JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, July 2011, Vol. 56, No. 4
doi: 10.1111/j.1556-4029.2011.01785.x
Available online at: onlinelibrary.wiley.com

TECHNICAL NOTE

GENERAL

Mary E. Cablk,[1] *Ph.D. and John C. Sagebiel,*[2] *Ph.D.*

Field Capability of Dogs to Locate Individual Human Teeth*

ABSTRACT: Avulsed teeth can be difficult if not impossible to recover in the outdoor environment, yet are important for victim identification. This study assessed dog teams as a resource to locate human teeth in a field setting and related performance in training with field capability. Standardized, objective training data were recorded and analyzed followed by double-blind capability trials. In the double-blind trials, 10 teeth were placed in each of six (10 m^2) plots. Search time per plot ranged from 27 to 50 min, and the proportion of teeth found by the teams varied between 0.20 and 0.79. Using 0.45 m as a distance criterion for a "find," the proportion of false positives ranged between 0.07 and 0.75. Results show that dog teams are capable of recovering individual human teeth in the field setting with high precision although capability varies. Training records support a team's expected field performance. Additional studies are needed.

KEYWORDS: forensic science, forensic dentistry, forensic anthropology, human remains detection, canine reliability, detection

The importance of teeth in forensic investigation is well documented. Teeth can be used collectively for positive victim identification through dental records as well as individually through DNA extraction (1). Postmortem, there are circumstances in which the teeth may avulse from the bones through natural processes or anthropogenic forces to the maxillo-mandibular region (2). Once separated, such as through a physical blow, animal ingestion, or disarticulation by animals, teeth can be difficult or impossible to recover (3). An individual tooth is relatively small and as such is difficult to see in the natural environment where it may quickly be covered by leaf litter, dirt, and other elements. Delattre (2) identified that after only a few days *in situ,* leaf litter and debris become intermixed with and even attached to decomposing human remains, requiring additional cleaning efforts for teeth. The potential importance of recovering as many teeth as possible should not be understated. In fact, the necessity to develop new techniques that can be used to locate human remains and specifically teeth is of such importance that the Department of Justice issued a call for proposals specifically toward this purpose in 2007 (Grants.gov funding opportunity no. 2008-NIJ-1698). One technique that is used in forensic investigations but which has not been studied is the use of dogs trained to find human teeth.

The use of dogs, referred to as "human remains detection (HRD) dogs," to locate human remains is recognized among law enforcement and the military, yet only a few scientific studies have been published in peer-reviewed literature. HRD dog handlers usually are volunteers associated with a Sheriff Offices' search and rescue unit, professional law enforcement officers, or military personnel. Volunteers comprise the majority of handlers in the United States, and because there are no universally accepted training or testing standards, capabilities vary tremendously between individuals, by team, and by agency affiliation. This variability results in part from the fact that the underlying science of canine olfaction applied to HRD, from describing odor signatures to measuring detection thresholds, remains unsolved (4,5). Similarly, standard terminology does not exist within and among the HRD dog team community. Because terminology varies, we present the following definitions. An HRD dog team is comprised of one dog and one handler. The term "alert" is based on Cablk and Heaton (6), which refers to the trained behavior the dog performs upon locating its target. A "passive alert" includes inactive behaviors such as a sit or a down; an "aggressive alert" includes active behaviors such as scratching or digging. The term "cross-trained" indicates having been trained for more than one target odor, such as both live and deceased humans (7). A defined search area in training is referred to as a "problem," and training problems can be "blind" or "known." A blind training problem is narrowly defined: the handler knows neither the number nor the location of the targets. In a known problem, the handler knows how many targets were placed and/or the target locations.

Dogs deployed to search crime scenes ideally are trained with a passive alert to minimize disruption to the scene, maximize preservation of evidence, and provide the greatest opportunity for pinpointing evidence. As such, an aggressive alert such as digging is unacceptable and of course urinating or defecating on a target, or picking up or ingesting a target is never acceptable. Working the dog on-leash, also the ideal, contributes toward scene preservation and optimizing area coverage. The dog is easily controlled when on-leash and works within close range of the handler. The exact coverage by the dog is known because it is always working within a set maximum distance from the handler, i.e., the length of the leash. Working a grid pattern minimizes disturbance while

[1]Division of Earth and Ecosystem Sciences, Desert Research Institute—DEES, 2215 Raggio Parkway, Reno, NV.

[2]Department of Environmental Health and Safety, University of Nevada, Reno, NV.

*Funded by the Desert Research Institute's competitive Institutional Project Assignment fund.

Received 21 May 2009; and in revised form 7 Mar. 2010; accepted 13 Mar. 2010.

maximizing coverage. Vegetation and terrain are not always conducive to working a dog on a leash, so the ability to deploy a dog off-leash and maintain detailed, directed pressure searching, where the dog's nose is at or near flush with the surface and focused on detection, is also an important team skill. Regardless of whether the dog is worked on- or off-leash, it is the handler's responsibility to ensure that the dog's nose is optimized to pass through the dog's minimum detection threshold, thus maximizing the likelihood that a target is found. The dog and handler are thus a team, and capability is assigned to the team, not just the dog.

Few published studies have evaluated HRD dog team capability. Komar (8) conducted experiments to quantify the ability of dog teams to locate simulated disarticulated human remains in Alberta, Canada. Lasseter et al. (9) quantified capability to locate simulated human burials in the southeastern United States. An early study by France et al. (10) evaluated capability for finding burials but used dead pigs as an analog to humans. Because of certain physiological similarities between pigs and humans, it was then believed that the two were analogs from an olfactory perspective. Recent work by Vass et al. (5) shows there are significant differences in vapor constituents between pigs and humans. Dog alerts on animal remains are generally regarded as undesirable by law enforcement. While Oesterhelweg et al. (4) utilized a rigorous scientific methodology to quantify dog capabilities to detect residual odor of recently and intact decedents on carpet, expected field performance of detection dogs outside a laboratory setting would be lower (11,12).

Of critical importance to a dog team's ability to locate their target, regardless of target, is training. Along with the variability in certification standards for HRD dog teams is the recording aspect of training activities. Handler training records are typically written in narrative format and as such do not lend themselves to quantification or analysis. We required handlers to maintain complete and objective training records for all training activities within a set time period before participating in trials.

This study sought to quantitatively capture a current state of practice of HRD dog team capability both in training and in simulated searches for individual human teeth in the field setting. The null hypotheses were (i) dog teams would not be capable of locating individual human teeth in the natural environment and (ii) performance in training would have no relationship with performance in the field. To address these hypotheses our objectives were to (i) quantify the capability of dog teams at locating individual teeth in the field setting and (ii) quantify the role of training relative to field performance.

Methods

The study consisted of two phases: data collection on training activities and experimental trials. All data were collected on three HRD dog teams that participated in both the training and the trial phases. Dog teams were selected for this study based on the following criteria:

- the team is certified for HRD by a recognized state or county agency;
- the dog is not cross-trained on live human odor;
- the dog has a passive alert;
- the handler is comfortable working the dog on-leash;
- the handler is willing to work within the parameters of a scientific experiment, including recording and submitting all training data prior to participation in the research trials.

Dog "A" is a 2-year-old German Shepherd Dog (GSD) certified for 1 year prior to participation in the trials. Dog "B" is a 7-year-old GSD certified for 7 months prior to the trials. Dog "C" is a 6-year-old Labrador retriever certified for 3 years prior to participation. All three dog teams were currently certified to the State of California Emergency Management Agency type 1 "cadaver dog," had either a sit or a down alert, and were not cross-trained on live people. Handlers, all females, maintained training logs prior to participation in the study. Each handler had trained and certified at least two dogs in HRD prior to participating in the study and had a minimum of 9 years of experience as a dog handler. Dogs A and B were acquired by their handlers at 8 and 10 weeks old, respectively, and began training at this age. Dog C was purchased as a partially trained 2-year-old dog, and its handler was the only owner of the dog after purchase. The handler for dog C completed the dog's training and certification.

This research was conducted under the University of Nevada Reno Animal Care and Use Protocol #00362.

Preparatory Training

Dog team training data used in this study were recorded beginning in August 2008. Handlers were instructed to continue their normal training regime and were advised that the research trials would focus only on teeth as targets. Handlers conducted training on dates and times of their own schedule without a required minimum number of trainings or targets. Allowing handlers to determine their own training regime mimics the current state of practice where there is no universally accepted training prescription for dog teams. As such, the capability of the dog teams quantified during the trials represents a realistic, albeit small, sample of the response one might expect when requesting an HRD dog team resource.

Data were recorded for each training session and also on each target, whether found or missed. Training logs included the following information about the training session: date, dog name, handler initials, number of targets placed, sequentially numbered search area, the date the problem was set, the time the targets were placed, and the time the problem was worked. Data were then recorded for each target find in the sequential order it was located by the dog team, not in the order placed. Data for each find included the general type of target (e.g., teeth, blood, etc.), whether the target was known or blind to the handler, whether or not the target was found, if found whether or not the dog performed its trained alert, if the dog alerted whether it was an independent alert or cued by the handler, whether or not the target was visible to the handler, the type of reward if any to the dog, target configuration (surface, buried, or elevated), and how the target was validated (e.g., visible to handler, set by assistant, etc.). Handlers also recorded the number of false positives during the training problem.

Training data were entered into a spreadsheet for analysis. The following basic summaries were calculated for each dog team for all target types and for teeth only: total available targets, total target encounters, total targets missed, percent recovery (proportion of targets found), proportion of known and blind problems worked, respectively, number of different training areas, number of training dates, and mean number of training problems per training date. Pearson's chi-squared test was used to assess significant differences in the proportion of blind problems versus known problems. False positives were reported both as a count and calculated proportion.

Calendar dates were converted to sequential training days. While in some instances handlers trained on the same date, the dates do not necessarily correspond to the same sequential training days between dog teams. However, the last training day for each dog team was the same—25 October, 6 days prior to the trials. Cumulative recovery rate was calculated across blind training problems and plotted against sequential training day. Cumulative recovery rate was calculated as

$$CR_T = \frac{\sum_{T=1}^{T} X_T}{\sum_{T=1}^{T} Y_T}$$

where CR_T is the cumulative rate of recovery at training day T, X_T is the number of targets found, and Y_T is the number of targets available at training day T, respectively. Cumulative recovery rate is a metric of the course of the dog's training, whereas the recovery rate (number found divided by number available to be found) per training session indicates the dog's performance only on that particular day. The cumulative recovery plot data are independent of number of days between trainings and specific dates of training.

Study Design—Experimental Trials

Experimental trials were held in western Washoe County, Nevada. A total of six square plots 10 m × 10 m in size were delineated in a flat area with a Jeffrey pine (*Pinus jeffreyi*) overstory and shrub/scrub understory of mountain mahogany (*Cercocarpus ledifolius*), sage brush (*Artemisia tridentata*), and native grasses. The plots were located in an area that had not been used for training prior to the trials. Teeth were provided by private individuals and were soaked in distilled water in a sterilized glass jar with a metal lid for 24 h followed by air drying in sunshine. Teeth with extensive dental work i.e., where the majority of the tooth appeared to be artificial, were excluded. The teeth were not placed in preservative or otherwise chemically treated. All teeth were then stored together in a single sterilized container. A total of 10 teeth were placed in each (10 m^2) plot on October 29, 2008. Each plot was delineated into 1 m^2 subgrids using a Cartesian coordinate system (*x*, *y*), and tooth placement was selected using random number–generated subgrid locations corresponding to each of the grid squares. Teeth were placed such that no more than one tooth was within a subgrid and if adjacent subgrids had a placement, the teeth could not be closer together than 0.90 m. Teeth were placed either partially buried with one end of the tooth flush with the ground surface so as not to be visible or at ground surface under pine litter. Each placement was photographed and cataloged by its unique random grid location to assist with recovery post-trial.

Eight additional teeth were also placed on the ground surface near the plots so that handlers had an opportunity to reward their dogs on known targets during the trials but outside of their search plots. The trials were double blind, and neither handler nor data collector knew how many or the locations of placements in the plots. If a handler was able to visually confirm a tooth in a plot once located by her dog, she could also reinforce the alert using those known (found) targets.

The trials were conducted on October 31, 2008. Each dog team was randomly assigned two of the six plots each. Only one dog team worked a plot, delineated with pin-flags at each of the four corners. The handlers were instructed to state out loud for data recording purposes when they were starting their search, when they were stopping their search, when the dog alerted, and when their search was completed. They were instructed to search their dogs on-leash using a two-pass grid strategy (Fig. 1) where the two grids ran perpendicular to each other. Handlers were advised that they had no time limit to complete the search of each plot. Start time and stop time for each grid pass was recorded. Handlers were

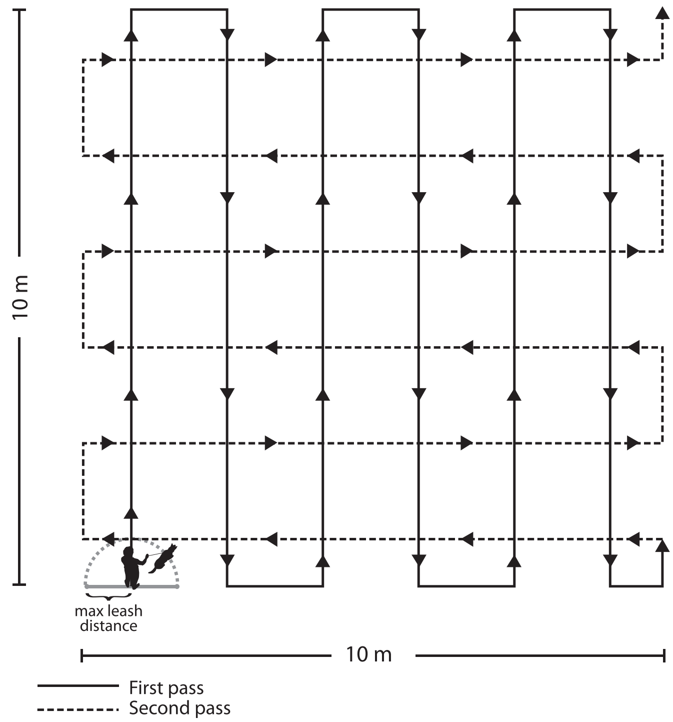

FIG. 1—*The two-pass grid strategy involves working the dog on-leash using a grid pattern to cover the search area. The first and second passes are worked perpendicular to each other to maximize coverage in an efficient manner.*

allowed to take breaks for any purpose, and start and stop times surrounding breaks were recorded so that both total time and actual work time could be calculated. Total time includes breaks, if any, while "work time" is the amount of time the dog actually spent searching within a plot calculated as total time minus total break time.

Dogs worked on-leash with detailed and directed pressure searching. When a dog alerted, as determined by the handler stating "I have an alert," the following data were recorded: time, sequential alert number, recorder's initials, plot number, search area number (1st or 2nd), team name, whether or not the dog performed its trained alert and if so whether it was dependent (cued by handler) or independent, target validation by handler (either visual or none), and reward type if any. Weather conditions were monitored via a local station which is part of the Remote Automated Weather Station network. This station was *c.* 1.9 km (1.18 miles) from and 90 m (*c.* 295 ft) lower elevation than the study site.

When the handler called an alert she placed a flag numbered sequentially at the location where the dog alerted. If the handler observed a tooth when placing the flag, she had the option of rewarding the dog with play, providing an intermediate level of reward such as verbal praise or petting, or offering no reward to the dog. The handler also had the option of determining that a tooth was likely present without the dog alerting based on the dog's change of behavior. In such an instance, the handler placed the flag in a location that she deemed was the highest probability location.

When the search of all plots was completed, the flags indicating potential tooth locations, either by dog alert or handler call, were mapped and teeth were recovered. Teeth had to be both placed and recovered to be considered "available" to be found. The distance between each flag and the nearest tooth were measured and recorded. The initial criterion for a handler's flag to be considered a "hit" was no farther than 0.45 m (*c.* 18″) from a tooth. Recovery was calculated as the number of teeth flagged by the handler and within the distance criterion divided by the total number of teeth recovered from the plot. This calculation was used to address the first hypothesis that dog teams would not be capable of locating individual human teeth in the natural environment. False positives were calculated as the proportion of alerts >0.45 m from a target divided by all alerts.

A binomial test for proportions was used to compare performance between the dog teams in training and trials where the null hypothesis was equal probabilities. This was a statistical means to address the hypothesis that training, calculated as recovery in blind problems, would have no relationship with performance in the field, calculated as recovery in the double-blind trials. Other analyses relevant to establish similarities and differences in the dog team pretrial training were conducted. These included comparing the proportion of blind versus known problems. Pearson's correlation was used to analyze the relationship between work time in a plot and recovery of teeth, to assess the possibility that a team could locate more teeth simply by spending more time searching. Because of the small sample sizes, a significance level of 0.01 was used.

Results

Pretrial Training

Table 1 shows the training data recorded and summarized for each dog team for both known and blind training problems for all target types within the human remains class which include but are not limited to blood and bone, as well as for teeth. Although there were differences in the number of targets encountered by each dog, at the end of the training time period, each dog team had >0.90 recovery for known targets and between 0.54 and 0.81 overall target recovery in blind problems and slightly lower recovery specifically for teeth. The number (proportion) of false positives recorded over the training period was 13 (0.09), 45 (0.22), and 3 (0.07) for dog team A, B, and C, respectively. The proportions of blind training problems between dog teams were not statistically significantly different ($p = 0.0135$ between dog team A and C); however, dog team A worked significantly more blind problems on teeth than either dog team B ($p = 0.0062$) or dog team C ($p = 0.0048$). For all target types, all dog teams worked more known than blind problems and more than half of the blind targets placed in each dog team's training were teeth.

Figure 2 shows the cumulative recovery rates for blind teeth problems during the pretrial training period. Each dog team experienced a decrease in recovery rate early in the series of working blind problems. Immediately prior to participating in the trials, the cumulative recovery rate for blind teeth problems for each of dog teams A, B, and C was 0.81, 0.67, and 0.65 respectively. Figure 3 shows the individual recovery rates by dog team for each of the sequential training days for blind trainings on teeth. The last seven trainings by dog team A yielded recoveries of 1.0, where all of the targets were found. Dog team B's recovery rate on the last training was 0.85 and Dog team C had recovery of 1.0 on trainings 4–6, while the last day of training dropped to 0.67.

Capability Trials

Table 2 presents a summary of the capability results from the experimental trials. Per plot dog team A had the highest recovery, followed by dog team B, and then dog team C. Within a plot, the highest recovery was seven of nine teeth and the lowest recovery was 0 of 10 teeth based on the maximum distance threshold of 0.45 m to be considered a "hit." Proportion of false positives for each of dog teams A, B, and C was 0.07, 0.27, and 0.75, respectively. Pearson's correlation between the work time in a plot and recovery was not statistically significant (DF = 4, $p = 0.0573$).

A total of five teeth were not recovered using both dogs and human methods and were presumed taken by animals. One tooth in each of three different plots and two reinforcement/reward teeth were not located over 2 days of search efforts totaling more than 5 h. A placement error resulted in only nine teeth being set in a fourth plot. Therefore, there were four trial plots that had a total of nine teeth.

Table 3 shows the length of time each dog team spent on each of their two plots. Dog teams A and B spent a similar amount of time working each of the two passes in each of the plots. The mean total time to complete a two-pass search of a (10 m^2) plot was dog team A, 45 min; dog team B, 47.5 min; and dog team C, 27.5 min.

Of the 29 teeth located by the dog teams, 24 were pinpointed precisely, while five were within 0.45 m of a tooth. Two flags were placed between 0.45 and 1 m from a target, and seven flags were placed within 2 m of a tooth. During the period the targets were out, the maximum temperature was 23.3°C (74°F) and the minimum was 6.1°C (43°F). Wind speed averaged 7.0 m/s (15.7 mph) the day of the trials.

TABLE 1—*This table summarizes the pretrial training for each dog team for all target types within the class of human remains (top) and for teeth specifically (bottom). The number of targets set and the number of targets found by the dog team ("finds") are reported by whether they were "known" or "blind." Recovery is the number found divided by number available to be found and is reported as overall ("all"), for blind targets and for known targets. The target proportions are the total number of targets worked "blind" or "known" divided by all targets worked.*

	All Target Types			Finds			Recovery			Target Proportions	
Dog	Known	Blind	Total	All	Blind	Known	All	Blind	Known	Blind	Known
A	84	69	153	138	56	82	0.90	0.81	0.98	0.45	0.55
B	115	75	190	162	50	112	0.85	0.67	0.97	0.39	0.61
C	60	28	88	70	15	55	0.80	0.54	0.92	0.32	0.68
	Teeth Only			Finds			Recovery			Target Proportions	
Dog	Known	Blind	Total	All	Blind	Known	All	Blind	Known	Blind	Known
A	45	50	95	81	37	44	0.85	0.74	0.98	0.53	0.47
B	77	53	130	105	31	74	0.81	0.58	0.96	0.41	0.59
C	44	24	68	51	11	40	0.75	0.46	0.91	0.35	0.65

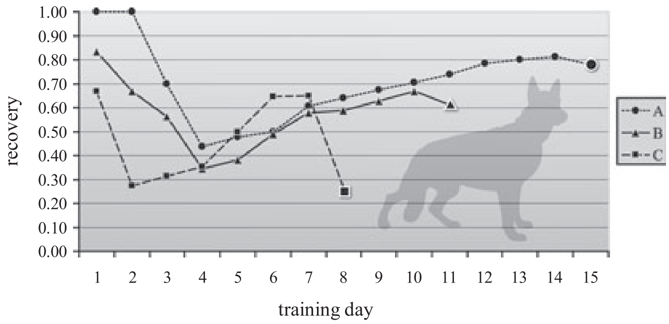

FIG. 2—*Cumulative recovery for blind problems during pretrial training for dog teams A, B, and C. The training day shown on the x-axis is sequential. Sequential training day numbers for each team may or may not coincide by calendar date. The highlighted (last) point in the plot is each team's recovery rate during the experimental trials. Two teams performed as might have been expected compared with previous training. One team performed at much lower capability than would have been expected based on prior capability in training.*

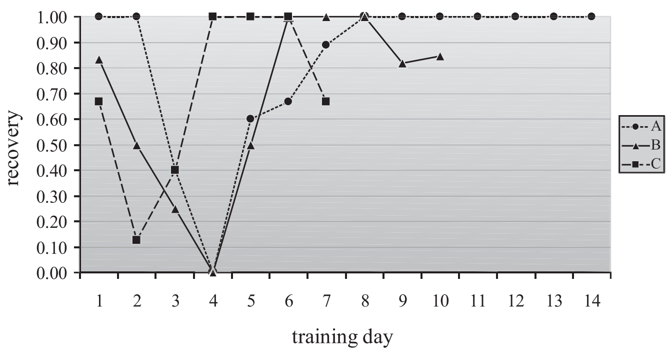

FIG. 3—*Individual team recovery rates for teeth during pretrial trainings are plotted for sequential training dates. The data presented here are for blind teeth problems. All teams experienced a large drop in capability before experiencing a subsequent large increase in recovery rate.*

TABLE 2—*Results from the double-blind capability trials. "Found (available)" is the number of teeth found by the dog team and the number that were available to be found within a plot. Although 10 teeth total were placed in each plot, unexpected circumstances such as animal theft reduced the number of teeth available to be found. Recovery was determined by the target being no more than 0.45 m from a flag. Total recovery is calculated over both plots. "Miss" is calculated as 1 minus the recovery. p-Values are from the binomial test for proportions comparing recovery in trials against the hypothesized proportion, recovery in training. False positives (false+) were determined by the number of alerts the dog performed, and the nearest target was >0.45 m.*

	Plot 1	Plot 2	Total			
Team	Found (Available)	Found (Available)	Recovery	Miss	*p*-Value	False+
A	7 (9)	7 (9)	0.78	0.22	0.76	1
B	5 (9)	6 (9)	0.61	0.39	0.62	4
C	4 (10)	0 (10)	0.20	0.80	0.0001	3

Comparison of Training with Capability in the Field

Results from the binomial tests for proportion showed that for all target types, neither dog team A ($p = 0.76$) nor dog team B ($p = 0.62$) performed statistically significantly different in the trials than in blind training problems, while dog team C did show a statistically significant difference ($p = 0.0001$). For teeth only results were the same, with dog team A ($p = 0.50$) and dog team B ($p = 0.78$) having no difference in recovery between training and trials, whereas dog team C performed significantly differently in the trials than in training ($p = 0.004$).

TABLE 3—*Time spent by each team searching each of their two areas during the trials. Total time is calculated from the time the handler indicates they are beginning their search to the time they state they have completed their search. Work time is only the amount of time the team spent actually searching their plot and does not include breaks for water, etc.*

	Total Time (Minutes)			Work Time (Minutes)		
Team	Pass 1	Pass 2	Total	Pass 1	Pass 2	Total
Search area 1						
A	31	9	40	30	9	39
B	30	15	45	30	15	45
C	11	16	27	8	15	23
Search area 2						
A	41	9	50	38	9	47
B	32	18	50	23	16	39
C	5	23	28	5	17	22

Discussion

This study was conducted to examine the capability of HRD dog teams when called for a specific investigative purpose, to locate human teeth. We sought to capture an example of the current state of practice of HRD dog teams in terms of training, to quantify their utility as a law enforcement resource to assist in evidence recovery, and to establish whether or not there was a relationship between performance in training and actual searches. Although we concede the sample size is small, the methodology in this study can be used as a template from which to build a larger database to gain greater statistical power.

The first null hypothesis is rejected as dog teams were shown to be able to locate individual human teeth in the field environment. However, not all dog teams were equally capable. Performance in training on blind teeth problems was correlated with performance in the field for two of three dog teams, suggesting that training for blind problems may establish an expected performance level in the field. With this finding, the second null hypothesis is rejected. We cannot establish cause and effect to explain why dog team C had a significantly lower recovery in the trials than the other teams. Overall, this team had the least number of trainings, the fewest exposures to and reinforcements on targets, and entered the trials with the lowest recovery rate. What is important to note from this finding is the demonstration that significant variability exists in dog teams, even among those that have met the same minimum criteria for certification, alert type, work style, and trained target odors (i.e., HRD only). From a law enforcement perspective when an HRD dog team resource is requested, there is no way of knowing which of these three dog teams will arrive. None of the teams performed better in the trials, which mimicked an actual search, than they did in training.

The relationship between training and performance in trials is relevant in the context of estimating the usefulness of a particular dog team to assist in evidence recovery. Dog teams A and B had recoveries in the trials that were slightly less than their respective cumulative recovery rate for the last training. The relationship between cumulative recovery and trial capability is illustrated in Fig. 2, which depicts each dog team's recovery during the trial plotted on the graph of pretrial training. The fact that for two dog teams, training and field performances were correlated is not surprising if one views a trial, or an actual search, for what it is—essentially a blind or double-blind problem. In this light, it might be expected that dog teams would perform at about the same level in the trials as they did in their last blind training problem but based on the data presented in Table 3 and Fig. 3, this was not the case. For example dog team A recorded 1.0 recovery (i.e., 100%) for the last seven blind problem teeth trainings but did not subsequently perform perfectly in the trials.

Of interest is that in the natural setting teeth may or may not have associated blood or decomposition of other tissues, depending on the circumstances surrounding death, postmortem processes, and time. Despite the relatively clean state of the teeth used in this study, dog teams were able to find and in many instances pinpoint the location of individual teeth to which they had not been previously exposed.

Plotting cumulative recovery provides the opportunity to identify directional trends in the dog team's training. The overall recovery statistic, as a proportion of total number of teeth found of those placed, provides a discrete measure that can be used to test for significant differences in performance between training and trials; however, it does not display trends. Combined, these data may provide a fairly realistic picture of the dog team's capability. Maintaining accurate and quantifiable training records such as those kept in this study would provide for on-going assessment and the ability for handlers to more accurately present probability of detection (POD) to search managers, detectives, and investigators.

False positives are an important aspect to consider with any tool's usefulness regardless of the technology. From a scientific perspective, false positives are challenging, in that, in a research study, a rule base is necessary to determine what constitutes a "hit" and what is deemed a "miss" for quantitative reasons. In the real-world setting, false positives may be overlooked when coupled with one or more alerts that result in a find, particularly when the goal is simply to locate a particular piece of evidence. As an example, many alerts lead to nothing, but one alert produces the murder weapon. Regardless of how many alerts were offered, the end goal of recovery was met. In reality, the relationship between the distances of an actual tooth from the dog's alert, or handler's flag, translates to processing effort. The farther away from a target the dog alerts and a flag is planted, the more time and effort required to find that target. This holds true for any target. Processing effort translates into time and cost. Most of the flags were planted very close to the actual teeth, <0.45 m (18″), which would make locating the tooth relatively easy with a sieve. False positives and false negatives are expected for any detection dog team, as no team is perfect. Minimizing both of these errors reduces the effort required for recovery and increases the chance that evidence is recovered. In this study, the false-positive responses of dogs were variable, both between dog teams and from training to the trials. Because there was a rule base established for what constituted a false positive based on alert distance from target, it is important to recognize that these false positives do not necessarily translate to the equivalent of a dog alerting in a blank area.

The results suggest that it is the number of days of training and the actual number of target encounters that is correlated with capability. The more often a dog is reinforced for alerting on target, the higher its capability under the pressure of an actual search. There were differences in the amount of time the dog teams spent searching the plots with the two teams that worked through the plots more slowly having higher success at recovery, although the time variable was not statistically significantly correlated with number of teeth located.

Training is fundamental to capability, and additional quantitative investigation with larger sample sizes will contribute toward a broader understanding of the range of capability and variability that HRD dog teams have. In the interim, the study results showing the relationship between training and field capability may be useful to handlers who work such assignments and for law enforcement. Certainly training specifically for a particular target to a known capability can build handler confidence and provide a means for better POD estimation in the field. From a practical perspective, these results might serve as a talking point between individuals requesting HRD dog team resources and the dog teams themselves. Requesting agencies may find value in working with their local HRD dog teams to better understand their capability or to outline specific requirements when requesting dog teams (i.e., have trained specifically on teeth or burned remains, for example, perform a passive alert, etc.).

Finally, using dog teams certified to an accepted, recognized standard was important from a research perspective, in that, the tool being measured demonstrated a minimum level of proficiency via testing to the same standard. Conducting capability studies whether for dog teams or gas chromatographs, for example, requires an established initial specification to be able to generate inferences from the resulting data. The idea of minimum level of proficiency, or baseline specification, is also important for actual crime scene deployments where the legal defensibility tied to any team's qualifications is a critical factor. The criteria used to select dog teams, which did not include a prescribed training requirement but did include certification to the same standard, yielded dog teams that had variable recovery rates in the field. Just because a dog team was "certified" did not mean that the team would perform equally well as another similarly certified team. Therefore, the criteria for selecting dog teams when ordering a resource to conduct a particular search, such as to locate teeth may merit additional requirements, such as demonstrated capability in training or demonstrated proficiency to a certification standard that more closely reflects the specific assignment. This finding also reinforces the need for handlers to conduct training that mimics actual search assignments and conditions and to train toward maximizing recovery through the use of blind problems. Because training was documented but not dictated, we cannot offer a prescription for proportion of blind versus known training problems to maximize team performance.

Conclusions

HRD dog teams can be an effective and efficient tool for locating individual human teeth in the field setting. Dog teams were able to not only pinpoint teeth, but teeth to which they had not been previously exposed prior to the experimental field trials. Capability varied although the dog teams in this study had all certified to the same standard prior to participating in the study. The results suggest that this might be related to both the type and amount of training done prior to the field trials. Working blind problems where both the number and location of sources is unknown to the handler is suggested to be an important means to develop the skills of the dog team. Because actual searches are essentially blind problems to the handler, and typically double-blind, conducting training and testing in this manner presents a realistic preparatory approach. As such, individual dog team qualifications are important when selecting dog teams. While the results suggest that a dog team's recovery rate in training, calculated as success during blind problems, may provide a good estimation of capability on actual search deployments, additional studies with larger datasets are needed to substantiate this with greater statistical power.

Acknowledgments

The authors wish to thank Susan Clark, Ph.D., of Dynamic Competence and Cindee Valentin of Applegate School for Dogs who provided assistance with the experimental trials. Lisa Wable contributed technical assistance with graphics preparation. Sara Marcus provided technical review assistance.

References

1. Bell GL. Dentistry's role in the resolution of missing and unidentified persons cases. Dent Clin North Am 2001;45(2):293–307.
2. Delattre VF. Forensic dental identifications in the greater Houston area. J Forensic Sci 2001;46(6):1379–84.
3. Haglund WD. Scattered skeletal human remains: search strategy considerations for locating missing teeth. In: Haglund WD, Sorg MH, editors. Forensic taphonomy: the postmortem fate of human remains. Boca Raton, FL: CRC Press, 1997;383–94.
4. Oesterhelweg L, Kröber S, Rottmann K, Willhöft J, Braun C, Thies N, et al. Cadaver dogs—a study on detection of contaminated carpet squares. Forensic Sci Int 2008;174:35–9.
5. Vass AA, Smith RR, Thompson CV, Burnett MN, Dulgerian N, Eckenrode BA. Odor analysis of decomposing buried human remains. J Forensic Sci 2008;53(2):384–91.
6. Cablk ME, Heaton JS. Efficacy and reliability of dogs in surveying for desert tortoise (*Gopherus agassizii*). Ecol Appl 2006;16(5):1926–35.
7. Lit L, Crawford CA. Effects of training paradigms on search dog performance. Appl Anim Behav Sci 2006;4:277–92.
8. Komar D. The use of cadaver dogs in locating scattered, scavenged human remains: preliminary field test results. J Forensic Sci 1999;44(2): 405–8.
9. Lasseter AE, Jacobi KP, Farley R, Hensel L. Cadaver dog and handler team capabilities in the recovery of buried human remains in the Southeastern United States. J Forensic Sci 2003;48(3):617–21.
10. France DL, Griffin TJ, Swanburg JG, Lindemann JW, Davenport GC, Trammell V, et al. A multidisciplinary approach to the detection of clandestine graves. J Forensic Sci 1992;37(6):1445–58.
11. Nussear KE, Esque TC, Heaton JS, Cablk ME, Drake KK, Valentin C, et al. Are wildlife detector dogs or people better at finding desert tortoises (*Gopherus agassizii*)? Herpetol Conserv Biol 2008;3(1):103–15.
12. Wallner W, Ellis T. Olfactory detection of Gypsy Moth pheromone and egg masses by domestic canines. Environ Entomol 1976;5:183–6.

Additional information and reprint requests:
Mary E. Cablk, Ph.D.
Desert Research Institute
2215 Raggio Parkway
Reno, NV 89512
E-mail: mary.cablk@dri.edu