Journal of Veterinary Behavior 52–53 (2022) 14–20

Contents lists available at ScienceDirect

Journal of Veterinary Behavior

journal homepage: www.journalvetbehavior.com

Discussion

The use of tracking/man trailing dog results as evidence in courts

Gertrud Adelheid Alice Schoon*

Animal Detection Consultancy, Vorchterenkweg 4, 8193KM Vorchten, The Netherlands

ARTICLE INFO

Article history:
Received 5 February 2021
Revised 30 March 2022
Accepted 1 April 2022
Available online 8 April 2022

Keywords:
Man trailing
Individual human scent
Evidence
Learning

ABSTRACT

Results achieved by man trailers/trackers is often accepted as evidence in courts. To be used as such, the results should be founded on scientific proof. With respect to the uniqueness of human individual scent, there is sufficient scientific foundation, but with respect to animal learning and cognition we now know that great care must be taken to ensure animals are responding to the intended cue. A review of man trailing studies indicate that these do not prove the essential question: are the dogs responding to the individual human odor cue when matching a scent article to a track? A suggestion is made for testing individual dogs based on guidelines set by the Scientific Working Group on the use of Dogs and Orthogonal Detector Guidelines. Finally, points that should be addressed when assessing such evidence in courts are provided.

Introduction

The performance of dogs in identifying people based on scent traces, and the subsequent use of such identifications in court, has been the subject of debate for several decades. Issues I tackled in my 1997 PhD thesis (Schoon, 1997) on scent lineups still arise in tracking and man trailing with respect to individuality of human odor, odor consistency and stability, and the differences between what we try to teach the dogs and what they actually learn. On top of that, tracking and man trailing include additional variables since tracking and trailing are not conducted in a controlled setting as lineups are, but are conducted outdoors under variable conditions and may include unexpected distractions.

Man trailing remains a subject of fascination. In "tracking" the dog is thought to follow mainly ground disturbance and in "trailing" the dog is thought to be following the individual scent of the person who laid the trail. In theory, man trailing fits a dog like a glove: just teach a dog that the person whose scent he is presented has his food, and he will track down this person as naturally as a wolf hunt down his prey. Dogs live in a scent world where they use their noses to hunt, we live in a visual world and have little understanding of what happens in this other scent world. It is a mysterious world that the dog naturally knows and uses, a world we can profit from by guiding them to look for people as their prey, a world that has been the inspiration for many legends, books, and films (Pemberton, 2013).

But in the world of evidence, this is not enough. As forensic evidence, the performance of dogs should be the subject of scrutiny and objectivity. We now know that wolves more often miss their prey than catch it – wolves are ranked 7th in a list of hunting mammals with a 14% successful kill rate (BBC Wildlife Magazine). So, how do dogs perform? What do we know of human scent and how it "behaves", what do we know of dogs and the way they smell and learn, what can we learn from scientific work conducted in this field? How does this all translate into guidelines for evaluating man trailing evidence presented in courts?

Man trailing is done by police and volunteers in many countries worldwide. In my own country (The Netherlands) it is used by volunteers for search and rescue, where any find is good, and misses are sad but inevitably happen. But man trailing is not conducted by the police, or presented as evidence in court. If used as court evidence, as it is in the US and in several other countries, it needs to meet different criteria than when used for search and rescue. And it certainly needs to meet more criteria than "common knowledge" to be gained from internet. The central part of this paper critically reviews some (not always published!) studies, and includes personal observations collected over the years. This is the foundation for the conclusions and resulting advice.

* Address for reprint requests and correspondence: Gertrud Adelheid Alice Schoon, Animal Detection Consultancy, Vorchterenkweg 4, 8193KM Vorchten, The Netherlands. Tel.: +31 578631683.

E-mail address: animaldetectionconsultancy@gmail.com

https://doi.org/10.1016/j.jveb.2022.04.001
1558-7878/

Individuality of human scent

Human scent is the combination of volatile organic components that are the product of bacterial breakdown of different gland excretion and the breakdown of corneocytes (also known as skin-rafts) (Eckenrode et al. 2022). The individuality of human scent has only been established quite recently with the advances of analytical chemistry. Most of the earlier evidence was based on the performance of dogs, varying from dogs who recognized their owner (Romanes, 1887), who could or could not discriminate between twins (Hepper, 1988), or who could identify perpetrators based on scent traces left behind at the crime scene (Schoon, 1997). But now it has been established that people do produce individually distinct odorants and that a person produces the same scent profile over a period of at least a couple of weeks. This makes it possible to use odor as a biometric (Schoon et al., 2009).

However, it is also clear that there is more information in scent than just human individuality. Sex differences and differences along ethnical lines have been shown through chemistry (Colón-Crespo et al., 2017). Differences in scent because of emotional state (especially stress) has been established in empirical research using both people and dogs as test subjects (e.g., de Groot and Smeets, 2017; D'Aniello et al., 2018), and recently in chemistry as well (Smeets et al., 2020). It is thus perfectly possible that dogs learn to respond to the scent of fear.

In conclusion: there is a lot of information available in the odor signature of a person at a given moment, which includes but is not limited to a personally unique odor.

Effect of surface differences, material differences, environmental circumstances, and aging

Once deposited, biological material interacts with the surface and this changes the scent profile in the air. This is a common phenomenon that has been studied with different types of bodily fluids such as blood, but also with human scent deposited by hand on gauzes made of different types of material (Prada, 2010). So: volatile molecules bind to different surfaces they are deposited on very differently; porosity has an influence (Rust et al., 2016); temperature directly influences the availability of molecules in the air and indirectly influences degradation. Biological activity (bacteria, fungi), moisture, UV light and air pressure are some other environmental factors that have a direct effect on the availability of odorants, as well as influencing degradation (Caraballo, 2014, Eckenrode et al., 2022). And time itself leads to differences through depletion of the original source material, as well as allowing degradation to set in. Processes outdoors affecting the change in odors available on the track have been described by Syrotuck in 1972, but to date they have not been confirmed by chemical analysis on tracks themselves due to the extremely low concentrations of relevant volatile organic compounds.

In conclusion: a constant individual profile in human scent deposed on a surface in time has not been established through chemical analysis yet. On the contrary, there is a lot of variability to be expected based on research of other biological traces.

Man trailing breeds

The original man trailing breed, the bloodhound, has been used to track people since the middle ages. Wikipedia describes the bloodhound as follows: "This breed is famed for its ability to discern human scent over great distances, even days later. Its extraordinarily keen sense of smell is combined with a strong and tenacious tracking instinct, producing the ideal scent hound, and it is used by police and law enforcement all over the world to track escaped prisoners, missing people, and lost pets."

Often the keen sense of smell of the bloodhound is attributed to their large olfactory epithelium, which is around 380 cm^2 as compared to the German shepherd's 150 cm^2. However, the highest sensitivity in a mammal for an odor (of a fox) has been measured in a mouse – with only 1.5 cm^2 of olfactory epithelium. Nowadays, the variety in olfactory receptors is thought to be the key factor in sensitivity, and rodents are superior to dogs in this respect.

Different dog breeds differ in genes for olfactory receptors, but they also differ in degree of trainability that may explain their differences in performance of typical scent tasks (Lesniak et al., 2008). It is for this reason that many man trailing handlers have switched to other (smaller) hunting breeds that are considered more trainable than the bloodhound. There are many types of breeds used in man trailing, as a quick perusal of man trailing websites on the internet shows.

In conclusion: man trailing is considered to be more the result of training, than a natural capability of a particular dog breed.

Human cues dogs to which respond

Dogs have been a domesticated species for a long time and have evolved a great awareness of human cues in the process. Even untrained, they pay close attention to our faces and in particular our eyes (e.g., Miklosi et al., 1998), they seem to understand that someone who "knows" is a more reliable source of information (Maginnity and Grace, 2014), and they naturally differentiate in odors between stressed and non-stressed people (D'Aniello et al., 2018). Through training, dogs learn to respond even more specifically to human signals, a process used in training "assistance dogs" that help their owners with their daily activities.

This tendency is a great disadvantage in scent training, where the dog is expected to follow his nose and respond based on an odor, and not to a human cue. People working with dogs have long known this problem (described as "Clever Hans") but the scientific detection canine world was alarmed by a paper demonstrating the effect of handler expectation (Lit et al., 2011). Also, dogs differ in the degree they will respond to olfactory cues over visual cues (Lazarowski et al., 2020), and in training, dogs build up persistent expectations that influence their performance depending on the context they are in (Gazit et al., 2005, Porrit et al., 2015).

In conclusion: dogs do not necessarily pick up the intended odor cues only in scent training, and great care must be taken they do not pick up unintended human or context specific cues.

Scent specific tracking/trailing studies

The debate over what cues a tracking or trailing dog focuses on is one that has never been resolved. The scent picture left on the ground has been described to be a mixture of human scent material the person has left behind directly on the ground or in skin grafts, volatile molecules that are immediately released through damage done by the person walking (crushed vegetation) and volatiles that are released more slowly when the changes caused by the footsteps find a new equilibrium (biological ground activity) (Syrotuck, 1972). The consensus is that a "tracking" dog follows the specific scent profile a person leaves behind while walking which (depending on the surface) consists of predominantly crushed vegetation and/or ground disturbance; and that a scent specific "trailing" dog focuses on the individual human odor in that specific scent picture, which he picks up based on the presentation of a starting scent (often called "smeller"). In both cases, dogs can be trained to be track-faithful and not be fazed by changes in sur-

face type, cross tracks, turns and obstacles, or distractions along the way.

Over the years, a sensible conclusion has been that dogs are capable of doing what they have been trained to do (e.g., Blum, 2017). So, if the dogs are trained to focus on ground disturbance, they will do so, if they are trained to air-scent (so not paying attention to the surface at all but sniffing the air to catch a whiff of human scent there) they will do that. The proof of the pudding is in the eating, and careful testing reveals what an animal has actually learned in the training. Man trailing test levels focus mainly on increasingly older tracks, but do not usually resolve the question if the dogs are really only focusing on the human scent all the time.

In a study on trailing bloodhounds (Harvey and Harvey, 2003), 8 bloodhounds were given 5 trails each, ending with a tracklayer to identify. The track was laid by 2 people who walked between 0.8 km and 2.5 km together, covering different surfaces, varying in degree of contamination by other people walking around, diverging apart about 15 m before the end of the trail (a so-called Y track). The dogs were brought in to search after 48 hours, just after the 2 tracklayers had been put back at the end of the trails. The presence of the tracklayers at the end of the track meant there was a significant scent plume available. The 3 "novice" dogs completed 2, 2, and 3 of these tracks and only one of these led to an incorrect identification of the tracklayer, the 5 "expert" dogs completed all but 1 track with perfect identifications. Despite the study being "blind" as to who the tracklayer was, there are a couple of problems with this study. A compulsory "negative check" was conducted prior to each start, which consisted of offering the dog a blank "smeller", claiming this proved the dogs would only work when there was a matching track. But this does not really prove anything with relationship to tracking since the dogs can easily learn that a "blank smeller" means "do not walk", and a "scented smeller" means "walk". Was the direction of the track known? Nothing is said on that point either. If even a general direction was known, the dogs could have come quite close to the 2 track layers without necessarily "tracking" or "trailing" if the tracklayers walked consistently in a general direction, as described in other studies by the same authors. Nothing is said of the wind direction. So, the tracklayers standing at the end of the tracks could be blowing their scent in the direction of the approaching dog, making it possible for the dog to be doing the discrimination and final identification through comparing the scent on the "scented smeller" and the fresh smell of the entire person, without actually having to pick it up from the track. Be aware that many dogs are extremely good at air scenting: they have a special method of breathing to allow for very long inhalations (Steen et al, 1996), building up odor concentrations in the olfactory chamber in their noses, and especially hunting dog breeds exhibit incredible air-scenting sensitivities. So, despite the very good results of the "expert" bloodhounds (trained over 18 months) the results of this study are debatable.

Setting up proper tests is really difficult. A simple test was conducted by myself and Jan Zoodsma, at the time chief of the Dutch canine division in 2004 with bloodhounds used then by the FBI. We stepped out of a car at a leisure center, and walked back along the way we had just driven to a restaurant – 3 straight legs, 1 left turn and 1 right turn, while the car followed us back. At the restaurant we got back into the car and drove away, taking care to not cross our walking path in any way. The next day we returned to the restaurant, again not crossing our walking path. The man trailing teams were given the starting point of our track at the leisure center, but they could not locate us. The track was about 750 m long and 1 day old which was well within their operational capability. The test was conducted double blind (the car driver was waiting with us in the restaurant) to exclude any cueing. They initially started off in the wrong direction and worked that side of the neighborhood quite thoroughly before deciding to restart, but after a couple of hours we stopped the test, having had our fill of coffee for the day. But these same dogs were quite capable of picking out who had worn T-shirts I had brought over from The Netherlands for this purpose in a "match to sample" setup. Five of those T-shirts had been worn by people that were not present, and one by my colleague, so no T-shirt had been worn by me. My colleague and I were the "lineup" and walked a typical Y track for this test (first together and then splitting up in different directions). The handlers were given the 6 anonymized T-shirts and randomly chose the sequence to use them. The dogs gave correct negative indications for the first 4 T-shirts and correctly followed my colleague's track – he had worn the 5th T-shirt they presented. Then expectations set in and the 6th shirt was incorrectly matched to me…. But the dogs responded correctly to 5 out of 6 comparisons of T-shirts to people.

An interesting review of man trailers used in experiments to illustrate the incredible preservation of human scent on debris left after explosions can be found in a paper by Curran et al. (2010). Most of these setups involve a "terrorist" and a decoy walking a short track together and then parting and hiding (Y-tracks). In these studies, dogs were given scent from the debris on a "smeller" made with a "Scent Transfer Unit" that collected odorants on the explosive debris onto a gauze pad to match to the "terrorist". Some studies had additional decoys present to ensure the dogs did not just indicate any person, and in some studies care has been taken to prevent other cues by conducting the experiments double blind. But in general, these tests at best indicate how persistent human scent is in the absence of biological degradation. They do not prove anything about the dogs actually using human odor in the track. The tracks were sometimes just a few meters so the dogs could be air scenting directly, or it was a longer Y setup where one combined track led to where the people were hiding, and the dogs could switch from following the ground disturbance trail to air scenting when getting close. Switching attention from one cue to the other depending on availability in searching is widespread in the animal kingdom.

The Essex police conducted a series of tests with their 4 bloodhounds in 2001. They were triggered to do this by Patrick Kirby, who as police officer worked a German Shepherd trained in the same way as the bloodhounds, a training method they had dubbed the "Scented Article Method". Kirby realized his dog only performed well in trails where he knew the trail layout. And that if he ran them blind, not knowing where the tracklayer was, the dog failed. There were 4 bloodhounds working with experienced officers at the time within the Essex police force that all had operational "finds" to their name when the decision was made to test them in a double-blind manner. Each of the bloodhound teams were given approximately 10 trails, between 3-9 hours old. The starting point was some landmark on a pavement so the track could go 1 of 2 ways. In about 66% of the 38 experimental trails the dogs started off in the right direction. This result was slightly biased by handlers familiar with the general area of search and knowing that in some cases, 1 direction could be excluded since it led to a dead end. After that there were series of "straight legs" and "turns". None of the tracks were completed. Analyzing the tracks, you could see the almost random choices the dogs made at each decision point – the first choice "straight on" was chosen correctly by 17 of the 25 of the dogs on track at that point, but only 9 made it past the first corner and only 2 made it past the second corner. Unfortunately, this material was never published but I was given permission to share it when relevant.

Our federal police colleagues in Belgium have also tried to train man trailing dogs several years under supervision of an experienced Swiss trainer. The 2 dogs were tested in a double-blind test on a soft surface in an uncontaminated area in 2013. This was a

complete disaster – the dogs showed no willingness to trail at all – and so these dogs were retired (Van Krunkelsven, pers. com). Police colleagues in Stuckenbrock, Germany, had also been training man trailing dogs for several years when we visited them in 2011. Their conclusion: sometimes it works, sometimes it does not, but they had no idea what factors contributed to success and what did not. At the time, they were deployed in missing people cases where they had run out of other options, but their tracking results were not used as evidence in criminal investigations.

Another study in Germany (Wolf, 2016) seemed to show excellent performance in aged tracks by bloodhounds. These tracks were relatively short (about 500 m) and did not contain negative controls. They were laid and then aged for a day, a week, or a month. Just before the dog was put to work on a track, the tracklayer was put back at the end of the track. Interestingly, the aging did not influence the performance of the dogs at all: some dogs could do all ages (even those of a month!), others none. Wolf's conclusion was the successful dogs were perhaps not paying attention to the odor on the track, but were particularly good air scenters, and advised future studies to not put tracklayers at the end of track to prevent this from happening.

In the Netherlands, instructors of the Canine Unit of the National police have evaluated many teams that were confident they could perform tracking/trailing. The purpose was to help volunteers to understand what the police was looking for. One example that I witnessed will illustrate this. In this case we cooperated with a volunteer group of man trailers. They had 3-man trailing dogs, and we laid double blind tracks for each of the dogs that were then aged 24 hours, since the group indicated that tracks of that age were optimal for their dogs. The tracks were not long, 500-700 m, but there was no tracklayer at the end to prevent air scenting from being a possible cue. This resembles a situation where a person is "picked up". In 2 tracks the dogs started off in the wrong direction and never came close to the track again, while the third track was run at such a distance from where it was laid that the handler himself was not happy. The group was dismayed at the poor results but open to our conclusion that they had been unintentionally influencing their dogs much more than they realized. In training, the handler would usually not know where the track was laid, but there would be an instructor following who did know. I witnessed some of their training runs. I could judge when the dog was "right" and "wrong" when he changed direction based on the behavior of the instructor, although I was blind to the direction of the track. And if I could see this, a dog certainly could also use that as a cue. To date, none of the dozens of teams evaluated in this manner could perform double blind tracks.

The final tracking/trailing study to which I want to refer is a recent 1 published by Woidtke et al. (2018). In his setup, dogs were about 100 m away from 2 people hiding behind different houses. The teams were given at random 1 of 3 odors to start from: 2 of these were from the 2 people hiding, the third was from someone not present at all. A simple setup where the dogs seemed to do pretty well in "identifying" the person whom they were given scent of, but again a setup with problems. First: a random odor should mean that 33% of the trials should have been "negatives." The results were significantly different: many less than expected trials were negatives – so was this really random? Or really blind (since these negatives almost never led to a mistake)? Second, there is nothing preventing air scenting in this setup. A "positive choice" was if a dog crossed an imaginary line 50 m from the chosen person. This is pure air-scenting. This means that at most the dogs could match an article scented by a person to that person. This has nothing to do with using human scent on a trail, which is what man trailing is all about. Third, an independent observer at a trial day observed inconsistencies in the way the results were noted by the non-blind experimenter, that would lead to biased results (Goss, pers. com.). After several letters to the editors, the journal published an "expression of concern" about this paper (Jackowski et al., 2020) where they "suggest that the study results are taken with care especially for the application to forensic casework."

No experimental study can do justice to operational reality involving identification of suspects that KNOW they are suspects in a crime. A dog may learn that identifying such nervous people leads to a very happy handler. This knowledge has been the reason for the Dutch police to stop direct interaction between dogs and people in scent identification lineups back in the 1960′s (Schoon and Haak, 2002). But direct physical contact in identification in man trailing is unfortunately still common.

Finally, a word of caution is appropriate when evaluating any work or studies with man trailers. Dogs have great memories. We do not know exactly where the limits are: people gave up when teaching them words (Pilley and Reid (2011) described a border collie who remembered 1000+ words) and odors (Williams and Johnston (2002) stopped at 10, a recent study by Waggoner et al. (2022) showed that dogs who had been taught 40 odors still remembered them all after a year). It is obvious that dogs recognize people they know by scent. It is therefore essential to use people whom the dog does not know when studying the scent matching capability of dogs (be it from article to a person, or from an article to a person's track, or from a track to a person). When familiar with the people, the dog may have a complete scent image in his mind and use this as mnemonic: "Oh this is John!" when smelling Johns glove, "and this is John's scent on grass!." Results obtained with familiar people are a lot better than those with unfamiliar people, and we should not fool ourselves (Hale, 2017).

In conclusion: studies that have shown good results of man trailers are questionable with respect to scientific rigor or at best demonstrate the acute air-scenting match-to-sample capabilities of dogs. The studies do not particularly prove the attention of the dogs to human scent in the track itself. One study with a good setup but poor results has unfortunately not been published. Bad experiences by police forces who are held accountable for their work have shown that at best, man trailers are limited in their deployment and are not used in criminal cases. Search and rescue groups often do not have the necessary scientific discipline in testing themselves in a correct manner. The theory of man trailing – a dog using the human scent cue within the scent picture of the track – is an attractive theory that attunes with the ideas on how canines hunt, but unfortunately there is no scientific proof of dogs being capable of doing this consistently based on a particular type of training. Direct interaction with people during identification has been abandoned in European lineup procedures for decades for the obvious reason that dogs can learn to respond to nervous behavior a suspect probably displays.

Protocols and testing criteria

Does this mean dogs are not capable of using human scent cues in the track? No, it only means that to date, there is no evidence that any particular kind of training leads to dogs doing this in a consistent manner. But we know the human scent cue is available in the scent picture left behind, although it changes depending on the surface it has been laid on, changes depending on environmental factors, and changes (degrades) with age until there is no scent left. So, there is a changing human scent cue that dogs may, within limitations, learn to pick up and use. It is easier to do this with the odors of familiar people. An additional cue that is operationally relevant is the scent a scared person leaves behind – a perpetrator fleeing from a crime scene is comparable to the scared per-

son described in the chapter on human scent above that the dogs naturally differentiate from happy and neutral people. Another additional cue is typical nervous behavior of a scared person a dog may learn to recognize and use during the identification part of the process.

So, the question is not if dogs are capable of using the human scent cue, but the question is if they DO. And that can only be discovered by testing each individual dog. Testing to establish which cue this particular dog is responding to, and testing to establish the range of this capability. This should be part of the certification process. Then, these dogs must be deployed within the established range, usually described in some kind of Standard Operational Protocol.

Tests must meet scientifically accepted criteria and general guidelines have been described by SWGDOG (Scientific Working Group on the use of Dog and Orthogonal equipment Guidelines). They include odor recognition testing, comprehensive assessments, and double-blind assessments.

Odor recognition assessments are conducted to examine if the dogs correctly respond to the target odors they have been trained on. For this part of "man trailing", it is my opinion it is best conducted by testing the initial start of the trails, where the dog matches the scent on the starting "smeller" to the track this person has laid (without the possibility of air scenting!). The requirements for testing when a dog is brought to a starting point and given an odor include does he:

a) track the correct person where this person is 1 of a group of people who were there at the same time and they all fanned out and then left;
b) not track if given a smeller from someone who was never at that location;
c) track correctly repeatedly with randomly presented "smellers" from a variety of new people to prevent memory effects.

Naturally, the tracks should be of people with whom the dog is not familiar.

Comprehensive assessments are conducted to examine the quality of the search in different environments where the dog will be deployed. In "man trailing", this would mean searches that vary in degree of human and animal contamination, types of underground, obstacles, track age, track length, environmental conditions, time of day, season, way the trail ends, and distractions along the way. In all cases it is mandatory to conduct these trails at least "blind" in the sense that the man trailing team has no prior information on where the track is or who the tracklayer is, and that the assessors prevent cueing the team by always staying at a fixed distance from the handler (also when the search team has taken an incorrect turn) or by using a GPS system. Evaluations should be based on how closely the dog followed the track, and limits as to how far the team may deviate from it have to be agreed upon in advance. Again, these results reflect operational reality only if the tracklayer is unfamiliar to the dog.

Double blind assessments can be part of a comprehensive assessment or can be part of quality control. Here a handler is just provided with a scenario and is sent out to conduct the search as if it were an operational call. This includes negatives (the person whose scent is provided to the handler was never at that location but someone else was), pick-ups (this excludes an identification at the end of a trail), incorrect prior information (to examine handler influence) and positives with a variety of unknown people of different ages, sex, occupations, and ethnic backgrounds.

The capabilities of organizations/agencies will differ, depending on their handlers, their dogs, training facilities and time. These capabilities (and limitations!) must be reflected in assessment and deployment parameters. It is not correct to conduct assessments in rural areas in the quiet of the night, and to deploy the dogs in busy urban areas during the day. It is also not correct to assess dogs on soft surface tracks of up to 1 hour old with the tracklayer at the end and deploy such a dog in an industrial area to follow a 2-day old trail of a burglar. It is also not correct to train and assess dogs using familiar handlers as tracklayers and expect the dogs to work equally well on tracks from unknown people. And great care must be taken the dogs do not learn to respond to nervous people during the identification.

In conclusion, regular assessments must be conducted, documented, and reviewed to evaluate the operational capabilities and limitations of the dog teams in any organization/agency using the odors of unfamiliar people, and the teams must be deployed within their limits to obtain reliable (=repeatable) results.

Evidential value

Dogs are not and never will be perfect. In olfactory research using dogs, one must understand the limitations of the methods used and the number of dogs involved in assessing reliability and validity of the outcome (Lazarowski et al., 2020). But the same factors must be understood by any police force deploying (man trailing) dogs, as well as by any forensic team or legal institutes evaluating the results they achieve as evidence. Similarly, to other detection technologies, there will be trails followed correctly, perhaps leading to identifications if the person is in the vicinity, and to correct "negatives" when a dog does not start trailing since there is no trail to follow. But there will also be mistakes.

There will be misses. This is when a person was actually at the location the man trailer was given the smeller, but the dog could not pick up the trail, or when he does pick up a trail at the start but loses it on the way, or when he fails to identify the tracklayer. This type of error can be limited by deploying the dog well within the limits of their confirmed capability (established by assessments described above). If the dogs are reliable at 2-hour old trails but not at 4, one should not exceed the 2-hour limit. If the dogs are only reliable on soft surfaces, do not deploy them in an urban environment. It is this type of mistake that makes it impossible to conclude that if a dog does not pick up a trail, that person was never there. At best, it means that at the time the dog was deployed, the scent picture of that particular person could not be picked up by that dog – it may have become too weak, or the dog was not at its peak performance, or the odor of that particular person could have been difficult for the dog involved. Absence of proof is not the same as proof of absence.

There will also be false positives. This is when a person was not at the location the dog started at, but the dog picks up a trail anyway, sometimes even identifying a person at the end. This can be the result of incorrect training. Not training a dog on "negatives" in a correct manner can lead to dogs picking up a fresh trail, or even any other trail when given the command to search. The capability to deal with "negatives" is important to assess. A "negative" can also be the result of (unintentional) cueing due to handler expectations. Handlers who have any kind of case knowledge will have expectations, and it is impossible to not communicate this expectation to the dog they are working with. Prior knowledge as to who the suspect is, where a suspect lives, where a person was last seen, or even the description of a suspect will influence the handler's expectations and in turn this will influence the behavior of the dog, generating false positives in tracks and in identifications.

No work has been done in this field that can provide guidance on how sensitive (measure for how often they miss) or how specific (measure for how often they falsely identify) man trailing dogs are. For medical tests, moderately performing tests are

described as having 80-90% sensitivity and specificity (Power et al., 2013). In laboratory circumstances using lineups where the scent sample being offered to the dog was identical to the 1 in the lineup, Jezierski et al. (2010) achieved a sensitivity of 58% and a specificity of 44%. Marchal et al. (2016) found approximately 90% sensitivity when the samples were identical and 70% when the sample presented to the dog was from a different body part than the sample in the lineup, and a 100% specificity but they described a long training process. I would therefore estimate that in laboratory conditions, acceptable sensitivities are 70%-90% (meaning dogs miss in 10%-30% of the cases), and acceptable specificities are above 90% (meaning they give up to 10% false indications).

Based on these laboratory figures, there will be limitations to the value that can be attributed to man trailing evidence. Any kind of man trailing evidence presented in court will have to be considered in the light of the following aspects:

- Was the dog deployed well within independently assessed and documented variables in terms of age and length of trail, surface types, environmental conditions?
  - *This is necessary to assess how potentially successful the dog could be.*
- Was the dog reliable in indicating "negative" trails in documented assessments?
  - *This is necessary to assess the chances for false tracks and identifications.*
- What did the handler know, or what could he deduce, prior to being deployed with his dog?
  - *The handler should know nothing to limit the chance he inadvertently cues his dog.*
- Was the suspect already aware he was a suspect when the dog identified him?
  - *This should not be the case, or the dog may have reacted to the suspect's stress.*
- How was the performance of the dog in assessments and its health around the time of the deployment?
  - *This should be good both before and after since it proves stability, and some diseases first manifest themselves in loss of olfactory sensitivity.*

This information provides some background to the value of the evidence provided by the man trailing team in a particular case. It must be evaluated by someone versed in forensic evidence, who is up to date in knowledge on the canine sense of smell, human scent, and canine learning, as well as having practical experience in studies on working canines.

Based on my own experience in court cases, I would advise to use results of canines as corroborative evidence only, and not as proof of identification or proof of absence. In detection tasks, dogs are clearly very fast and sensitive, and can help determine places of interest in a very efficient manner. These places are then investigated using forensic techniques. The results of such investigations can be presented as evidence in court where their value can be queried. If human scent work is followed up by forensic investigations such as fingerprint identification or DNA profiling, the dog has done his work and the forensic results can be presented as evidence in court. But without independent forensic validation of identifications made by dogs I would strongly advise courts to not use them as proof.

In conclusion, the validity of canine evidence is limited if the dog was not deployed within his proven capabilities and in a manner that excluded the possibility of human influence. Even when used within proven capability limits it is wise to only use such evidence in conjunction with results of other reliable forensic techniques.

Ethical statement

For this paper no animal experiments were done. Compliance with regulations on the ethical treatment of animals is supported.

Conflict of interest

The authors declare no conflict of interest.

References

BBC Wildlife Magazine. https://www.discoverwildlife.com/animal-facts/mammals/hunting-success-rates-how-predators-compare/.

Blum, B., 2017. The hounds of empire: forensic dog tracking in Britain and its Colonies, 1888-1953. Law History Rev. 35 (3), 621–665.

Caraballo, N.I., 2014. Identification of characteristic volatile organic compounds released during the decomposition process of human remains and analogues. PhD thesis. Florida International University.

Colón-Crespo, L.J., Herrera-Hernández, D., Holness, H., Furton, K.G., 2017. Determination of VOC marker combinations for the classification of individuals by gender and race/ethnicity. Forensic Sci. Int. 270, 193–199.

Curran, A.M., Prada, P.A., Furton, K.G., 2010. Canine human scent identifications with post-blast debris collected from improvised explosive devices. Forensic Sci. Int. 199 (1-3), 103–108.

D'Aniello, B., Semin, G.R., Alterisio, A., Aria, M., Scandurra, A., 2018. Interspecies transmission of emotional information via chemosignals: from humans to dogs (Canis lupus familiaris). Anim. Cogn. 21 (1), 67–78.

de Groot, J.H., Smeets, M.A., 2017. Human fear chemosignaling: evidence from a meta-analysis. Chem. Senses 42 (8), 663–673.

Eckenrode, B., Riley, P., Dailey, A., Couch, R., 2022. Towards the development of a human scent model. In: DeGreeff, L.E., Schultz, C.A. (Eds.), Canines: The Original Biosensors. Jenny Stanford Publishing, pp. 275–315.

Gazit, I., Goldblatt, A., Terkel, J., 2005. The role of context specificity in learning: the effects of training context on explosives detection in dogs. Anim. Cogn. 8 (3), 143–150.

Hale, E., 2017. Canine human-scent-matching: The limitations of systematic pseudo matching-to-sample procedures. Forensic Sci. Int. 279, 177–186.

Harvey, L.M., Harvey, J.W., 2003. Reliability of bloodhounds in criminal investigations. J. Forensic Sci. 48 (4), 811–816.

Hepper, P.G., 1988. The discrimination of human odour by the dog. Perception 17 (4), 549–554.

Jackowski, C., Cattaneo, C., Broccard, A., Duembgen, L., 2020. Expression of concern regarding "Individual human scent as a forensic identifier using man trailing" by Woidtke L, Dreßler J and Babian C. Forensic Sci. Int. 282, 111–121.

Jezierski, T., Gorecka-Bruzda, A., Walczak, M., Swiergiel, A.H., Chruszczewski, M.H., Pearson, B.L., 2010. Operant conditioning of dogs (*Canis familiaris*) for identification of humans using scent lineup. Anim. Sci. Pap. Rep. 28 (1), 81–93.

Lazarowski, L., Krichbaum, S., DeGreeff, L.E., Simon, A., Singletary, M., Angle, C., Waggoner, L.P., 2020. Methodological considerations in canine olfactory detection research. Front. Vet. Sci. 7, 408.

Lesniak, A., Walczak, M., Jezierski, T., Sacharczuk, M., Gawkowski, M., Jaszczak, K., 2008. Canine olfactory receptor gene polymorphism and its relation to odor detection performance by sniffer dogs. J. Hered. 99 (5), 518–527.

Lit, L., Schweitzer, J.B., Oberbauer, A.M., 2011. Handler beliefs affect scent detection dog outcomes. Anim. Cogn. 14 (3), 387–394.

Maginnity, M.E., Grace, R.C., 2014. Visual perspective taking by dogs (*Canis familiaris*) in a Guesser–Knower task: evidence for a canine theory of mind? Anim. Cogn. 17 (6), 1375–1392.

Marchal, S., Bregeras, O., Puaux, D., Gervais, R., Ferry, B., 2016. Rigorous training of dogs leads to high accuracy in human scent matching-to-sample performance. Plos One 11 (2), e0146963.

Miklösi, Á., Polgárdi, R., Topál, J., Csányi, V., 1998. Use of experimenter-given cues in dogs. Anim. Cogn. 1 (2), 113–121.

Pemberton, N., 2013. The bloodhound's nose knows? dogs and detection in Anglo-American culture. Endeavour 37 (4), 196–208.

Pilley, J.W., Reid, A.K., 2011. Border collie comprehends object names as verbal referents. Behav. Processes 86 (2), 184–195.

Porritt, F., Shapiro, M., Waggoner, P., Mitchell, E., Thomson, T., Nicklin, S., Kacelnik, A., 2015. Performance decline by search dogs in repetitive tasks, and mitigation strategies. Appl. Anim. Behav. Sci. 166, 112–122.

Power, M., Fell, G., Wright, M., 2013. Principles for high-quality, high-value testing. BMJ Evid. Based Med. 18, 5–10.

Prada, P., 2010. Evaluation of contact and non-contact trapping efficiencies of human scent chemical profiles and their stability under different environmental conditions. PhD thesis. Florida International University.

Romanes, G.J., 1887. Experiments on the sense of smell in dogs. Zool. J. Linn. Soc. 20 (117), 65–70.

Rust, L., Nizio, K.D., Forbes, S.L., 2016. The influence of ageing and surface type on the odour profile of blood-detection dog training aids. Anal. Bioanal. Chem. 408 (23), 6349–6360.

Schoon, G.A.A., 1997. The performance of dogs in identifying humans by scent. PhD thesis. Leiden University.

Schoon, A., Haak, R., 2002. K9 Suspect Discrimination. Detselig Enterprises, Calgary.
Schoon, G.A.A., Curran, A.M., Furton, K.G., 2009. Odor and biometrics. In: Li, S.Z., Jain, A.K. (Eds.), Encyclopedia on Biometric Recognition. Springer-Verlag, Heidelberg, Germany, pp. 1003–1008.
Smeets, M.A., Rosing, E.A., Jacobs, D.M., van Velzen, E., Koek, J.H., Blonk, C., Semin, G.R., 2020. Chemical fingerprints of emotional body odor. Metabolites 10 (3), 84.
Steen, J.B., Mohus, I., Kvesetberg, T., Walløe, L., 1996. Olfaction in bird dogs during hunting. Acta Physiol. Scand. 157 (1), 115–119.
SWGDOG. https://ifri.fiu.edu/research/detector-dogs/swgdog/). 2022
Syrotuck, W.G., 1972. Scent and the scenting dog. Arner publications, New York.
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. Appl. Anim. Behav. Sci. 247, 105568.
Williams, M., Johnston, J.M., 2002. Training and maintaining the performance of dogs (*Canis familiaris*) on an increasing number of odor discriminations in a controlled setting. Appl. Anim. Behav. Sci. 78 (1), 55–65.
Woidtke, L., Dreßler, J., Babian, C., 2018. Individual human scent as a forensic identifier using man trailing. Forensic Sci. Int. 282, 111–121.
Wolf, A., 2016. Untersuchung des Einflusses der Alterung menschlicher Geruchsspuren auf die Ausarbeitung der Fährten durch Personensuchhunde. Dissertation Tierärztliche Hochschule Hannover.