. Session accuracy remained high during probe sessions.

Figure 16

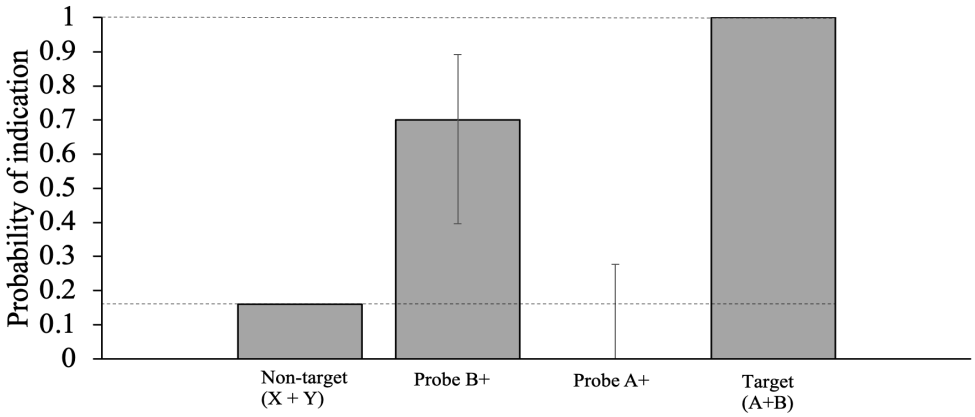
Session-by-Session Probe Outcomes Alongside Concurrent Discrimination Accuracy for Shiner



Note. Panel A shows session-by-session discrimination accuracy alongside outcomes of probe trials for the increased odour concentrations B+ and A+. Panel B shows session-by-session discrimination accuracy alongside outcomes of probe trials for the higher concentration B++. Session accuracy represents the percentage of correct responses on discrimination trials within each session. Probe outcomes are plotted on the secondary axis as binary responses (1 = indication; 0 = no indication).

Figure 17

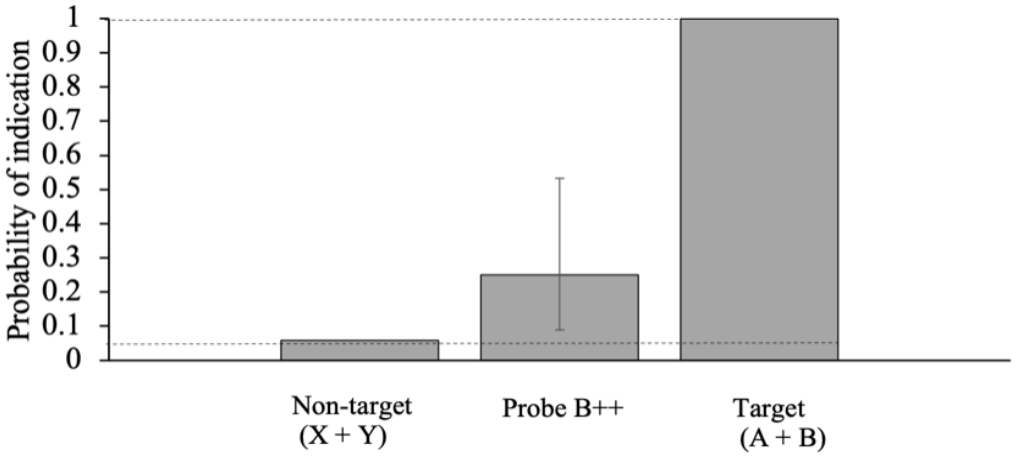
Probability of Indication for Probe Trials (A+ and B+), False Alarms, and Trained Target Trials for Shiner with Binomial Confidence Intervals



Note. A and B denote the target odours used during training (A = cinnamaldehyde; B = hexanoic acid).

Figure 18

Probability of Indication for Probe Trials (B++), False Alarms, and Trained Target Trials for Shiner with Binomial Confidence Intervals



Note. A and B denote the target odours used during training (A = cinnamaldehyde; B = hexanoic acid).

Discussion

The present study examined whether dogs trained to detect target odours at a single concentration would spontaneously generalise responding to higher concentrations of those same odours. Consistent with the study hypothesis, the results suggest that concentration changes affected established stimulus control. Across the two dogs that completed probe testing, responding on higher concentration probe trials did not match responding on trained target trials. Instead, probe performance showed a graded and odour-specific pattern. For both Ellie and Shiner, indications on the higher concentration hexanoic acid probes (B+ and B++) occurred more frequently than false-alarm indications, but less frequently than indications on trained target trials. In contrast, indications on the higher concentration cinnamaldehyde probe (A+) occurred at levels comparable to false alarms for both dogs. This pattern is consistent with partial generalisation across concentration, but not full transfer of stimulus control from the trained concentration to higher concentration variants of the same odour.

Summary of Overall Findings

Probe responding differed across odours when concentration was increased, indicating that concentration changes influenced stimulus control differently for each odour. For both dogs, responding to the B+ probe (hexanoic acid for these dogs; see Table 2 for odour assignments) occurred at levels above false alarms but below the hit rate, with confidence intervals overlapping these ranges. This pattern suggests that the dogs treated the B+ stimulus as partially similar to the trained target, producing intermediate levels of responding (DeChant et al., 2021a; Moser et al., 2019). In contrast, responding to the A+ probe (cinnamaldehyde for these dogs) overlapped substantially with false-alarm responding, indicating that the dogs did not reliably treat this probe as a target stimulus. Intermediate responding was also observed for the B++

probe, although confidence intervals were wide due to the limited number of probe trials. Overall, the pattern across dogs suggests that the B+ and B++ probes preserved stimulus control to some extent, whereas the A+ probe did not reliably occasion the trained indication response.

The probe data are most readily interpreted within a generalisation-gradient account of stimulus control. In this framework, the training stimulus occasions the highest probability of the trained response, and response probability declines as test stimuli differ from the training value along a relevant dimension (Guttman & Kalish, 1956; Jenkins & Harrison, 1960; Moser et al., 2019). In the present study, trained target odours produced the highest indication rates, B+ and B++ produced intermediate levels of responding, and false alarms (i.e., indications to non-target stimuli) occurred at the lowest rates. This interpretation is consistent with prior work suggesting that generalisation across odour concentration changes in detection dogs may be limited (DeChant et al., 2021a).

The asymmetry between odours may also reflect differences in vapour-phase behaviour rather than differences in generalisation processes alone. Hexanoic acid and cinnamaldehyde differ in volatility and vapour pressure, which influences how liquid dilutions translate into vapour stimuli delivered to the sampling port. Hexanoic acid has a reported vapour pressure of approximately 0.0435 mm Hg at 25 °C (PubChem, n.d.-b), whereas cinnamaldehyde has a lower vapour pressure of approximately 0.0289 mm Hg at 25 °C (PubChem, n.d.-a). Odorants with higher vapour pressure may tend to enter the headspace more readily, which can produce more stable vapour delivery under airflow conditions (Sloan et al., 2025; Conchou et al., 2026). Under the apparatus conditions used in the present study, increases in liquid concentration did not necessarily produce proportional increases in vapour-phase concentration. The PID traces (see Appendix C) showed relative increases in signal but were not sufficiently stable to establish

precise vapour equivalence across odours. Consequently, increasing the liquid concentration of cinnamaldehyde may not have produced the same proportional change in vapour stimulus as the corresponding increase in hexanoic acid concentration.

The present findings are consistent with previous work showing that concentration changes can alter stimulus control in detection dogs. DeChant et al. (2021a) reported that dogs trained at a single baseline concentration showed reliable responding to nearby concentrations but declining responding as test concentrations moved further from the training stimulus. Similarly, Aviles-Rosa et al. (2021) described failures of explosive detection dogs to respond reliably to substantially larger target quantities than those encountered during training. Although the present study used different odours and procedures, the pattern of probe responding observed here is broadly consistent with these findings.

Discrimination Training

The earlier training phases provide important context for understanding the probe findings. During initial discrimination training, hit rates rose before correct rejection rates stabilised across all dogs. Early in training, the indication response occurred on both target and non-target trials. Across repeated sessions, responding on non-target trials declined while responding on target trials remained high, resulting in gradual increases in correct rejection performance. This pattern is consistent with the contingencies arranged in the task. Target indications produced reinforcement, whereas responding on non-target trials did not. Under these conditions, high rates of early indicating and gradual reduction in responding on non-target trials would be expected as extinction effects on non-target trials accumulated across sessions (Bouton, 2018).

The increase in the indication threshold likely contributed to this change in response pattern. As the required nose-hold duration increased, incorrect indications became more

effortful and less likely relative to quickly rejecting the stimulus and moving on to the next trial (Edwards et al., 2022). Edwards et al. (2022) reported that increasing the response requirement for indication can shift response bias while leaving detection sensitivity relatively stable. Early in training, high rates of indicating may have been maintained because indication was the response class most directly associated with reinforcement. As the threshold criterion increased, the effort cost of incorrect indications also increased, making indiscriminate indicating less efficient. The lag in CRR is therefore best interpreted as an expected product of the contingencies arranged during discrimination training rather than as an anomalous feature of the procedure.

Faster Acquisition of the Second Discrimination

A second pattern in the training results was that discrimination of the second target odour was established in fewer sessions than discrimination of the first. Ellie required 111 sessions to meet criterion for odour A but 23 for odour B; Shiner required 26 versus 10; and Atlas required 73 versus 36. The mixed target discrimination phase was established even more rapidly. These results suggest that, by the time the second discrimination phase was introduced, the relevant response topography was already well established. Similar patterns have been reported in detection dog training, where acquisition of additional target odours can occur more rapidly once initial odour discriminations have been learned (Waggoner et al., 2022).

Novel Non-Target Generalisation

Performance differed across dogs. Shiner rejected the first novel non-target combination at criterion levels on initial exposure. Ellie did not show that immediate pattern, but stable rejection responding emerged across successive novel combinations, and a later combination was rejected on first exposure. Atlas did not progress beyond this phase. These differences suggest variability in how stimulus control developed across individual dogs (Guest et al., 2020). For

Shiner, rejection responses generalised beyond the specific non-target combinations used during training relatively quickly. For Ellie, broader rejection control appeared to emerge more gradually across additional exemplars. For Atlas, one possible interpretation is that stimulus control by the target odours was weaker, with responding more strongly organised around the specific non-target combinations encountered during training. Including successive novel non-target combinations also served a methodological purpose, as tests of generalisation are strengthened when they assess responding to novel distractors rather than only familiar training stimuli (Lazarowski et al., 2021).

Implications

The present findings have applied relevance because variation in odour concentration is a common difference between training conditions and operational detection environments. Detection dogs may encounter targets at concentrations far outside the range represented in training due to variation in quantity, packaging, concealment, environmental dilution, or airflow conditions (Furton & Myers, 2001). If dogs trained at a single concentration do not reliably respond to other concentrations of the same odour, operational performance may be affected even when the target substance itself has not changed.

The findings are therefore consistent with concerns raised in operational research. Aviles-Rosa et al. (2021) reported that detection dogs trained on smaller quantities of explosive material did not reliably respond to substantially larger quantities until additional exposure occurred. The present study does not directly replicate that case, but it supports a compatible interpretation: changes in concentration alone may weaken established stimulus control. In that sense, concentration should be treated as a training-relevant stimulus dimension rather than as an incidental physical property. Exposure to multiple concentration exemplars during training may

support more reliable generalisation under operational conditions (DeChant et al., 2021a; Moser et al., 2019).

The broader research programme from which the present study was drawn aimed to test whether training across multiple concentrations of a secondary target odour could broaden generalisation to a primary target odour. Because time constraints prevented progression to that phase, the present study addresses spontaneous generalisation only. Nevertheless, the probe results clarify why such procedures may be useful. If spontaneous generalisation across concentration is limited, then systematic concentration training may be required to produce broader and more stable stimulus control without requiring direct training with hazardous or operationally impractical quantities of target material.

Limitations

Several limitations should be acknowledged. First, the number of dogs in the study was small, and only two dogs progressed to the probe phase. Second, the number of probe trials was limited. Third, the two odours differed in volatility and vapour behaviour, and the liquid dilution procedure used here may not have translated linearly into vapour-phase concentration. Finally, although PID measurements documented relative changes in signal, they did not provide fully stable quantitative validation of vapour concentration. Interpretations of odour-specific probe differences should therefore remain cautious.

Future Research

Future research would benefit from more precise control of vapour delivery. Odour delivery systems incorporating electronically controlled valves or direct vapour-phase regulation could allow concentration to be manipulated more accurately and reduce uncertainty about how liquid concentration maps onto the stimulus encountered at the sampling port. Such systems

would strengthen interpretation of concentration generalisation effects. A second priority would be the inclusion of additional target odours with clearer or more predictable vapour profiles. Comparing odours with different volatility characteristics would help determine whether the asymmetry observed in the present study reflects physical delivery properties or broader features of concentration generalisation in detection dogs. A further extension of this research programme would be to examine how explicit training across multiple odour concentrations influences generalisation. The present study assessed spontaneous generalisation following training at a single concentration, providing an initial indication of how changes in concentration affect stimulus control.

Conclusion

The present study suggests that dogs trained at a single odour concentration may not reliably generalise responding to higher concentrations of that same odour. Evidence from the probe phase showed partial responding to higher concentration hexanoic acid probes but little responding to the higher concentration cinnamaldehyde probe. These findings support the interpretation that concentration changes can disrupt established stimulus control and that such disruptions may vary across odours depending on both behavioural and physical factors. More broadly, the results highlight the importance of treating concentration as a relevant training dimension in scent detection work. Training procedures that incorporate systematic variation in concentration may therefore be necessary to promote robust generalisation across operational conditions.

References

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

Bouton, M. E. (2018). Extinction of instrumental (operant) learning: interference, varieties of context, and mechanisms of contextual control. *Psychopharmacology*, *236*(1), 7–19. https://doi.org/10.1007/s00213-018-5076-4

Caldicott, L., Pike, T. W., Zulch, H. E., Mills, D. S., Williams, F. J., Elliker, K. R., Hutchings, B., & Wilkinson, A. (2024). Odour generalisation and detection dog training. *Animal Cognition*, *27*(1), 73. https://doi.org/10.1007/s10071-024-01907-0

Cleland, T. A., Morse, A., Yue, E. L., & Linster, C. (2002). Behavioral models of odor similarity. *Behavioral Neuroscience*, *116*(2), 222–231. https://doi.org/10.1037/0735-7044.116.2.222

Conchou, L., Trelea, I., Gévar, J., Monsempès, C., Peron, A., Renou, M., Souchon, I., & Lucas, P. (2026). A method to estimate absolute odorant concentration of olfactory stimuli. *PLoS ONE*, *21*(1), e0337336. https://doi.org/10.1371/journal.pone.0337336

Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied Behavior Analysis*. Pearson UK.

DeChant, M. T., Bunker, P. C., & Hall, N. J. (2021a). Stimulus control of odorant concentration: Pilot study of generalization and discrimination of odor concentration in canines. *Animals*, *11*(2), 326. https://doi.org/10.3390/ani11020326

DeChant, M. T., & Hall, N. J. (2021b). Training with varying odor concentrations: implications for odor detection thresholds in canines. *Animal Cognition*, *24*(4), 889–896. https://doi.org/10.1007/s10071-021-01484-6

Edwards, T. L., Giezen, C., & Browne, C. M. (2022). Influences of indication response requirement and target prevalence on dogs' performance in a scent-detection task. *Applied Animal Behaviour Science*, *253*, 105657. https://doi.org/10.1016/j.applanim.2022.105657

EpiTools - Home. (n.d.). Retrieved March 15, 2026, from https://epitools.ausvet.com.au/

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. (2004). The role of context specificity in learning: the effects of training context on explosives detection in dogs. *Animal Cognition*, *8*(3), 143–150. https://doi.org/10.1007/s10071-004-0236-9

Gross-Isseroff, R., & Lancet, D. (1988). Concentration-dependent changes of perceived odor quality. *Chemical Senses*, *13*(2), 191–204. https://doi.org/10.1093/chemse/13.2.191

Guest, C. M., Harris, R., Anjum, I., Concha, A. R., & Rooney, N. J. (2020). A lesson in standardization – subtle aspects of the processing of samples can greatly affect dogs' learning. *Frontiers in Veterinary Science*, *7*, 525. https://doi.org/10.3389/fvets.2020.00525

Guttman, N., & Kalish, H. I. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology*, *51*(1), 79–88. https://doi.org/10.1037/h0046219

Howard, R. W. (2000). Generalization and Transfer: an interrelation of paradigms and a taxonomy of knowledge extension processes. *Review of General Psychology*, *4*(3), 211–237. https://doi.org/10.1037/1089-2680.4.3.211

Jenkins, H. M., & Harrison, R. H. (1960). Effect of discrimination training on auditory generalization. *Journal of Experimental Psychology*, *59*(4), 246–253. https://doi.org/10.1037/h0041661

Lazarowski, L., Krichbaum, S., DeGreeff, L. E., Simon, A., Singletary, M., Angle, C., & Waggoner, L. P. (2020). Methodological considerations in canine olfactory detection research. *Frontiers in Veterinary Science*, *7*, 408. https://doi.org/10.3389/fvets.2020.00408

Lazarowski, L., Waggoner, P., Hutchings, B., Angle, C., & Porritt, F. (2021). Maintaining long-term odor memory and detection performance in dogs. *Applied Animal Behaviour Science*, *238*, 105301. https://doi.org/10.1016/j.applanim.2021.105301

Moser, A. Y., Bizo, L., & Brown, W. Y. (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. *Applied Animal Behaviour Science*, *166*, 112–122. https://doi.org/10.1016/j.applanim.2015.02.013

PubChem. (n.d.-a). *Cinnamaldehyde*. PubChem. https://pubchem.ncbi.nlm.nih.gov/compound/Cinnamaldehyde

PubChem. (n.d.-b). *Hexanoic acid*. PubChem. https://pubchem.ncbi.nlm.nih.gov/compound/Hexanoic-Acid

Sloan, K. B., Maughan, M. N., Sharpes, C. E., Greubel, R. R., Gallegos, S. F., Miklos, A. E., & Waldrop, L. D. (2025). The manipulation of odor availability of training aids used in detection canine training. *Frontiers in Allergy*, *5*, 1445570. https://doi.org/10.3389/falgy.2024.1445570

Stokes, T. F., & Baer, D. M. (1977). AN IMPLICIT TECHNOLOGY OF GENERALIZATION1. *Journal of Applied Behavior Analysis*, *10*(2), 349–367. https://doi.org/10.1901/jaba.1977.10-349

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

Waggoner, P., Lazarowski, L., Hutchings, B., Angle, C., & Porritt, F. (2022). Effects of learning an increasing number of odors on olfactory learning, memory and generalization in detection dogs. *Applied Animal Behaviour Science*, *247*, 105568. https://doi.org/10.1016/j.applanim.2022.105568

Appendix A

Standard Operating Procedure for Training Dogs to Use Olfactory Stimulator

Purpose

This document outlines the training procedure used to teach dogs to operate the olfactory stimulator and to establish reliable discrimination between target and non-target odours. The protocol describes the sequential training phases used to shape the required response behaviours and establish stimulus control prior to experimental testing.

Apparatus Setup

Position the apparatus in the lab room without additional objects which could lead to distraction. Only the front of the apparatus should be accessible to dogs. This can be achieved using the positioning of the device and barriers. Ensure the room doors close and latch securely, and the room contains suitable cameras for observation. Control systems and observation stations for the apparatus should be prepared in an adjacent room.

Training

Introduction and Habituation

The dog should be allowed to explore the room and facility and become habituated to the environment and researchers prior to commencement of training sessions. During early stages, sessions should be terminated at the first sign of fatigue or disinterest. Early sessions should not exceed 10 minutes.

Conditioned Reinforcer Establishment

Use the manual remote to dispense food from the automatic feeder until the dog immediately approaches upon activation. The dog should approach the feeder upon hearing the sound of activation and consume food within 3 seconds 3 times in a row before continuing training.

Shaping Nose in Sample Port

Once sound of the feeder has been established as a conditioned reinforcer, the remote is used to train the dog to put its nose into the sample port. Using differential reinforcement of successive approximations to this target behaviour. Prompting may be used but the prompt must be faded and removed before proceeding to next steps. Run air system throughout the shaping process to improve habituation.

As soon as the nose is inserted into the port for the first time, and the feeder triggered, take the dog out of the room and load cinnamaldehyde (target odour 1) into the apparatus. Set the target odourant to flow constantly and set to treat nose holds of 1,000 ms as a positive indication. All trials should be set as positive trials for shaping training. Continue shaping as appropriate until the dog begins to trigger the feeder automatically. Once the dog completes a session at the 1,000 ms indication threshold, increase the threshold to 1500ms. Once a session at 1500ms is complete, proceed to the next steps.

Shaping: Lever Activation

With the apparatus active and non-target 1 (benzyl-acetate) present, use the method of differential reinforcement of successive approximations to shape lever pressing. Depending on the specific dog, appropriate topography should be selected for shaping. Prompting may be used but must be faded and removed before proceeding to the next step. Once the lever has been activated 10 times without prompts (and reinforced through activation of the feeder), the next step can begin.

1. **Discrimination Training – Odour A Vs X**

To train attending to the odour port, begin with a discrimination procedure between two distinct odours. For this experiment using benzyl-acetate (non-target) and cinnamaldehyde (target odour

1). Start by loading cinnamaldehyde into odour channel 1 and benzyl-acetate into odour channel 2.

Ensure the odour rotameter is set to 0.2 L/min and the air flow carrier is set to 4 L/min. Both the odour rotameter and the air flow carrier remain unchanged. To test odours, sample manually by sniffing the port with either channel active. The single odorants when presented separately in trials should be just noticeable.

Start training by setting *Subjecttrialorder.csv* with 10 positive and 9 negative trials in alternating order. Positive trials should only have target odour channel active, and negative trials should only have non-target odour channel active. Set *subjectconfig.txt* with poss-niff = 1500 and minsniff = 500. Allow subject to sample the port. If no nose hold is performed for more than 2 positive trials, use prompting but fade as soon as possible. The same should be done for lever press with negative trials. Once pos sniff improves to 2000, immediately introduce randomization for the following trials with poss-niff 2000 and above and reduce the number of positive trials to 8 and 12 negative trials.

Once the subject is reliably performing both behaviours, record the number of correct indications and number of correct rejections. Increase sniff time incrementally by 500ms to around 5,000–6,000 ms, only after the dog consistently achieves a perfect hit rate at each threshold. Continue increasing until the correct rejection rate (CRR) stabilises above 80%. This should occur between 5000–6000ms, adjusted for depending on the individual dog.

Continue procedure until both correct rejection and correct indication >75% for 3 sessions. At this point move to the next part of the experiment.

2. **Discrimination Training – Odour B Vs Y**

The procedure for training the second odour (Target B) with hexanoic acid mirrors the process used to train the first target odour (cinnamaldehyde). Begin by loading hexanoic acid into odour channel 1 and linalool into odour channel 2. Ensure the odour rotameter is set to 0.2 L/min and the air flow carrier is set to 4 L/min. Both odour rotameter and carrier flow remain unchanged. To test odours, sample manually by sniffing the port with either channel active; the single odorants used in both positive and negative trials should be noticeable. Set up *Subjecttrialorder.csv* with 10 positive and 9 negative trials in randomized order, where positive trials only activate target odour channels and negative trials only activate non-target odour channels. In *subjectconfig.txt,* set poss-niff = 1500 and minsniff = 500. Allow the dog to sample the port. If the dog fails to perform a nose hold on more than 2 consecutive positive trials, use prompting and fade as soon as possible; the same applies to lever pressing on negative trials. Once poss-niff improves to 2000 ms, immediately adjust the ratio to 8 positives and 12 negatives. When the dog is working at poss-niff values previously established, reduce to 7 positives and 13 negatives. Once the dog reliably performs both behaviours, record correct indications and correct rejections, adjusting poss-niff thresholds if required, and keep minsniff at 500 should remain unchanged. Continue until performance exceeds 75% on both measures for 3 consecutive sessions, then proceed to the next part.

3. **Mixed Target Discrimination with Multiple Non-Targets (A + B vs X + Y)**

Following successful discrimination training with individual target odours, the next phase involves presenting both cinnamaldehyde and hexanoic acid (A & B) as positive odours, and benzyl acetate and linalool (X and Y) as non-target odours. The *SubjectTrialOrder.csv* file specifies the full sequence of 20 trials per session, incorporating all four odours (A, B, X, Y). Target odours (A & B) are assigned to target odour channels, and non-target odours (X & Y) to

non-target channels. Each session includes 8 target trials (4 A, 4 B) and 12 non-target trials (6 X, 6 Y), with trial order randomised across the session. Maintain poss-niff between 5000–6000 ms and minsniff at 500ms. Prompting may be used if the dog fails to respond correctly on more than two consecutive positive or negative trials but should be faded as soon as reliable performance is observed. Continue this phase until both the hit rate and correct rejection rate reach at least 75% for three consecutive sessions, calculated separately for each target odour (A & B) and each non-target odour (X & Y). Once a consistent performance threshold is reached and possniff duration improves to previously established threshold, the ratio of positive to negative trials may be adjusted to 7:13 to reduce target frequency.

4. **Novel Non-Target Introduction for Generalization Control**

Once the dog has successfully discriminated between target odours A and B and non-target odours X and Y, the next step involves introducing novel non-target odours to confirm that the dog is responding specifically to A and B, rather than excluding only previously seen distractors. Begin by introducing the non-target odour combination Y + Z (where Z is a novel non-target odour not previously encountered by the dog). The dog must correctly reject Y + Z in at least 75% of trials over two consecutive sessions before progressing to the next combination. Subsequent sessions will introduce additional novel non-target combinations in sequence (e.g., Z + W, then W + V), with each new odour (W, V) being unfamiliar to the dog until its introduction. This phase continues until the dog demonstrates immediate and consistent rejection (≥75% correct rejection in the first session of exposure) of a novel non-target combination, indicating generalisation of non-target discrimination. Non-target odours should be loaded into odour channels not currently used for A and B, with odour delivery randomized automatically by the software. Maintain poss-niff values previously established and minsniff stays at 500ms. Once

the dog correctly rejects a novel non-target odour on its first encounter within a combination at a rate of 75%, continue this phase until the correct rejection rate reaches 75% across two consecutive sessions, demonstrating that the dog can consistently reject unfamiliar non-target odours while maintaining at least a 75% hit rate for A and B.

5. Intermittent Reinforcement

During the initial intermittent phase, reinforcement will occur on approximately 75% of correct responses, corresponding to a variable-ratio schedule of VR1.33. Each session will consist of 20 trials: 8 target trials and 12 non-target trials. Of the target trials, 6 will be reinforced and 2 will be unreinforced, with reinforced trials randomly distributed within each session. A stable hit rate ≥75% and correct rejection rate ≥75% across two consecutive sessions will be required before progressing to the next phase. In the second phase (VR2; 50% reinforcement), each session will again include 20 trials (12 negative, 8 positive), with 4 of the 8 target trials reinforced and 4 unreinforced. Trial order and reinforcement positions will be randomized for each session. Once stable performance is observed (≥75% hit rate and ≥75% correct rejection rate across two consecutive sessions), probe trials will be introduced to assess generalisation to higher concentrations of the target odours.

6. Probe Trials for Generalisation Across Concentration

The final phase will involve non-reinforced probe trials to assess whether the dog generalises its target responses across different concentrations of trained odours. These probe trials will present either cinnamaldehyde or hexanoic acid (A+ or B+) at a 20x concentration. Before introducing probe trials, a PID test will be conducted to confirm that the olfactometer is delivering the correct concentrations as required. Only after validating the equipment's accuracy will probe

trials be introduced. During this phase, probe trials will not be reinforced to avoid a testing effect.

Each probe session will comprise 20 randomized trials in total. Of these, eight will be target trials (A, B), ten will be non-target trials (X, Y, Z), and two will be probe trials presenting the target odours (A+, B+) at twenty times the trained concentration. Six of the target trials (six out of eight; 75%) will be reinforced on a VR1.33 schedule, while the remaining two target trials will be unreinforced. Both probe and non-target trials will remain unreinforced.

7. Concentration-Training Procedure for Odour B

If probe trials show no generalisation to the higher concentration of Odours A+ & B+, a concentration-training phase will be conducted. Odour B will be presented at the originally trained concentration (B) and at the high concentration (B+) within the same session. Each session will contain 4 B trials at the normal concentration, 4 B+ trials at the 20× concentration, and 12 non-target trials, for a total of 20 randomized trials.

For the first two sessions, B+ indications will be accepted at 1000 ms instead of the full possniff threshold, and all correct B+ responses will be reinforced during the first session before shifting to VR75% reinforcement in later sessions. Normal B trials will retain the possniff threshold the dog was already performing at (e.g., 5000–6000 ms), and reinforcement for B trials will remain at VR75%.

After the first two sessions, the possniff threshold for B+ trials will be gradually increased in stepwise increments across subsequent sessions until it reaches the same threshold as normal B trials. Training will continue until the dog achieves at least a 75% hit rate across both B and B+ trials and at least a 75% correct-rejection rate for non-targets across two consecutive sessions.

Once these criteria are met and performance is stable, the concentration-training phase will be discontinued, and the protocol will move on to testing probes of A+ & B+ again.

Appendix B

Guidelines for Shaping

1. Procedure

The researcher should position themselves near the apparatus, ideally near the door, while avoiding the dog's gaze to reduce unintentional cueing.

2. Shaping of Sample Port Entry

Turn the apparatus on and start a session so that sample port beam breaks produce a "beep."

• Reinforce moving progressively further away from the feeder.

• Reinforce attending to the apparatus.

• Reinforce placing the nose near the port.

• Reinforce placing the nose in the port.

• Reinforce breaking the beam.

3. Shaping of Lever Press

Turn the apparatus off.

• Reinforce any movement toward the lever.

• Reinforce movement of the nose or paw toward the lever.

• Reinforce any contact with the lever.

• Reinforce any movement of the lever.

• Reinforce movement of the lever that produces the "click."

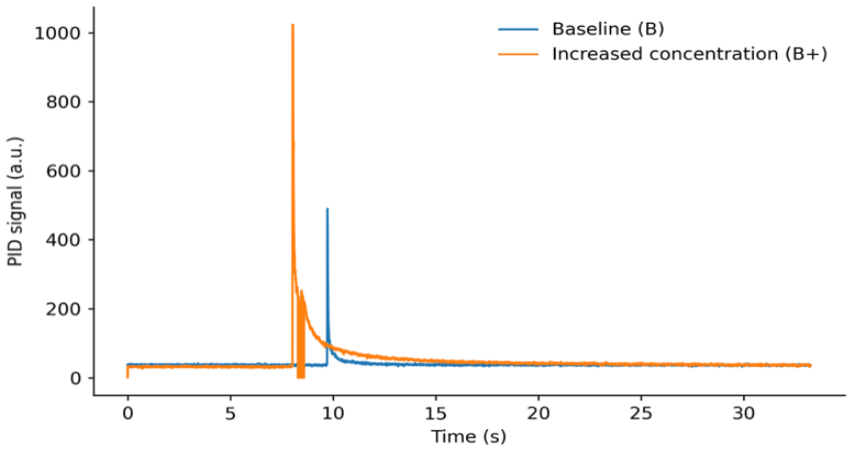
Appendix C

Validation of Vapour Concentration Using Photoionisation Detector

The following figures present raw photoionisation detector (PID) output recorded during validation of vapour concentration for the target odours used in the present study. Signals are displayed as continuous traces sampled at 1 ms resolution. Output values are expressed in arbitrary units (a.u.) as provided by the miniPID sensor. These data are presented to document relative differences in vapour concentration between baseline and increased concentration conditions.

Figure 19

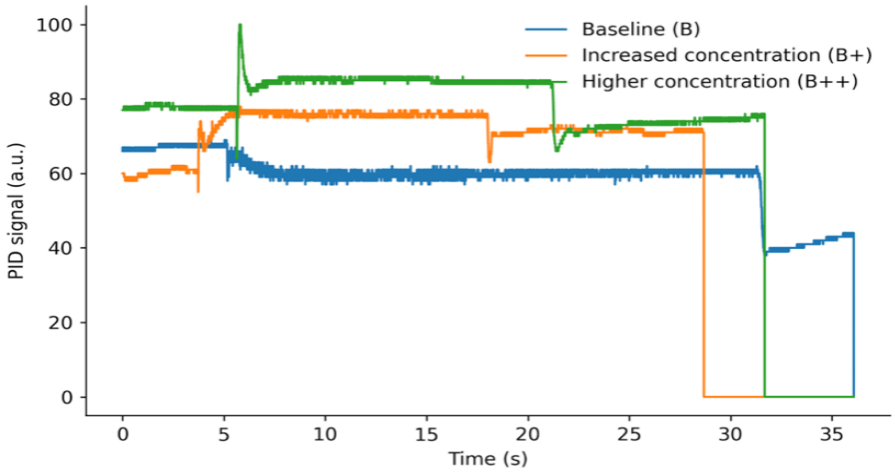
PID Response to Hexanoic Acid at Baseline Concentration (B) and Increased Concentration (B+) During Initial Validation Test



Note. Continuous traces sampled at 1 ms resolution. Output values represent raw PID signal (a.u.). The rise in signal corresponds to the onset of odour delivery, and the decline toward baseline reflects cessation of odour flow and clearing of the sampling line.

Figure 20

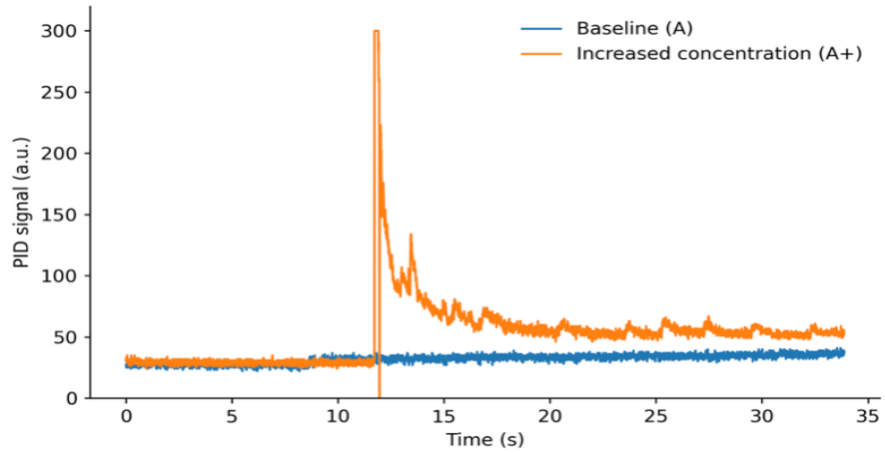
PID Response to Hexanoic Acid at Baseline Concentration (B), Increased Concentration (B+), and Higher Concentration (B++) During Later Validation Test



Note. Continuous traces sampled at 1 ms resolution. Output values represent raw PID signal (a.u.). The rise in signal corresponds to the onset of odour delivery, and the decline toward baseline reflects cessation of odour flow and clearing of the sampling line.

Figure 21

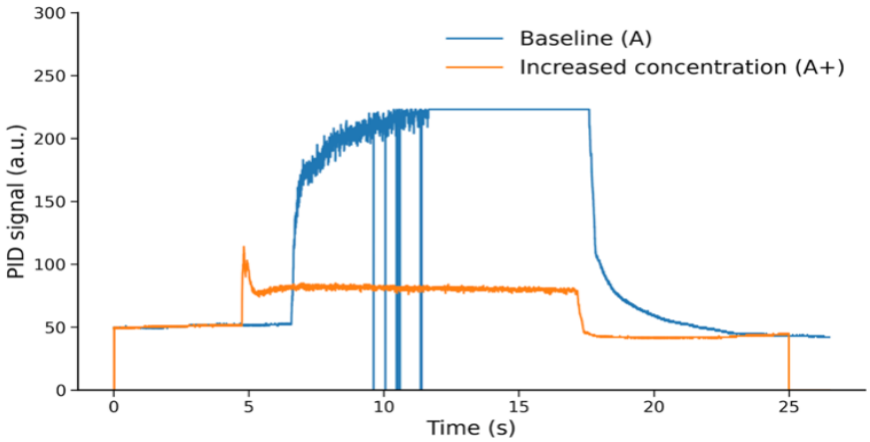
PID Response to Cinnamaldehyde at Baseline Concentration (A) and Increased Concentration (A+) During Initial Validation Test



Note. Continuous traces sampled at 1 ms resolution. Output values represent raw PID signal (a.u.). The rise in signal corresponds to the onset of odour delivery, and the decline toward baseline reflects cessation of odour flow and clearing of the sampling line.

Figure 22

PID Response to Cinnamaldehyde at Baseline Concentration (A) and Increased Concentration (A+) During Later Validation Test



Note. Continuous traces sampled at 1 ms resolution. Output values represent raw PID signal (a.u.). The rise in signal corresponds to the onset of odour delivery, and the decline toward baseline reflects cessation of odour flow and clearing of the sampling line.