Teaching old dogs and young dogs new tricks: canine scent detection for seabird monitoring

Mark Bolton[1*], Greg Morgan[2], Susannah E. Bolton[1], Jack R. F. Bolton[1], Sarah Parmor[2] and Laura Bambini[3]

* *Correspondence author. Email: mark.bolton@rspb.org.uk*

[1] *RSPB Centre for Conservation Science, The Lodge, Sandy, Bedfordshire, SG19 2DL, UK;*

[2] *RSPB Ramsey Island, St Davids, Pembrokeshire, SA62 6PY, UK;*

[3] *Global Species Recovery, RSPB UK Headquarters, The Lodge, Sandy, Bedfordshire, SG19 2DL, UK.*

Abstract

Dogs *Canis familiaris* have been domesticated for over 11,000 years and have been trained to perform a vast array of tasks. Scent dogs are routinely used to detect elusive animals of conservation concern as well as the presence of invasive non-native predators. However, a recent review of studies on scent dogs for nature conservation found few peer-reviewed papers on detection of seabirds, and we suggest the potential of scent dogs for seabird monitoring is currently under-utilised. To stimulate wider consideration of the use of scent dogs for seabird monitoring we: (i) document the training, testing and performance of a one-year old Golden Retriever, to detect scent from European Storm Petrels *Hydrobates pelagicus*, and to differentiate between scent of European Storm Petrels and Manx Shearwaters *Puffinus puffinus*, and (ii) report on the field performance of a 12-year old Border Collie to detect Manx Shearwaters breeding in natural nest burrows and artificial nest boxes. We show that these individual dogs had a high success rate in locating hidden seabird targets by scent in experimental and field settings and we highlight some of the set-backs encountered during the training process, and their solutions. We show that the detection of occupied Manx Shearwater burrows is dependent on their depth and recent history of occupation. Our results highlight the potential value of scent dogs to establish the presence of particular seabird species at sites where breeding status is currently uncertain, and to map the extent of occupied areas. Further work is needed to validate the use of scent dogs to determine seabird breeding density in real-world situations, where detection probability may be influenced by nest site characteristics and breeding density. We stress the need to involve professionals in the selection, training and testing of scent dogs.

Introduction

Dogs *Canis familiaris* have a long history of domestication, stretching back over 11,000 years (Bergström *et al*. 2020), and have been trained to perform a vast array of tasks, including protection and herding of livestock, search and rescue, assistance of the visually impaired and scent detection of a range of targets including narcotics, explosives, disease, plants and animals (Bird 1996; Gazit & Terkel 2003; Browne *et al*. 2006; Jezierski *et al*. 2016; Gerritsen & Haak 2017; Bennett *et al*. 2020). Canine olfactory neurophysiology is well-understood (Uemura 2015) and widely recognised to be immensely superior to human olfactory ability. In New Zealand and North America, scent dogs have been used for conservation purposes for many decades, to locate species of conservation concern, and their non-native predators (Dahlgren *et al*. 2012). The New Zealand Government Department of Conservation has operated a Conservation Dog Programme for many years (www.doc.govt.nz/our-work/conservation-dog-programme, accessed 27/12/2020), to promote, manage and set the standards for the use of scent detection dogs for conservation purposes. Within Europe, the use of scent dogs is much less widespread, despite their potential for surveys of species such as Capercaillie *Tetrao urogallus* being recognised many years ago (e.g. Gilbert *et al*. 1998). A recent review (Grimm-Seyfarth *et al*. 2021) of over 2,400 published cases of the use of scent dogs for wildlife detection purposes found 619 employed dogs to detect avian targets, but only nine referred to the detection of seabirds. The paucity of peer-reviewed studies on canine detection of seabirds is surprising, since canine detection has been shown to be effective for species that breed in cavities and are only active above ground nocturnally (e.g. kiwis *Apteryx* sp. and Kakapo *Strigops habroptilus*). Many seabird species exhibit these same characteristics and some, such as the storm petrels, are well-known for their strong odour, which suggests that scent detection may provide an effective method for establishing the presence of nesting birds at potential breeding locations, and possibly to quantify breeding density.

The location and extent of breeding colonies of cavity-dwelling, nocturnal seabirds are often poorly known, leading to uncertainty regarding the presence of particular species at a given site. Even where the occurrence of breeding birds has been established, delimiting the extent of the breeding colony/sub-colonies (which potentially may cover vast areas in difficult terrain) is problematical, leading to large extrapolation errors in estimating population size from density of sample plots. The use of scent dogs to (i) detect the presence of a particular species at a site, and (ii) to indicate the extent of the area(s) occupied by the target species, offers considerable potential that is currently under-utilised in a European context, likely due to lack of experience of scent dog capabilities by those responsible for seabird monitoring.

Our primary aim is to highlight to the community of seabird researchers, conservation managers and statutory agencies the potential of scent dogs for seabird monitoring and to encourage further research into the efficacy and efficiency of scent dogs compared to existing methods. Here we (i) document the training, testing and performance of a one-year old Golden Retriever to locate targets treated with scent of European Storm Petrels *Hydrobates pelagicus (*hereafter Storm Petrel) in a variety of experimental settings, and (ii) quantify the reliability of a self-trained 12-year old Border Collie sheepdog to detect occupied Manx Shearwater *Puffinus puffinus* (hereafter Shearwater) nests in natural and artificial nest sites. We examine the dog's assessment of Shearwater burrow occupancy in relation to burrow depth and the number of Shearwater occupants and the local occurrence of European Rabbits *Oryctolagus cuniculus* (hereafter 'Rabbits'), which may all influence scent detection. We document problems, set-backs and solutions, and quantify the relative time effort required for canine scent detection compared to conventional survey methods, acknowledging that many of these issues would not have arisen had we employed fully-trained dogs and professional handlers. We hope that our study will encourage others to consider the use of scent dogs for seabird monitoring, under professional guidance and to conduct appropriate assessment of precision and sensitivity of scent dog performance.

Methods

Training the scent dogs

Dog A was a one-year old Golden Retriever, from a working pedigree, selected from the litter due to her high drive, in order to maximise suitability for scent work (see Beebe *et al*. 2016). Professional one-to-one training in scent detection was provided at the outset and the dog was trained by handlers (authors MB, SB) with no previous experience, to detect and indicate scent from the feathers of Storm Petrels using well-established methods (Hewings 2019; Centre for Protection of National Infrastructure 2020; Lazarowski *et al*. 2020). The dog was trained using a clicker device to mark correct behaviour. Scent targets were prepared by placing feathers (from ≥ 8 full-grown individuals of unknown sex) of either Storm Petrels or Shearwaters in an airtight glass jar with cotton pads for at least 24 hours. Control samples (not containing seabird scent) were prepared in exactly the same manner, but omitting feathers. Care was taken to avoid cross-contamination of scent between jars and sample tubes. Use of cotton pads allowed the preparation of a large number of identical scent targets, which were disposed of after each session to avoid the risk of cross-contamination associated with prolonged use. However, cotton pads may not absorb the entire scent profile of Storm Petrel feathers, and so may not fully reflect the whole scent spectrum of the feathers themselves (Kotthoff & Nörenberg 2016). For training or testing sessions, a cotton pad was placed in an uncapped plastic sample tube (90 x 25 mm). The target scent of Storm Petrel was first classically conditioned with a food stimulus by presenting the scent alongside a food treat (Hewings 2019). After a week of such training, search behaviour was trained by hiding a sample tube containing the target scent either in one of a number of identical plastic pipes in a choice test, or outdoors (e.g. a crevice in a wall, under a bush, in long vegetation etc). When the target was located, the dog's response was to indicate the location with its front paw, which was marked by a click (to indicate correct behaviour) and a food reward, to condition the indication of the target. Training and testing took place on an almost daily basis, for up to 30 minutes each day, over a four-month period.

During the training process to establish that the dog was cueing on the Storm Petrel scent, rather than the scent of the sample tube, or the scent of the individual who had concealed the tube, a variety of choice tests were performed, to discriminate between Shearwater and Storm Petrel scent (to ignore the former and indicate the latter), and latex gloves were worn by the individual placing the sample tubes. In order to maintain the dog's interest and motivation during the training and testing processes, a wide variety of equipment and locations were used to train, test and improve identification and discrimination of target from non-target scent (see Hewings 2019 for examples).

Dog B was a 12-year old Border Collie and working sheepdog, trained and handled by author GM, and used on a daily basis to assist with livestock management on the RSPB nature reserve of Ramsey Island, Pembrokeshire, UK (51°33'N 5°20'W). The dog has accompanied GM for many years on playback surveys to detect breeding Shearwaters that nest in burrows on Ramsey Island (Perkins *et al.* 2017). Playback surveys involve playing a recording of a Shearwater call at the entrance to potential nesting burrows to elicit a vocal response from birds within. Typically, responses will be elicited from around 50–65% of occupied burrows. Whilst accompanying GM on such surveys over several years, the dog learned to discriminate between occupied and unoccupied burrows, and developed behaviour to search for burrows with Shearwater scent, lying down at the next burrow it detected to be occupied, ignoring those that were not. In this manner it was 'self-taught' to identify and indicate burrows occupied by Shearwaters.

Experimental protocols

1. Reliability of Storm Petrel detection

When the training indicated that dog A had developed an ability to detect and indicate the required target, a choice experiment was conducted to measure the dog's precision (i.e. the proportion of alerts that were correct, Bennett *et al.* 2020) using a low test wall, similar to collapsed drystone walls in which Storm Petrels nest (Bolton *et al.* 2017). The test wall was of loose brick construction, and contained 20 (Experiments 1, 2a) or 10 (Experiment 2b) similarly sized numbered cavities approximately 14 cm apart. Scented cotton pads were placed in uncapped sample tubes that were concealed in cavities selected at random. Scent samples were discarded at the end of each session, so training and testing used different scent samples. The sample tubes were located at the far end of the cavities and were not visible to the dog. The dog was led along the wall, on a lead, testing each cavity in turn. The handler, who was blind to the location of the target, reported the cavity indicated to contain the scent. If correct, the dog's indication was marked with a click, and rewarded with food and play. An incorrect indication was recorded as incorrect, and ignored by the handler. Each daily session involved up to 16 successive trials, with scent tubes being repositioned between successive trials. Up to nine successive sessions were conducted (usually daily) for each experiment. Experiment 1 tested the ability of the scent dog to correctly indicate the location of a Storm Petrel target placed at random in one of the 20 cavities. Experiment 2 tested the dog's ability to locate a Storm Petrel target and differentiate from a Shearwater target placed at random elsewhere in the test wall. Although efforts were made to avoid cross-contamination of scent between sample tubes, poor discrimination during the initial trials Experiment 2 (see results), led to concerns over the possibility of cross contamination and hence confusion for the dog. As a result, scent training was recommenced from the start, to re-establish the conditioning with Storm Petrel scent, and discrimination from blank and Shearwater scent. Once the training sessions indicated that correct discrimination of Storm Petrel scent had been re-established, Experiment 2 was recommenced, implementing strict protocols to avoid cross-contamination of scent between treatments.

The experiments described above allow assessment of the proportion of the dog's indications that correctly indicate a target. However, because the dog had been trained, and was tested, to detect a single target (and the trial concluded once a single indication had been given, even if the correct target had not yet been encountered), these experiments do not provide an effective assessment of sensitivity, that is the proportion of available targets that are detected. To assess sensitivity, we performed a further experiment, in which a variable number of Storm Petrel scented targets were placed at random in 40 cavities located in a 25 m stretch of drystone wall which resembled Storm Petrel breeding habitat. In order to represent breeding densities as closely as possible, the number of targets chosen for each trial was selected at random from the distribution of the number of nests within 25 m sections of drystone wall in a survey of a large Storm Petrel colony in Scotland (Bolton *et al.* 2017). The number of targets per 25 m ranged from five to 26. The locations of the Storm Petrel scented targets were selected at random from the 40 possible cavities. Unscented control samples were placed in the remaining cavities. The dog was led on a lead along the wall and indicated the location of Storm Petrel targets. Correct indications were marked with a 'click' and rewarded with food. Incorrect indications were noted, but not marked or rewarded. The time required to complete each trial was noted. Up to four trials were repeated daily.

2. Field trials to detect occupied Shearwater burrows on Ramsey

In order to assess the reliability of detection of burrows occupied by breeding Shearwaters, dog B was tested repeatedly on a set of 83 study burrows (nest boxes) where the occupation status was known from regular visual inspection of the box contents. The dog assessed the contents of each nest box on four separate occasions between 30 May and 20 June 2020 (the peak incubation period of Shearwaters on Ramsey), and after it had indicated whether the box was occupied or not, the contents were inspected, and recorded as 'unoccupied, non-breeding' (no birds or egg present), 'occupied, non-breeding' (if one or more adults were present, but no egg), 'occupied, breeding' (if at least one adult was present with an egg), or 'unoccupied, breeding' if no adult was currently present, but the box contained an egg, indicating the box was being used for breeding in the current year, but both adults were temporarily absent, or the nest had been abandoned. Since the dog might have detected scent lingering in a nest cavity that was currently unoccupied, but which had been occupied recently, we also recorded the occurrence of occupation within the previous seven days, from regular nest box inspections. Since the dog had not been actively trained to detect Shearwater nests, it was not rewarded for correctly indicating occupied sites.

Due to the modest anticipated sample size of nest boxes used for breeding, we also tested dog B at natural burrows where occupancy status was assessed using playback. Earlier work on Ramsey indicated that on average, responses are obtained from around 57% of occupied nests on any given occasion (Perkins *et al*. 2017). We aimed for a sample of c. 50 occupied and c. 50 unoccupied sites, so on 27 May we conducted playback in an area known to be occupied by breeding Shearwaters until 40 responses had been obtained, and a further 60 sites had elicited no response. The dog assessed all the sites prior to conducting the playback and indicated whether each was occupied. Each burrow was individually marked and on three subsequent occasions, on separate days between 31 May and 23 June, the dog indicated if the site was occupied, and then each was subjected to further playback, and any vocal response, and the number (one or two) and sex of any responding birds (determined from the pitch of the calls, Perkins *et al*. 2017) was noted. Since Shearwaters nesting in deep burrows may be more difficult to detect by scent compared with those nesting closer to the surface, where responses to playback occurred, we scored the depth of the burrow as deep or shallow, according to the perceived volume of the vocal response. We also recorded the presence/absence of Rabbit droppings at the burrow entrance, since Rabbit odour could potentially mask the detection of Shearwater scent, or the dog may simply be indicating the presence of a burrow-dwelling animal that represents potential prey.

Since we conducted playback and scent dog detection concurrently, we were unable to directly measure the time required to assess all sites by the scent dog alone. Instead, we estimated the unique components of each. Firstly, we measured the total amount of time required to complete the playback inspection of all 100 locations on the last three trials (which did not include marking burrows). Then, we calculated the travel time required to move among all sites as the total time to conduct playback at all sites, minus the time required to complete each playback and note any response at the set of 100 natural sites (55 seconds/site x 100 sites). Finally, we estimated the total amount of time required to survey the 100 natural sites by a scent dog as the sum of travel time + 500 seconds, assuming the time taken for the dog to indicate, and the outcome to be noted, was five seconds per site. To compare search effort of scent dog detection with that of conventional playback, we assumed that a single fieldworker could undertake both the handling and data recording associated with each dog.

Data analysis

Following Bennett *et al*. 2020 we calculated 'precision' (the proportion of indicated targets that were correct), 'sensitivity' (the proportion of available targets that were detected) and 'accuracy' (the proportion of all cases that were correctly identified as targets or non-targets). We used mixed models to assess any improvement in detecting Storm Petrel targets (Experiment 4) across trials. Sample type (Storm Petrel scent or control) and trial number were fitted as fixed effects and cavity identity was fitted as a random term. Models were fitted using package lme4 (Bates *et al*. 2015) in R v3.6.1 (R Core Team 2019). Model comparisons were conducted on models fitted by Maximum Likelihood estimation, and top-performing models were refitted using adaptive Gauss-Hermite quadrature to estimate parameter values, where the number of quadrature points was determined as the smallest value at which log likelihood stabilised (Lesaffre & Spiessens 2001). For Shearwater nest boxes, true occupancy status was known with certainty from visual inspection of the chamber contents. However, assessment of occupancy of natural sites (where visual inspection is not always possible, or reliable, Perkins *et al*. 2017), relied on results of playback trials. Where vocal responses were obtained, a nest could be confidently classified as 'occupied', and this subset was used to assess sensitivity, which we calculated separately for shallow and deep occupied burrows. However, burrows where no response was elicited could not be judged as unoccupied with certainty, due to non-response of occupants on occasion. For nest boxes we examined whether the dog's indication of occupancy (modelled as a binomial response with a logit link) was related more closely to recent occupancy than current occupancy by comparing AICc values of separate models including each of these terms. We were unable to examine the effect of Rabbit presence on the dog's indication since Rabbit droppings were absent from all nest box entrances. Nest box identity was included as a random term to account for repeated measures from each nest box. For the subset of natural sites where occupancy was confirmed by playback, the dog's indication of occupancy was modelled as a binomial response with a logit link in relation to burrow depth, Rabbit presence and the number and gender of the responding nest occupant(s). Burrow identity was included as a random term.

Ethical considerations

This study was carried out in full compliance with the UK Animal Welfare Act 2006, and the guidance notes on working dog welfare published by the UK Centre for Protection of National Infrastructure (2020). The dogs were never subjected to any form of punishment for incorrect behaviour; rather, training relied entirely on positive reinforcement of correct behaviour. All activities complied with the RSPB Code of Practice 14 'The Use of Detection dogs on Islands', and were fully Covid compliant (being conducted whilst Covid-19 restrictions applied to UK). Feathers for preparing scent targets were obtained from Storm Petrels and Shearwaters killed by predators.

Results

Storm Petrel detection

Experiment 1. Ability to locate a Storm Petrel target

After an initial period of two weeks' scent training with a Storm Petrel scent target, dog A's precision (proportion of indications that were correct) was 67% and increased to 100% over the course of 11 subsequent daily trials (Figure 1).

Experiment 2. Ability to differentiate Storm Petrel and Shearwater scent

Dog B's ability to differentiate between Storm Petrel and Shearwater scent increased from 50% to 94% over the course of the first four sessions, but thereafter was highly variable, with performance on half the subsequent six trials of < 63%

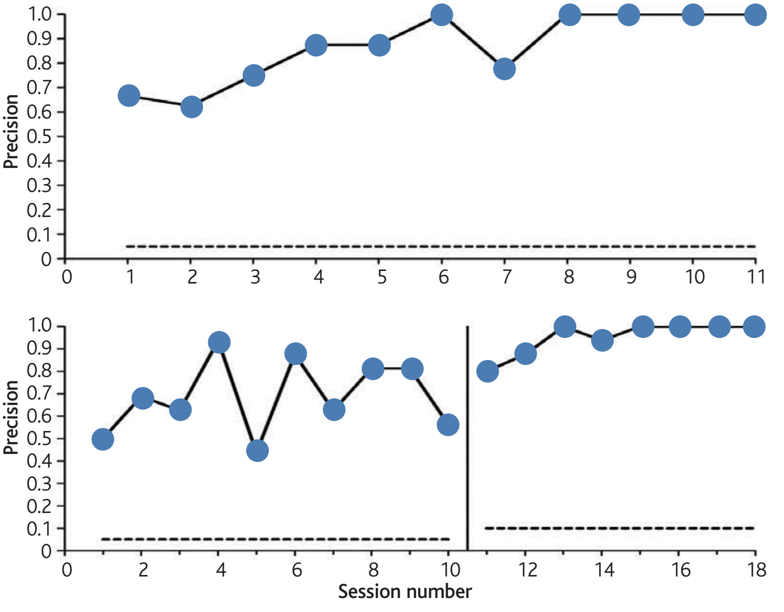

Figure 1. Ability of a dog *Canis familiaris* to locate a European Storm Petrel *Hydrobates pelagicus* scent target hidden in a cavity in a test wall, in the absence of any other hidden scent sample (Experiment 1, upper panel) and when a Manx Shearwater *Puffinus puffinus* scent target was also hidden elsewhere in the wall (Experiment 2, lower panel). Experiment 2 was paused after session 10 for a period of retraining, indicated by the vertical line. The dotted lines indicate the likelihood of detecting the correct target at random, higher from trial 11 of Experiment 2 due to the smaller number of cavities.

(Figure 1). All unsuccessful detections related to identification of the cavity holding the Shearwater scent. Reasons for indication of non-target scent sample were not known, but may have been caused by cross-contamination between scent samples, or cueing on the scent of the plastic bottles, or latex gloves used to handle samples rather than the Storm Petrel target, leading to confusion for the dog. Following a further period of training, re-establishing the conditioning of Storm Petrel scent, rewarding indication of Storm Petrel targets and ignoring indication of Shearwater targets, ability to differentiate Storm Petrel targets was restored. Storm Petrel targets were reliably located on 80% of trials when testing resumed (session 11), and all incorrect indications related to identification of the Shearwater targets, rather than empty cavities. Performance improved further to 100% thereafter, such that all Storm Petrel targets were successfully located on all of the four final sessions (each involving 16 trials, i.e. 64 consecutive correct indications).

Experiment 3. Ability to locate multiple Storm Petrel targets

When presented with a larger and variable number of Storm Petrel scent targets hidden in 40 cavities in a 25 m stretch of drystone wall, precision was initially much lower than the 100% achieved in the previous single-target experiment, but significantly improved across the 23 trials, conducted over six days (Figure 2). At the start of the experiment, the probability of a target being detected (sensitivity) was around 50%, rising to 95% at the end of the experiment (Figure 2). After four days of training (from trial 12), there was a strong positive correlation between the total number of alerts by the dog (including false positives), and the number of targets available for detection on each trial (Figure 2, $R^2 = 0.914$, $P < 0.0001$).

The mean time required to compete each trial was 03:28 minutes (range 01:42–06:25 minutes), and varied according to the number of targets located (since correct indications were rewarded, requiring a break from searching of a few seconds). Conventional playback typically uses a sound recording of 38 seconds (which includes time to note the response, Bolton *et al*. 2017), which would require approximately 16 minutes to survey 25 m of wall.

Scent detection of Shearwaters

1. Nest boxes

All four trials indicated 100% sensitivity: all nest boxes that were occupied by birds at the time of the trial were correctly indicated by dog B (Table 1). Of the 14 nest boxes in which breeding activity occurred (i.e. an egg was laid), four were not occupied by birds in all trials, and at one of these nests an unattended egg that was present at the first trial, was subsequently ejected. The box was unoccupied on the second trial and occupied by non-breeding birds on the last two trials. Among the nest boxes that were not used for breeding and not occupied by non-breeders at the time of the trials, correct indications of non-occupancy averaged 92.5% (i.e. a false positive rate of 7.5%). Nests that were unoccupied at the time of the trial were more likely to be indicated as occupied if they had been occupied in the previous week or had been used for breeding (i.e. contained an unattended egg): 90% of the false positive indications related to instances where nest boxes had been recently occupied by non-breeding birds. If all boxes that were currently occupied (by breeding or non-breeding birds), or which had been occupied in the previous seven days were considered as 'active sites', across the four trials precision (proportion of indications that were correct) averaged 97.5% (range 94.7–100.0%), sensitivity (proportion of occupied sites detected) averaged 86.5% (range 85.7–87.5%) and accuracy (proportion of all sites where occupancy was correctly assessed) averaged 95.8% (range 85.2–96.4%, Table 2). Note that boxes classified as 'unoccupied non-breeding' in Table 1 are included as

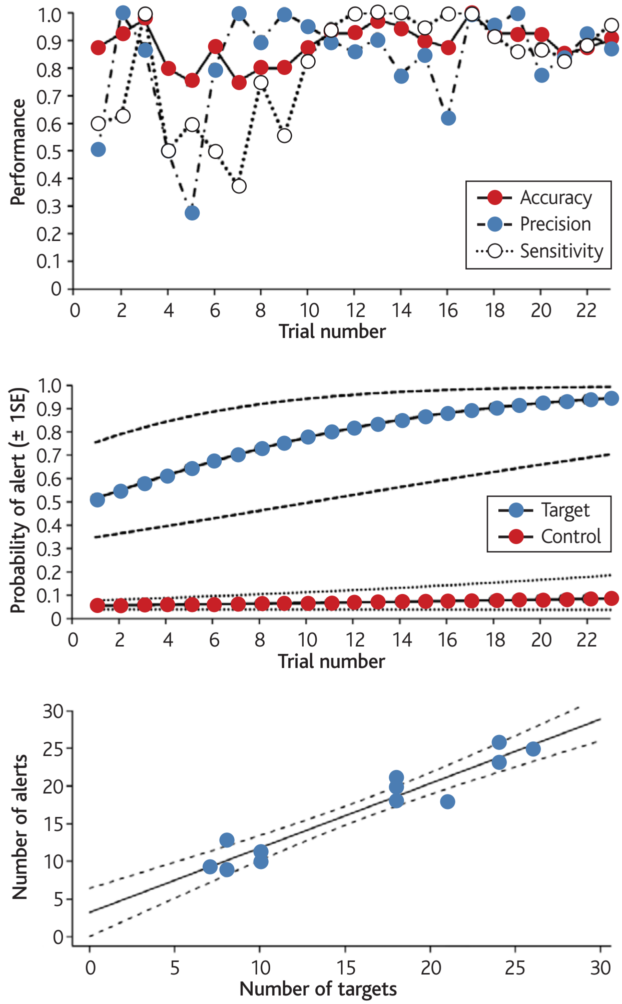

Figure 2. Precision, sensitivity and accuracy of detection of European Storm Petrel *Hydrobates pelagicus* scented targets concealed within a stone wall (upper panel); modelled improvement in discrimination of targets and controls (middle panel); relationship between the number of alerts and the number of true targets, from day 4 (trial 12) onwards (lower panel, alerts = 3.232 (S.E. 1.439) + 0.855 (S.E. 0.083) x targets).

Table 1. Confusion matrix of indication of occupancy status of Manx Shearwater *Puffinus puffinus* nest boxes by a scent dog *Canis familiaris*.

Trial #	Date	Dog indication	Nest box occupancy status			
			Occupied, breeding	Unoccupied, breeding	Occupied, non-breeding	Unoccupied, non-breeding
1	30 May	Unoccupied	0	0	0	65
		Occupied	11	2	1	4
2	5 June	Unoccupied	0	1	0	63
		Occupied	11	1	3	4
3	12 June	Unoccupied	0	2	0	59
		Occupied	10	1	3	8
4	20 June	Unoccupied	0	3	0	60
		Occupied	10	0	6	4

Table 2. Confusion matrix of scent dog *Canis familiaris* indication of occupancy and recent (within the previous seven days) usage of Manx Shearwater *Puffinus puffinus* nest boxes.

Trial #	Date	Dog indication	Nest box recent usage status	
			Active	Inactive
1	30 May	Unoccupied	3	62
		Occupied	18	0
2	5 June	Unoccupied	3	61
		Occupied	18	1
3	12 June	Unoccupied	3	58
		Occupied	21	1
4	20 June	Unoccupied	3	60
		Occupied	20	0

'active' in Table 2, if occupied during the seven days prior to a trial. The response of the dog was more closely related to the box occupation over the previous seven days than the current occupancy (Table S1).

2. Natural sites

Among the subset of burrows where subsequent playback revealed they were currently occupied, sensitivity averaged 62.6% (range 0.50–0.70) over the four trials. Dog B was more than twice as likely to correctly indicate a burrow was occupied where burrows were shallow (58% of measured burrows), compared with deep burrows (Figure 3). Hence, the probability of a positive indication was strongly related to burrow depth (included in all three top models, Table S2). The dog was more likely to indicate a burrow was occupied if only one bird subsequently responded, compared with nests where two birds replied to playback (Table S1, Figure 3), and models including the number of responding birds performed considerably better (> 2.0 delta AICc) than those including the sex of the responding occupants. There was also weak evidence that, in addition, the presence of Rabbits reduced the probability of detection (second-ranked model with delta AICc of 1.99). Considering the subset of 41 burrows where vocal responses to playback were never detected (which were most likely to be genuinely unoccupied), the proportion identified as unoccupied averaged 97.6% (range 92.7–100%), indicating a maximum false positive rate of 0–7.3%.

The mean duration of playback surveys of all 100 natural nest sites was 1:39:40 hours (travel time = 8 minutes, total playback duration = 1:31:40 hours). We estimate the time required to survey all sites using the scent dog alone as 16:20 minutes (travel time = 8 minutes + 5 seconds x 100 sites), representing a reduction in effort of 84% for this set of sites.

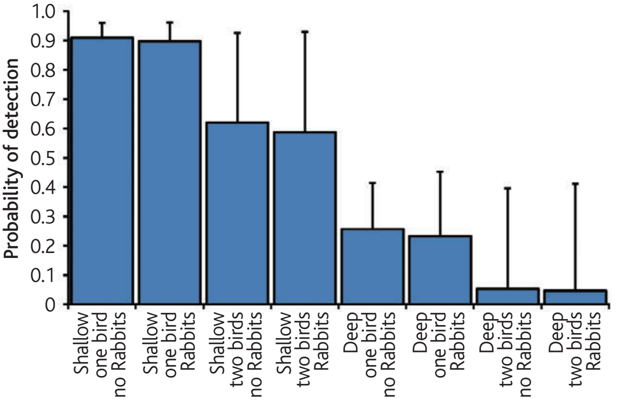

Figure 3. Probability of scent dog *Canis familiaris* indicating an occupied natural nest site (where playback subsequently indicated birds present), in relation to burrow depth (shallow or deep), the number of responding birds (1 or 2) and the presence of European Rabbits *Oryctolagus cuniculus* indicated by droppings at the burrow entrance. Estimates are derived from weighted average parameters of the top three models in Table S1.

Discussion

To our knowledge, this study provides the first peer-reviewed assessment of the performance of scent dogs for detection of seabirds in both experimental and field settings (see Grimm-Seyfarth *et al*. 2021 for a recent review and Cristescu *et al*. 2015 for a similar study to detect Koalas *Phascolarctos cinereus*). Our results show that dogs can reliably detect particular seabird species, and can differentiate between the scent of one species and another. Our experiments showed that precision, sensitivity and accuracy of detection of multiple Storm Petrel scented targets, that were placed in a stone wall to reflect breeding densities at a large colony, improved with training and reward, and all measures averaged > 90% after five days of trials. There was a strong positive correlation between the number of sites indicated by the dog, and the true number of targets, over the range of target densities tested. We stress however, that the scent targets were prepared from a small number of individual birds, and present only the 'residual' odour that had been transferred from feathers to cotton pads. As such, the targets are unlikely to present the full odour spectrum of a colony of Storm Petrels and the intensity of odour presented may be considerably lower than that emitted from live birds, especially where breeding densities are high and burrows shallow. The extent to which scent dogs would be able to differentiate the scent of individual nests where nest density is high, is unclear. These findings should be treated with caution therefore, and field trials in study plots of known breeding density are required to assess dog performance, and to determine whether high success rates can be achieved across a range of breeding densities and nest site characteristics. Detection of occupied Shearwater nests was reliable for birds breeding in nest boxes, where nests were close to the surface and the confines of the box will tend to have retained the birds' scent. Indeed, we found that the dog's indication of Shearwater presence more closely reflected the history of occupation during the

Dog B 'Dewi' at work on Ramsey Island.
© Greg Morgan

preceding week than the presence of birds at the time of the trial, leading to false positives, where boxes were occupied by non-breeding birds, which were not resident every day. Detection of occupied natural breeding sites, which were often associated with Rabbit burrows, was less reliable, likely because nest chambers tended to be further from the surface (up to several metres), and the interconnected nature of nesting burrows will tend to disperse Shearwater scent. We found that occupied burrows were less likely to be detected by the dog where burrows were deep, where Rabbits were clearly present, and where burrows were known to be occupied by two birds, rather than a single bird, at the time of the trial. Visual inspection of the nest boxes revealed that cases of daytime occupation by two birds during the survey period (N = 9) always related to non-breeders, and it is possible that the lower probability of detection of these nests was due to the intermittent nature of their occupation by Shearwaters, resulting in lower scent concentrations. The difference in the likelihood of detection of cavities occupied by non-breeders between natural and artificial sites is noteworthy and highlights the importance of validation of scent dog performance in circumstances that conform as closely as possible to the detection task the dog is required to undertake. We also highlight that dog B had not been specifically trained for detection of Shearwater burrows, and the detection of more challenging targets (i.e. deeper burrows and those associated with Rabbits), had not been reinforced by conditioning. It is possible that the dog was responding to cues other than scent, such as sound from the burrow. A period of training, in which detection of more challenging targets is rewarded, and false alerts (e.g. of burrows occupied by non-breeders) are not rewarded, might result in a higher level of performance. However, it is currently not clear to what extent the detection of deep burrows, and non-detection of those occupied by non-breeders, might be mutually incompatible. In situations where burrow depth, or high density of nests prevents the reliable identification of all nests sites, the use of a scent dog may nevertheless be beneficial, in order to indicate the presence of the target species within a much larger area of search, where the occurrence of the species is uncertain (see Barros *et al*. 2020 for a remarkable example of the former, involving scent dogs to locate Elliot's Storm Petrels *Oceanites gracilis* breeding in the Atacama Desert). In our own study, after the conclusion of the field trials, dog B located a previously unknown sub-colony of Manx Shearwaters on the study island. In a survey context, scent dogs may have great value in delimiting the approximate area of occurrence of the target species, which is often a source of considerable uncertainty in the estimation of population size of burrow nesting birds. The use of a scent dog could provide a rapid and effective means of identifying the location and extent of areas to be surveyed subsequently by more labour-intensive methods.

The search effort required for canine assessment of potential nesting cavities was considerably lower than that required by conventional playback methods. We estimate that scent dog detection of Storm Petrel nests in a wall would achieve on average 4.8 (range 2.5–9.3) times greater coverage per unit time of sampling effort than playback methods, and it would be possible to survey six times as many Shearwater burrows using a scent dog, compared with playback. However, any efficiencies in terms of survey rate (area of habitat per hour), need to be weighed alongside the number of hours per day a scent dog and human fieldworkers can operate, the number of fieldworkers required per dog, and the relative accuracy of each method.

Whilst canine scent detection and discrimination ability is extremely acute, it will never be 100% accurate and the consequences of false positives and false negatives vary hugely according to the context. There is likely to be a trade-off between high precision (which minimises false positives) and high sensitivity (which minimises false negatives). False positives are not problematic where alerts can be verified directly (e.g. the identification of wind turbine collision casualties, Mathews *et al*. 2013), and in such studies sensitivity should be maximised, even at the expense of reduced precision (Bennett *et al*. 2020). However, for burrow nesting seabirds it is not easy to directly verify the precision of detections and follow-up verification relies on methods that typically have high false-negative rates (e.g. playback). Here, a more equitable balance between false positives and false negatives rate might be preferred. As has been demonstrated in many contexts such as the detection of disease and narcotics (e.g. Bird 1996), in situations where incidence is low (e.g. if individuals are tested at random for a rare condition, or locations are tested at random where target densities are extremely low), a high proportion of alerts may be false, even when the false-positive rate is low (since the false positive rate relates to the number of true negatives, which will be very common in these situations). The consequence of alerts being erroneous are often severe. In an avian survey context, in the absence of independent variation, a high proportion of alerts being false (i.e. low precision) would lead to large over-estimation of population size, if sensitivity is high (i.e. most genuine targets are located). The proportion of erroneous detections can be reduced by ensuring that testing is carried out in a suitably targeted manner, for example by surveying areas known to hold, or deemed to be highly suitable, for the target species. In situations where scent dogs are being used to determine the presence of a species which is very scarce, or not known with certainty to occur in the area (e.g. detection of invasive predators on seabird islands), it is important to regularly test and reward the dog with planted targets, and wherever possible, independently verify any alerts.

The performance of all tasks by dog A improved considerably over time, as correct responses were rewarded and false positives were ignored. We recognise that a fully-trained detection dog and professional handler would be expected to show consistently higher levels of performance than were achieved in the initial stages of our study. However, our results highlight the importance of regular, ongoing data collection and real-time analysis to indicate any lapses in performance, such as in Experiment 3.1. Although the cause of the lapse in performance was not identified with certainty, revision of the methods of sample preparation and storage, to minimise risk of cross-contamination of samples, and a period of retraining led to greatly improved performance. Methods of scent preparation and experimental protocols for testing detection and discrimination ability vary widely across studies, and may confound the target scent with the scent of individuals involved in preparation of different scents, or packaging material (Johnen *et al*. 2017), or preparation equipment (Centre for Protection of National Infrastructure 2020; DeGreeff 2020; Lazarowski *et al*. 2020). Reliable detection and discrimination of potential targets under field conditions requires training and testing on samples obtained from multiple sources. Elliker *et al*. (2014) found that the performance of dogs trained to detect cancer samples dropped when samples from new patients were introduced, indicating that the earlier high accuracy observed to the training odours was due to dogs memorising the small number of the individual target samples rather than the common odour profile. The training of scent dogs for detection of seabirds should use targets prepared from samples from as large a number of individuals and locations as possible, to create a common odour profile.

We hope this study will stimulate discussion and motivate researchers, conservation managers and others to consider the wider use of scent dogs for a range of seabird-related fieldwork. Whilst there is no doubt that canine olfactory ability is sufficient to detect, and differentiate among, seabird species, the extent to which these abilities may be employed for seabird monitoring is currently not fully clear, and potentially under-utilised. Scent dogs may certainly be employed to search efficiently for burrow-nesting seabirds on islands where breeding is suspected, or is known to have occurred in the past, but current breeding status is unknown. In addition, in areas where breeding is known to occur, but colonies cover huge areas in difficult terrain, scent dogs may play an important role in delimiting the spatial extent of breeding areas, as a preliminary to more intensive survey methods. It is possible that with further refinement and validation, scent dogs could be used to quantify the density of cavity-nesting seabirds where nest spacing is sufficiently wide to enable resolution of individual nest sites. Currently it seems improbable that dogs would be able to survey complex habitats, such as deep talus and boulder beaches, where breeding densities may be extremely high and nests may lie several metres below the ground surface.

There are several models of conservation dog deployment, including contracting professional handlers with trained dogs, to researchers training their own animals. The costs and benefits of each approach will vary according to circumstances. It should be noted that whilst authors of popular accounts claim that no prior expertise is required to deliver effective scent training and that any dog can be scent trained, many authors of peer-reviewed studies (e.g. Beebe *et al*. 2016) emphasise the need to involve professional dog trainers and to select individual dogs for scent work with extreme care (DeMatteo *et al*. 2019 and references therein), since they require a high drive and in consequence may not make easy pets, or be able to readily switch between periods of working and non-working. Furthermore, an effective working scent dog must also possess an appropriate temperament to avoid undue disturbance to sensitive wildlife and remain task-focussed for long periods of time. A strong bond between the dog and handler, and the handler's ability to correctly interpret the dog's signals are also of paramount importance (Beebe *et al*. 2016; DeMatteo *et al*. 2019; Lazarowski *et al*. 2020).

Figure 4. Dog A 'Islay' training in North Yorkshire. © Mark Bolton

There are numerous commercial scent dog training programmes in UK, some of which have a national network of trainers and bespoke tiered accreditation systems for the handler. The UK government Centre for Protection of National Infrastructure has also recently produced a number of extremely detailed guidance notes on the training, testing and welfare of scent dogs and for specifying, procuring and implementing detection dog services, across the full range of sectors (www.cpni.gov.uk/canine-detection-guidance-notes). Porritt *et al*. (2015) describe an odour discrimination test developed by the UK Defence Science and Technology Laboratory which is designed to allow rapid quality assurance of odour recognition by detection dogs, as a component of accreditation. There is a need for rigorous, standardised protocols to scientifically validate conservation detection dogs' capabilities, in settings that reflect the environment of the search task (Mathews *et al*. 2013; Long *et al*. 2007; Reed *et al*. 2011). This will lead to better informed decisions regarding capability, vulnerability, and risk analysis (Lazarowski *et al*. 2020).

Acknowledgements
We thank Sally Sanford for stimulating discussion on the potential of scent dogs for conservation monitoring and for demonstrating scent dogs in action. We also thank two anonymous reviewers for helpful comments on an earlier draft of this paper.

References

Animal Welfare Act 2006. www.legislation.gov.uk/ukpga/2006/45/contents

Barros, R., Medrano, F., Silva, R., Schmitt, F., Malinarich, V., Terán, D., Peredo, R., Pinto, C., Vallverdú, A., Fuchs, J. & Norambuena, H. V. 2020. Breeding sites, distribution and conservation status of the White-vented Storm-petrel *Oceanites gracilis* in the Atacama Desert. *Ardea* 108: 203–212.

Bates, D., Maechler, M., Bolker, B. & Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67: 1–48.

Beebe, S. C., Howell, T. J. & Bennet, P. C. 2016. Using scent detection dogs in conservation settings: a review of the scientific literature regarding their selection. *Frontiers in Veterinary Science* 3: 96.

Bennett, E. M., Hauser, C. E. & Moore, J. L. 2020. Evaluating conservation dogs in the search for rare species. *Conservation Biology* 34: 314–325.

Bergström, A., Frantz, L., Schmidt, R., Ersmark, E., Lebrasseur, O., Girdland-Flink, L., Lin, A.T., Storå, J., Sjögren, K-G., Anthony, D., Antipina, E., Sarieh Amiri, S., Bar-Oz, G., Bazaliiskii, V. I., Bulatović, J., Brown, D., Carmagnini, A., Davy, T., Fedorov, S., Fiore, I., Fulton, D., Germonpré, M., Haile, J., Irving-Pease, E. K., Jamieson, A., Janssens, L., Kirillova, I., Kolska Horwitz, L. K., Kuzmanovic-Cvetković, J., Kuzmin, Y., Losey, R. J., Ložnjak Dizdar, D., Mashkour, M., Novak, M., Onar, V., Orton, D., Pasarić, M., Radivojević, M., Rajković, D., Roberts, B., Ryan, H., Sablin, M., Shidlovskiy, F., Stojanović, I., Tagliacozzo, A., Trantalidou, K., Ullén, I., Villaluenga, A., Wapnish, P., Dobney, K., Götherström, A., Linderholm, A., Dalén, L., Pinhasi, R., Larson, G. & Skoglund, P. 2020. Origins and genetic legacy of prehistoric dogs. *Science* 370, 557–674.

Bird, R. C. 1996. An examination of the Training and Reliability of the Narcotics Detection Dog. *Kentucky Law Journal* 85: 405–433.

Bolton, M., Sheenan, D., Bolton, S. E., Bolton, J. A. C. & Bolton, J. R. F. 2017. Resurvey reveals arrested population growth of the largest UK colony of European Storm-petrels *Hydrobates pelagicus*, Mousa, Shetland. *Seabird* 30: 15–30.

Browne, C., Stafford, K. & Fordham, R. 2006. The use of scent-detection dogs. *Irish Veterinary Journal* 59: 97–104.

Centre for Protection of National Infrastructure 2020. Canine detection guidance notes. Available online at www.cpni.gov.uk/canine-detection-guidance-notes. Accessed 27/12/2020.

Cristescu, R. H., Foley, E., Markula, A., Jackson, G., Jones, D. & Frere, C. 2015. Accuracy and efficiency of detection Dogs: a powerful new tool for koala conservation and management. *Scientific Reports* 5: 1–6.

Dahlgren, D. K., Elmore, R. D., Smith, D. A., Hurt, A., Arnett, E. B. & Connelly, J. W. 2012. Use of dogs in wildlife research and management. 7th edition. In: Silvy N. J. (ed.) *The Wildlife Techniques Manual Volume 1*: 140–153. The John Hopkins University Press, Baltimore.

DeGreeff, L. E. 2020. *Elucidation and Durability of Odor Profiles of K9 Nose Work Materials. (Report No.6180/0306)*. Naval Research Lab, Washington, DC.

DeMatteo, K. E., Davenport, B. & Wilson, L. 2019. Back to the basics with conservation detection dogs: fundamentals for success. *Wildlife Biology* doi.org/10.2981/wlb.00584

Elliker, K. R., Sommerville, B. A., Broom, D. M., Neal, D. E., Armstrong, S. & Williams, H. C. 2014. Key considerations for the experimental training and evaluation of cancer odour detection dogs: lessons learnt from a double-blind, controlled trial of prostate cancer detection. *BMC Urology* 14: 22.

Gazit, I. & Terkel, J. 2003. Explosives detection by sniffer dogs following strenuous physical activity. *Applied Animal Behaviour Science* 81: 149–161.

Gerritsen, R. & Haak, R. 2017. *K9 Explosive and Mine Detection*. Brush Education Inc, Canada.

Gilbert, G., Gibbons, D. W. & Evans, J. 1998. *Bird Monitoring Methods: a manual of techniques for key species*. RSPB, Bedfordshire, UK.

Grimm-Seyfarth, A., Harms, W. & Berger, A. 2021. Detection dogs in nature conservation: A database on their world-wide deployment with a review on breeds used and their performance compared to other methods. *Methods in Ecology and Evolution* 12: 568–579.

Hewings, R. 2019. *Scent training for every dog*. First Stone, Warminster.

Jezierski, T., Ensminger, J. & Papet, L. 2016. *Canine Olfaction Science and Law: Advances in Forensic Science, Medicine, Conservation, and Environmental Remediation*. CRC Press, Florida.

Johnen, D., Heuwieser, W. & Fischer-Tenhagen, C. 2017. An approach to identify bias in scent detection dog testing. *Applied Animal Behaviour Science* 189: 1–12.

Kotthoff, M. & Nörenberg, S. 2016. Odor and nutrition. Part 2: traits of odors. *Ernahrungs Umschau* 63: 22–30.

Lazarowski, L., Krichbaum, S., DeGreeff, L. E., Simon, A., Singletary, M., Angle, C. & Waggonner, L. P. 2020. Methodological Considerations in Canine Olfactory Detection Research. *Frontiers in Veterinary Science* 7: 408.

Lesaffre, E. & Spiessens, B. 2001. On the effect of the number of quadrature points in a logistic random effects model: an example. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 50: 325–335.

Long, R. A., Donovan, T. M., Mackay, P., Zielinski, W. J. & Buzas, J. S. 2007. Comparing scat detection dogs, cameras, and hair snares for surveying carnivores. *Journal of Wildlife Management* 71: 2018–2025.

Mathews, F., Swindells, M., Goodhead, R., August, T. A., Hardman, P., Linton, D. M. & Hosken, D. J. 2013. Effectiveness of search dogs compared with human observers in locating bat carcasses at wind-turbine sites: a blinded randomized trial. *Wildlife Society Bulletin* 37: 34–40.

Perkins, A. J., Douse, A., Morgan, G., Cooper, A. & Bolton, M. 2017. Using dual-sex calls improves the playback census method for a nocturnal burrow-nesting seabird, the Manx Shearwater *Puffinus puffinus*. *Bird Study.* 64: 146–158.

Porritt, F., Mansson, R., Berry, A., Cook, N., Sibbald, N. & Nicklin, S. 2015. Validation of a short odour discrimination test for working dogs. *Applied Animal Behaviour Science* 165: 133–142.

R Core Team 2019. R: A language and environment for statistical computing. R Foundation for Statistical computing, Vienna, Austria. www.R-project.org.

Reed, S. E., Bidlack, A. L., Hurt, A. & Getz, W. M. 2011. Detection distance and environ mental factors in conservation detection dog surveys. *Journal of Wildlife Management* 75: 243–51.

Rosell, F. 2017. Die Welt der Gerüche. Spezial-Spürhunde im Einsatz. Kynos Verlag, Nerdlen/Daun, Germany.

Uemura, E. E. 2015. *Fundamentals of Canine Neuroanatomy and Neurophysiology*. 1st ed. Wiley, Danvers, Massachusetts.