Title
The Steel Helmet Project: Canine Olfactory Detection of Low Concentrations of a Surrogate Chemical Warfare Agent

Journal
International Journal of Comparative Psychology, 16(4)

ISSN
0889-3675

Author
Hilliard, Stewart

Publication Date
2003-12-31

DOI
10.46867/ijcp.2003.16.04.04

International Journal of Comparative Psychology, 2003, **16**, 193-208.

The Steel Helmet Project: Canine Olfactory Detection of Low Concentrations of a Surrogate Chemical Warfare Agent

Stewart Hilliard
Department of Defense Military Working Dog Veterinary Service, U.S.A.

The Steel Helmet project was meant to assess the feasibility of the chemical warfare agent (CWA) detector dog concept. A relatively benign organophosphate pesticide called dichlorvos was used as a surrogate for CWAs. Using conventional training techniques, U.S. Department of Defense military working dogs were taught to discriminate scent boxes containing dichlorvos from “vehicle” scent boxes. Experiment 1 appeared to show that two out of three subjects were capable of criterion accuracy (0.95 or better) at the lowest test concentrations of dichlorvos— 3 and 1 parts per billion by volume (ppbv). An additional manipulation showed that, when differential contamination of the scent boxes in Experiment 1 was accounted for, all three subjects fell short of criterion accuracy when tested at 1 ppbv. The canine dichlorvos detection "threshold" was therefore estimated at equal to or less than 3 ppbv, but not so low as 1 ppbv. Experiment 2 demonstrated that detection responding was specifically controlled by dichlorvos, rather than concomitant odors, and that the subjects were not merely reacting to the novelty or salience of dichlorvos vapor. The implications of these results for the feasibility of the CWA detector dog concept are discussed in terms of safe canine CWA exposure levels.

Although dogs have been commonly trained for law enforcement and security applications since the early 20th Century (Ceulebroeck, 1983; de Caluwé, 1995; Most, 1951; Stephanitz, 1923), they have in recent years assumed unprecedented importance because of the ability of trained dogs to detect and localize target substances and devices, especially explosives, arms, and land mines (Cain, Mason, & Morton, 1985; Chao, 1977; Home Office Standing Advisory Committee, 1996).

Detector dogs are trained to respond to target odors through associative conditioning, using appetitive reinforcers such as food and retrieve objects (e.g. rubber balls). Because the olfactory "signatures" of detector dog target substances

The views expressed in this article are those of the author and do not reflect the official policy or position of the Department of the Army, Department of Defense, or the U.S. Government. The Steel Helmet project was funded through Technical Support Working Group Task 154E. This report describes merely a part of an ambitious program of work spanning almost three years of dog training and experimentation. The effort required the cooperation and help of a large number of people, some of whom volunteered their time, and many of whom went above and beyond the call of their paid duties. We are grateful to the following parties: Dr. Jim Dykes and Dr. David Pillow, University of Texas at San Antonio; Robert Ehlers, Dr. Bill Barclay, and Russ Barajas, Southwest Research Institute; Russ Jones, Bryant Willett, Dawn Bayles, Danny Diller, J. Lacy Smith, and Diana Thomas of the Behavioral Medicine Section, Department of Defense Military Working Dog Veterinary Service; LT Keith S. Crim, United States Marine Corps, and SSGT Paul Schaaf, United States Air Force; and Dr. Paul Waggoner, Institute for Biological Detection Systems, Auburn University. This paper is dedicated to the memory of Dr. Paul Geno, Southwest Research Institute. Correspondence to Stewart Hilliard, DOD/MWD Veterinary Service, 1219 Knight Street, Lackland AFB, TX 78236-5519, U.S.A. (stewart.hilliard@lackland.af.mil).

such as smokeless gunpowder and nitroglycerine are highly complex (Williams et al., 1997), and because such odors normally cannot be presented in isolation from concomitant odors (such as the smell of the canvas bag containing the gunpowder, or the scent of the person who handled the canvas bag), the study of detector dog performance must include detailed consideration of stimulus control.

For instance, Williams et al. (1997) have shown that, when dogs are reinforced for instrumental responding to the odor of smokeless gunpowder, their detection performance does not necessarily come under the control of the active explosive ingredient, nitroglycerine. Rather, dogs learn to respond to the matrix in which the nitroglycerine is embedded. Furthermore, individual dogs exhibit different "odor detection signatures" when trained to respond to smokeless powder. That is, when the complex odor is broken down into constituent parts, different constituents will exert different degrees of stimulus control over responding in different animals.

The lore and science of detector dog training are filled with examples in which attempts were made to teach dogs to detect some substance by its characteristic odor, but unfortunately the animals learned to detect some odor associated with, rather than produced by, the substance. For instance, in a Defense Evaluation and Research Agency (DERA) study of the low-concentration TNT detection capabilities of British Army detector dogs, the animals displayed nearly perfect levels of proficiency on improbably (even incredibly) low titrations of a TNT solution. It was eventually discovered that, at low concentrations of TNT, the dogs had smoothly switched from solving the problem by responding to TNT odor to responding to the odor of trace amounts of pipette plastic that had been dissolved by the solvent used to titrate the TNT (Nicklin, 2000).

The Chemical Warfare Agent Detector Dog Concept

The real possibility of terrorist use of weapons of mass destruction, such as nerve gas, and the paucity of available countermeasures against such terrifying threats, have recently led to interest in the concept of a "chemical warfare agent detector dog." Chemical warfare agent (CWA) detector dogs might be deployed in a number of ways. They might be used prior to release, to “screen” critical locations for the presence of a CWA device. This application assumes that associated with such a device would be certain levels of CWA residue that had escaped from the device or that had contaminated it during manufacture. CWA dogs might also be utilized postrelease to delineate areas of contamination resulting from activation of a CWA device. Alternatively, CWA detector dogs might be used to assess whether specific locations or facilities had ever been used to manufacture or store CWAs.

The class of CWAs that merits the greatest concern is that of the organophosphate neurotoxins, such as GB (sarin, or Isopropyl methylphosphonofluoridate) and VX (0-ethyl S-[2-(diisopropylamino)ethyl] methylphosphonothiolate). In order for the concept of the CWA detector dog to be feasible, it would be necessary to demonstrate that dogs are capable of accurately detecting substances like GB and VX at concentrations below harmful levels. Most available standards for subharmful chemical exposure define it in terms of “occupational” exposure for workers in the chemical industry. For instance, the permissible Airborne Exposure

Limit (AEL) for GB is an 8-h time-weighted average of 0.0001 mg/m^3, or 0.01 parts per billion by volume (ppbv), while for VX the AEL is 0.00001 mg/m^3, or 0.001 ppbv (U.S. Army Soldier and Biological Chemical Command, 1996; 2001). These AEL's represent vapor concentrations of CWAs to which human beings can be exposed for 40 h per week throughout their working lives with little risk of injury.

To provide an initial proof of the CWA detector dog concept, a common organophosphate pesticide called dichlorvos (2, 2 – Dichloroethenyl phosphate) was used in place of GB and VX. Dichlorvos is an eminently suitable CWA surrogate because it exhibits volatility similar to that of GB and VX (Spectrum Laboratories, 1998), but it is low in toxicity when compared to CWAs. The Occupational Safety and Health Administration Permissible Exposure Limit (OSHA PEL) for dichlorvos is 1 mg/m^3, or 110 ppbv (Ehlers and Barclay, 2001).

The approach used in this exploratory research, called the Steel Helmet project, was to ascertain whether working dogs could be trained to detect concentrations of dichlorvos near the AEL's for GB and VX. Another way to put this is that the Steel Helmet project was meant to provide an estimate of the *canine detection threshold* for dichlorvos. The term threshold is here used loosely, as the lowest dichlorvos vapor concentration at which subjects could achieve a criterion *accuracy* of 0.95 (Accuracy = (Hit Rate + (1 – False Alarm Rate)) / 2).

The Steel Helmet project relied on a "semioperational" paradigm adapted from the standard method of training a United States Department of Defense (DoD) detector dog (Craig, 1989). In this paradigm, odors are presented to the subject in "scent boxes." These boxes are made of cardboard, with a hole cut in the top into which the dog is taught to insert its nose in order to sample the contents. The subject's task is to move along an array of these boxes, sampling each in turn. The detection response is sitting next to a box containing the target, and rejection of a box consists of passing it by without sitting. Although still not a completely realistic detector dog task (it lacks the localization component that is a crucial aspect of a detector dog's work), this protocol features a number of advantages. Because the dog must locomote from station to station this task is topographically related to an operational detector dog problem. In addition, because each stimulus sampling event is spatially discrete and marked by an overtly observable behavior (insertion of the nose into the box), it is possible to tally target and nontarget stimulus presentations and compute Hit and False Alarm rates. Finally, the scent boxes represent partially closed systems (with sharp gradients between the vapor concentrations inside vs. outside of the boxes) that lend themselves to sampling and quantification of intrascent box vapor concentrations by means of gas chromatography.

Experiment 1: Estimation of the Canine Dichlorvos Detection Threshold

Experiment 1 was designed to estimate the "canine detection threshold" for dichlorvos. Subjects were first trained to detect and indicate a "training aid" that provided the maximum safe concentration of dichlorvos vapor, and then their accuracy on this aid was tested. The subjects were subsequently trained and tested with three additional levels of training aid producing successively lower concentrations of dichlorvos vapor.

Method

Subjects. Three Belgian Malinois (Robby I, Robby II, and Ringo) and one Labrador retriever (Tara), aged 1.8, 1.9, 2, and 9.5 years, respectively, served in the study. Robby I and Robby II were naïve with respect to substance detection. Ringo had received introductory substance detection training using smokeless gunpowder. Tara was a certified Federal Aviation Administration Explosives Detector Dog that had been retired from duty for health reasons.

Materials. Dichlorvos odor was presented to the subjects by means of training aids prepared by Southwest Research Institute of San Antonio, Texas (Ehlers and Barclay, 2001). These aids consisted of various masses of neat dichlorvos pipetted onto a square pad of chromatography filter paper (Whatman 17Chr). The largest mass of dichlorvos used to prepare training aids was 70 mg, and successively "weaker" training aids were fashioned by employing 7 mg, 1.4 mg, and 0.50 mg. "Vehicle" aids were prepared by treating filter paper pads in an identical fashion, except that no dichlorvos was applied. Dichlorvos and vehicle aids were individually wrapped in aluminum foil, bundled according to type of aid, sealed into plastic bags, and stored in a freezer.

For training and testing, the dichlorvos and vehicle training aids (without aluminum foil) were placed inside plain cardboard boxes (30.5 X 30.5 X 23 cm). The boxes were assembled with masking tape that sealed all seams and edges, and each had a scent hole (from 10 to 14 cm in diameter) cut in the top. A single thumb tack was used to fix a training aid (one dichlorvos or vehicle aid per box) to the inside wall of each box near the top, so that the aid was not visible through the scent hole, and so that the side of the training aid onto which dichlorvos had been adsorbed was exposed. To prevent contamination with human scent, aids and scent boxes were handled with nitrile (Ansell Edmont) or plastic (Handgards, Inc.) gloves by grasping the rims of the scent holes, using the left hand to transport vehicle boxes and the right to transport dichlorvos boxes. Gloves were changed periodically during training and testing.

Scent boxes were placed in twelve stations arranged equidistantly (about 1.8 m apart) around the perimeter of the training room (6.7 X 7.3 m), by means of four wooden racks (with three stations per rack) resting on the floor against the walls of the training room. A variety of food, rubber balls, and other toy items were used as reward objects to reinforce the dogs for correct detections of dichlorvos training aids.

Instrumentation and Dichlorvos Monitoring. To measure the concentrations of dichlorvos vapor in ambient air and inside the scent boxes, two instrument systems were used. Constant, real-time monitoring was conducted using a small, automated gas chromatograph (Miniature Chemical Agent Monitoring System, or MINICAMS; CMS Field Products Group, Birmingham, AL). The MINICAMS was configured in a fixed-site installation, with a pump (Gast Mfg. Corp., Model 1531-107B-6557X) drawing samples of air through three heated sample lines (HSL-4XXX, CMS Field Products Group) that could be used to measure dichlorvos concentrations both in ambient air and inside the scent boxes.

To assess the MINICAMS' accuracy, technicians provided by Southwest Research Institute of San Antonio, Texas used polyurethane foam (PUF) cartridges and a personal monitoring pump to sample concurrently with the MINICAMS (Ehlers and Barclay, 2001). The PUF cartridges were subsequently extracted in hexane and analyzed using an Agilent 5973 gas chromatograph/mass spectrometer (GC/MS) at Southwest Research Institute.

Concentrations of dichlorvos vapor (as measured by MINICAMS during concurrent sampling with PUF-GC/MS) produced by 70, 7, 1.4, and .5 mg training aids spanned more than two orders of magnitude, from about 115 ppbv to about 1ppbv. Agreement between the two instrument systems at each level of training aid (see Figure 1) varied from excellent (within 6% deviation) to only fair (up to 100% deviation). At higher intrascent box concentrations produced by 70 mg and 7 mg aids, where safety of the dogs was a concern, MINICAMS consistently overestimated PUF-GC/MS (probably the more accurate of the two systems as they were set up and used here).

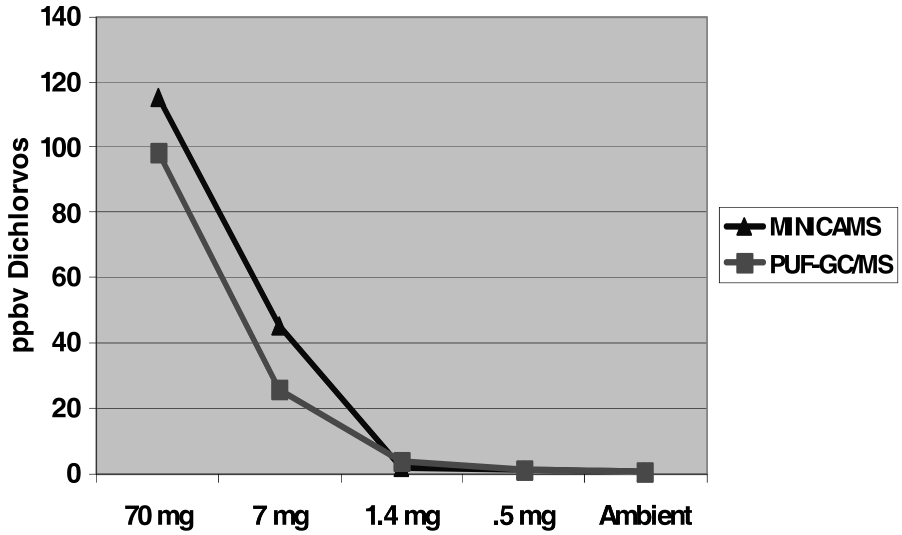

Figure 1. Average MINICAMS versus PUF-GC/MS dichlorvos readings across 70 mg, 7 mg, 1.4 mg, and 0.50 mg intrascent box concentrations. Also shown are the mean MINICAMS and PUF-GC/MS estimates of the ambient concentrations in the training room when six 70 mg training aids were exposed inside the scent boxes. OSHA PEL for dichlorvos is 110 ppbv.

During dog testing, the average MINICAMS estimate of the intrascent box dichlorvos concentration produced by 70 mg aids was about 166 ppbv (*SD* = 53, range = 73 – 316), or 51% in excess of the 110 ppbv OSHA PEL for dichlorvos. However, we judged that these concentrations would pose no hazard to the dogs, because subject exposure to the contents of scent boxes was very limited in duration (recall that the PEL is an occupational standard applying to chronic exposure for 8 hours a day and 40 hours per week for a working lifetime). Intrascent box exposure consisted in most cases of one "sniff" inside the box, followed by the dog withdrawing its head from the box, sitting, and directing its attention to its handler. During the most intensive 70 mg aid exposure, the subjects averaged 14.2 (*range* = 1 – 40) encounters with 70 mg aids per day. Analysis of video showed that subjects spent 0.7 s on average (*SD* = 0.7, *range* = 0.2 – 2.6) sniffing each "hot" box in the course of indicating it. During the course of several months of 70 mg testing and training, the subjects worked an average of 467.3 "hot" boxes (*SD* = 134.6, *range* = 331 – 637), translating to a roughly-estimated average total exposure of only about 5.6 min. Studies with animals have shown that even continuous airborne exposure with up to 222 ppbv of dichlorvos for 28 days (Walker et al., 1972), and up to 555 ppbv for two years (Blair et al., 1976) results in no harmful effects. The highest ambient concentrations of dichlorvos to which handlers and dogs were exposed in the air inside the training room was estimated by MINICAMS at 0.72 ppbv, or less than 1% of the OSHA PEL.

Initial Training. Initial training of a substance detector dog consists of teaching the animal what is referred to as "odor recognition." The subjects were taught dichlorvos odor recognition using one of two methods. Two dogs (Robby I and Tara) were trained using the standard DoD protocol (Craig, 1989). The handler led the dog up to a scent box containing one 70 mg aid, gave the command "Seek," and used a hand gesture to induce the animal to place its nose inside the box. Once the dog had investigated the box, the handler gave the command to "Sit," along with any assistance required to make the animal assume the desired position. Upon completion of the sit response, the handler rewarded the dog with a play object such as a rubber ball. Once the animal was responding reliably to a single 70 mg scent box with an unassisted sit, it was presented with two, then three, then four scent boxes in each trial. On any trial, only one of these boxes contained a training aid. All other boxes were empty, or "blank." Sitting in response to a blank box was not rewarded. The number of trials per day were limited to 30 or less, accumulated across two or three sessions separated by rest periods of 15 to 45 min. Robby I and Tara required 16 days and 290 trials, and 13 days and 211 trials, respectively, to achieve the criterion of 15 consecutive correct four-box trials with the handler "blinded."

Two subjects (Ringo and Robby II) did not exhibit odor recognition after many trials using the standard DoD protocol, and these animals were therefore subsequently trained using a technique called "reward from source." 70 mg dichlorvos training aids were inserted into 30- to 40-cm sections of PVC plastic or aluminum tubes, and then these tubes were used to play retrieving and tug-of-war games with the dogs. During the course of play, the animals presumably formed an association between the target odor and object play. Subsequently, when an object was thrown into low groundcover or a similar concealed location, the subjects immediately fell to searching for it by olfaction, and exhibited the classical "head-snap" and other behavioral changes upon encountering the plume of odor originating from the toy. Ringo and Robby II received five and six days of training, respectively, each day consisting of about 15 min of tug-of-war and retrieving play, and three to four brief search exercises of no more than 45 s in duration.

Once Ringo and Robby II began to exhibit strong odor recognition and olfactory search behavior, they were taught to indicate the presence of a 70 mg dichlorvos training aid in a scent box, and to sample a number of boxes sequentially until they located the "hot" box, using methods similar to the standard DoD protocol. Ringo and Robby II required 18 days and 152 trials, and 11 days and 170 trials, respectively, to achieve the criterion of 15 consecutive correct four-box trials.

Testing Procedures. Following basic training, tests were conducted using training aids made with progressively decreasing masses of dichlorvos (70, 7, 1.4, or 0.50 mg). The testing of each mass required 3 days, with three trials per day. A trial consisted of one search of the 12-box array, the array consisting of six dichlorvos and six vehicle boxes, arranged in the racks described in the Materials section. During each trial, the handler commanded the dog to sample each box/station in turn, moving clockwise around the training room. The starting box for each trial was determined using a random number table, and the locations of the vehicle and dichlorvos boxes in the 12-box array were also randomly plotted. The dog handlers were blind to these locations. An experimenter controlled each trial, telling the handler where to begin searching and when to reward the dog for a Hit, and recording all responses.

Prior to each testing series a new set of cardboard boxes was made. This set of 12 boxes was used for the entire 3-day testing series. Half of the boxes were dedicated to dichlorvos and half to vehicles, and the two kinds of boxes were never mixed or stored in close proximity.

Robby I, Robby II, and Tara were tested with 70 mg, 7 mg, and 1.4 mg aids. Tara had begun to display limping due to aging-related arthritis, so she was then retired from the study. Ringo was selected and "purpose-trained" as rapidly as possible to provide a third subject for testing with 0.50 mg aids, along with Robby I and Robby II.

Results and Discussion

The results of testing are presented in Table 1 in terms of overall accuracy (Accuracy = (Hit Rate + (1 – False Alarm Rate)) / 2) for each dog across all 3 days of testing. Two out of three subjects displayed accuracy better than the 0.95 criterion when tested with 70 mg aids. These results were obtained with a mean dichlorvos concentration inside the scent boxes of 166 ppbv (*SD* = 53, *range* = 73 – 316). All three subjects exceeded criterion accuracy with 7 mg aids, which produced a mean dichlorvos scent-box concentration of 42 ppbv (*SD* = 20, *range* = 8 – 113). Two out of three dogs exceeded criterion accuracy with 1.4 mg aids, which produced an estimated average dichlorvos concentration of 3 ppbv (*SD* = 1.3, *range* = 0.9 – 6.1). Finally, two out of three subjects also exceeded criterion at 0.50 mg, while the third narrowly missed criterion. These results were obtained with a mean dichlorvos concentration inside the scent boxes of only 1 ppbv (*SD* = 0.3, *range* = 0.6 – 1.7).

Table 1
Accuracy During Testing with Four Levels of Dichlorvos Training Aids.

Level of Training Aid	Ringo	Robby I	Robby II	Tara	Mean Accuracy	Dichl. Conc.
70.0 mg	-----	0.99	0.99	0.93	0.97	166
7.0 mg	-----	0.97	1.0	0.99	0.99	42
1.4 mg	-----	0.99	0.97	0.90	0.95	3
0.50 mg	0.97	0.97	0.94	-----	0.96	1

When examined over each 3-day testing series, the data showed that accuracy tended to improve over days (see Figure 2). This trend was most notable in the case of 0.50 mg testing, where all three subjects failed to achieve criterion accuracy on Day 1 (M = 0.90), but all three achieved criterion on Day 2 (M = 0.98), and they exhibited perfect performance on Day 3 (M = 1.0). Similar but less dramatic trends were also noted in the 70 mg and 7 mg data.

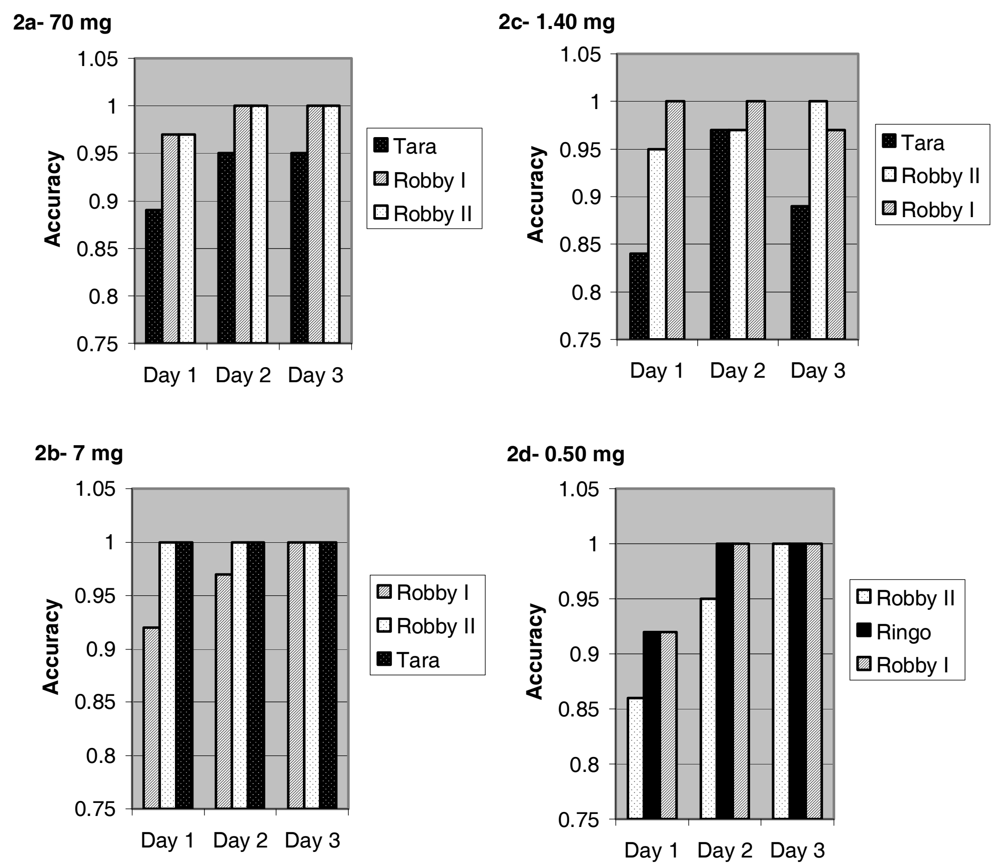

Figure 2. Experiment 1. Daily accuracies for each subject across Days 1 – 3 of (a) 70 mg testing, (b) 7 mg testing, (c) 1.40 mg testing and, (d) 0.50 mg testing.

The improvement in performance across test days was unlikely to have been the result of any practice effect, because the trend began only with the first day of testing. In blinded training with 0.50 mg aids during the week prior to testing, all three subjects exhibited *higher* accuracies (0.98, 0.95, and 0.89 for Ringo, Robby I, and Robby II, respectively) than they did on the first day of 0.50 mg testing (0.92, 0.92, and 0.86, respectively). The improvement in accuracy over days during 0.50 mg testing was probably related to another factor.

New scent boxes were constructed for 0.50 mg testing (as for 70, 7, and 1.4 mg testing). These new boxes were used for all 3 days of testing, with two segregated groups of boxes. One group was dedicated to dichlorvos aids, and one to vehicle aids. It is possible that, during these 3 days, the two groups of boxes were differentially contaminated in a way that incrementally contributed to discrimination between them. It is certain that the dogs contaminated the scent boxes with their own hair and saliva-the boxes became visibly "dirty" with use. This contamination could have been differential, because the subjects tended to interact more intensively and over longer durations with dichlorvos boxes than they did with vehicle boxes. Such differential contamination might have accounted for some part of the detection performance displayed by the subjects with 0.50 mg aids.

To evaluate this possibility, twelve scent boxes previously used for 0.50 mg testing were employed for 6 more days of testing, with two 12-box trials a day. For each trial, six vehicle and six 0.50 mg dichlorvos scent boxes were randomly placed in the 12 stations of the 12-box array. The 6 days of testing were divided into three 2-day blocks. Each 2-day block included one Normal day and one Reversed day. On Normal days, the dichlorvos aids were placed in the same boxes that had previously contained dichlorvos aids during 0.50 mg testing, and vehicle aids were placed in the same boxes that had contained vehicle aids during testing. On Reversed days, 0.50 mg dichlorvos aids were placed in boxes that had previously contained vehicles, and vehicles were placed in boxes that had contained dichlorvos aids.

All three subjects exhibited lower accuracy in the Reversed condition than in the Normal condition (see Figure 3). These reduced accuracies were a function of both decreases in Hit rates, and also increases in False Alarm (FA) rates (see Table 2).

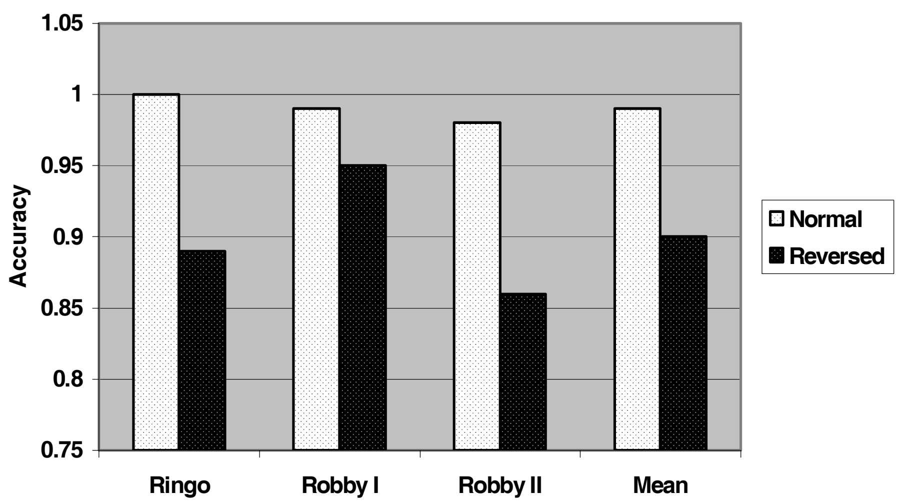

Figure 3. Experiment 1: Box-reversal results. Mean accuracies for each subject, and for all subjects together, in Normal- and Reversed-box conditions.

Table 2
Detection Performance in the Normal Box and Reversed Box Conditions.

	Normal		Reversed	
Subject	**Hit Rate**	**FA Rate**	**Hit Rate**	**FA Rate**
Ringo	1.0	0.0	0.83	0.06
Robby I	0.97	0.0	0.89	0.0
Robby II	0.95	0.0	0.86	0.14
Means	**0.97**	**0.0**	**0.86**	**0.07**

Note. FA = false alarms.

The result of box-reversal testing is consistent with a mechanism in which repeated use of dedicated boxes for dichlorvos training aids results in contamination of these boxes with residual dichlorvos. Such a hypothesis would not militate against the conclusion that two out of three subjects performed at criterion levels on 1 ppbv during testing with 0.50 mg aids (because any dichlorvos contamination would be included in this 1 ppbv estimate).

However, the overall pattern of results in Experiment 1 is also consistent with another mechanism, in which dichlorvos boxes were differentially contaminated by the subjects (with saliva, hair, etc.) in the course of repeated searches. According to this hypothesis, over the course of 0.50 mg testing, as the dichlorvos boxes became progressively more contaminated than the vehicle boxes (because the dogs interacted with them more intensively and for longer durations), the dogs were able to use the canine-contamination cues to help them discriminate between dichlorvos boxes and vehicle boxes. This mechanism would explain why accuracy improved over the three days of 0.50 mg testing. It would also explain why performance degraded on Reversed days. If, at some point during Steel Helmet training and testing, "canine contamination odors" had gained a degree of stimulus control over detection responding, then we would expect the Reversed condition vehicle boxes (heavily contaminated with saliva, hair, etc.) to elicit more FA's than the Normal condition vehicle boxes (lightly contaminated). In addition, we would also expect the Reversed condition target boxes (that had previously contained only vehicles and were contaminated by the dogs to a lesser degree) to elicit fewer Hits than the Normal condition target boxes (that had previously contained dichlorvos aids and were heavily contaminated by the dogs).

According to this "dog contamination" hypothesis, performance on 0.50 mg aids during the initial test series (0.97, 0.97, and 0.94 for Ringo, Robby I, and Robby II, respectively) reflected not only the sensitivity of the subjects to the exceedingly low concentrations of dichlorvos produced by 0.50 mg aids, but also a degree of stimulus control exerted by saliva and hair contaminating the target scent boxes. Although there was no improvement in performance over days in the 1.4 mg data, and trends in 7 mg and 70 mg data were not as dramatic as those in the 0.50 mg data, without conducting box-reversal procedures at each level of dichlorvos concentration, we cannot eliminate the possibility that contamination effects also played a role in the detection performance displayed by the subjects during 1.4, 7, and 70 mg testing. Therefore, the best available measure of canine dichlorvos detection accuracy is given by the subjects' performance on Day 1 of testing with each level of aid (see Table 3), when the contribution of differential contamination can be assumed to have been at its weakest. Before pursuing this line of reasoning further, it should be noted that no more than two out of three subjects achieved criterion at any level of Day 1 testing, even with 70 mg aids. Accordingly, we redefined our canine detection threshold as the lowest level of dichlorvos concentration at which a *majority* of the subjects achieved criterion performance.

Table 3
Accuracy During Testing with Four Levels of Dichlorvos Training Aids,in Day 1.

Level of Training Aid	**Ringo**	**Robby I**	**Robby II**	**Tara**	**Mean Accuracy**	**Dichl. Conc.**
70.0 mg	-----	0.97	0.97	0.89	0.94	166
7.0 mg	-----	0.92	1.0	1.0	0.97	42
1.4 mg	-----	1.0	0.95	0.84	0.93	3
0.50 mg	0.92	0.92	0.86	-----	0.90	1

Two out of three subjects exceeded criterion on the first day of testing with 70 mg aids. This was also true with 7 mg aids, and with 1.4 mg aids. However, when tested with 0.50 mg aids, all three subjects failed to achieve criterion on Day 1. We must therefore conclude that the canine dichlorvos detection threshold falls at or below the 3 ppbv produced by 1.4 mg aids, but not so low as the 1 ppbv produced by 0.50 mg aids.

Experiment 2: Effects of Distracter Substances

It is natural to assume that the canine dichlorvos detection response (when contamination effects are accounted for) is controlled by the specific odor(s) produced by dichlorvos. However, in principle the dogs could solve the problem of identifying dichlorvos boxes simply by responding whenever they encounter an odor that is discrepant from the background, or especially salient (see Johnston & Waggoner, 1998, for an example in which instrumental responding by dogs in operant conditioning chambers is elicited by any unfamiliar odor). In such a case, the detection performance displayed by the Steel Helmet dogs in our semioperational scent-box paradigm might not generalize to truly operational settings (e.g., engine rooms, flight lines, and kitchens) that incorporate countless substances generating unique and salient odors. The specificity of Steel Helmet detection performance was therefore assessed by comparing responding to dichlorvos with responding to novel distracting agents. This testing was also an opportunity to directly assess the strength of behavioral control exerted by vehicle boxes compared to blank/empty boxes.

Such investigation is essentially a problem in testing for stimulus generalization (Guttman and Kalish, 1956) or stimulus control (Terrace, 1966). When comparing responding to different stimuli, it is crucial to avoid procedures resulting in differential reinforcement of responding to one or another of the stimulus alternatives. Differential reinforcement provides the opportunity for within-testing discrimination learning, making it impossible to correctly characterize the stimulus control of the behavior of interest that existed *prior* to testing. The solution traditionally applied to such difficulties in stimulus control testing is to test in extinction (e.g., Hanson, 1959). However, this solution also provides the opportunity for within-testing learning, because suppression of the previously conditioned response will result.

The problem of within-testing learning was addressed by testing the subjects in partial extinction, with responses to dichlorvos, distracter, vehicles, and empty boxes all yoked together on identical 29% reinforcement schedules. This manipulation did not preclude the possibility of within-testing learning, but the yoked arrangement should have served to discourage discrimination between dichlorvos and distracter, while the random and attenuated reinforcement schedule should have retarded association of nontargets with reward (e.g., Stevenson & Zigler, 1958). In addition, any adventitious learning that did take place would have tended to contribute to responding to nontarget boxes. This would only pose a problem for interpretation of the result if a high rate of responding to nontarget boxes were found. On the other hand if, as we predicted, responding to nontarget boxes were negligible, then this result would be clearly interpretable.

Method

Prior to testing, the two subjects were given a number of training trials (12 days and 37 trials, and 13 days and 40 trials, for Robby I and Tara, respectively) during which they were gradually worked down to random reinforcement schedules with reinforcement probability as low as 0.25. (Note: Tara's participation in this procedure antedated her participation in the 1.4 mg testing of Experiment 1, after which she was retired due to arthritis and adopted into a private home.)

The subjects then experienced 2 days of testing, with four 12-box trials per day. In each trial, there were three 7 mg dichlorvos boxes, three boxes containing a distracter aid, three vehicle boxes, and three blank (empty) boxes, for a total of 24 of each kind of box over the 2-day test. Reinforcement was scheduled so that there was a 29% probability of reinforcement for any detection response that occurred, regardless of the type of box that elicited this response (24 of each type of box were presented, and 7 rewards were programmed for responses to each type). However, the constraint was imposed that no reinforcements were available for any detection responses that might occur during the first 12-box trial.

During the testing procedure, the handler was blind to the locations of the dichlorvos, distracter, vehicle, and blank boxes. At each box, he or she judged from the dog's behavior whether the box contained dichlorvos and indicated this decision by saying the position (1-12) of the box. Using the previously-arranged schedule of random reinforcement, the experimenter informed the handler which detection responses should be reinforced, but without telling him/her whether these responses were to dichlorvos, distracter, vehicle, or empty scent boxes.

Each distracter aid was prepared using one of seven different volatile/odorous chemicals—anisole, n-undecane, trans-anethole, endo-borneol, 4-5-dimethylthiazole, geraniol, and vanillin. In each case, 5 mg of the chemical (diluted or dissolved in methanol) was adsorbed onto Whatman 17 chromatography filter paper pads. Each of the distracting odors appeared in three boxes over the course of testing, except endo-borneol, vanillin, and 4-5-dimethylthiazole, which appeared in four boxes (for a total of 24 distracting boxes over the 2 days of testing).

Results

Both subjects displayed strong responding to the boxes containing 7 mg training aids, but little or no responding to distracter, vehicle, and empty/blank boxes (see Figure 4). Robby I and Tara exhibited 22 and 19 detection responses, respectively, to dichlorvos boxes, 1 and 0 responses, respectively, to the boxes containing distracter aids, 1 and 0 responses, respectively, to vehicle boxes, and 2 and 0 responses, respectively, to empty boxes.

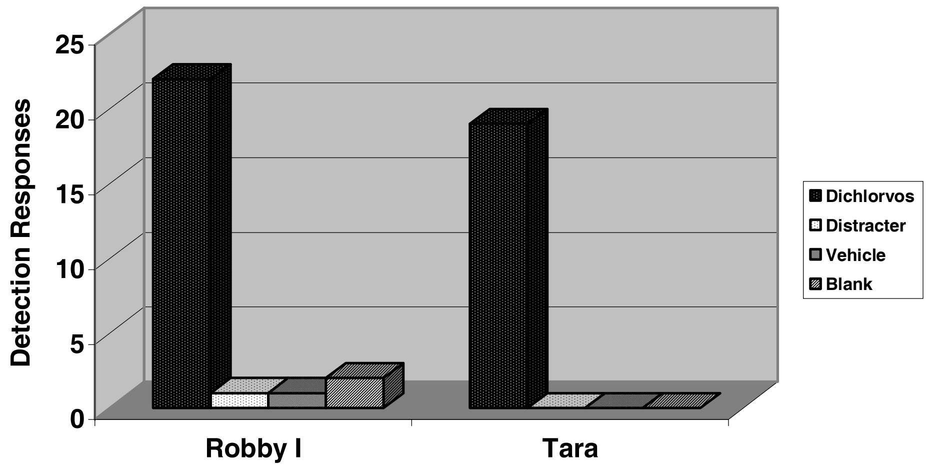

Figure 4. Experiment 2. Numbers of responses by each subject to dichlorvos, distracter, vehicle, and blank (empty) boxes.

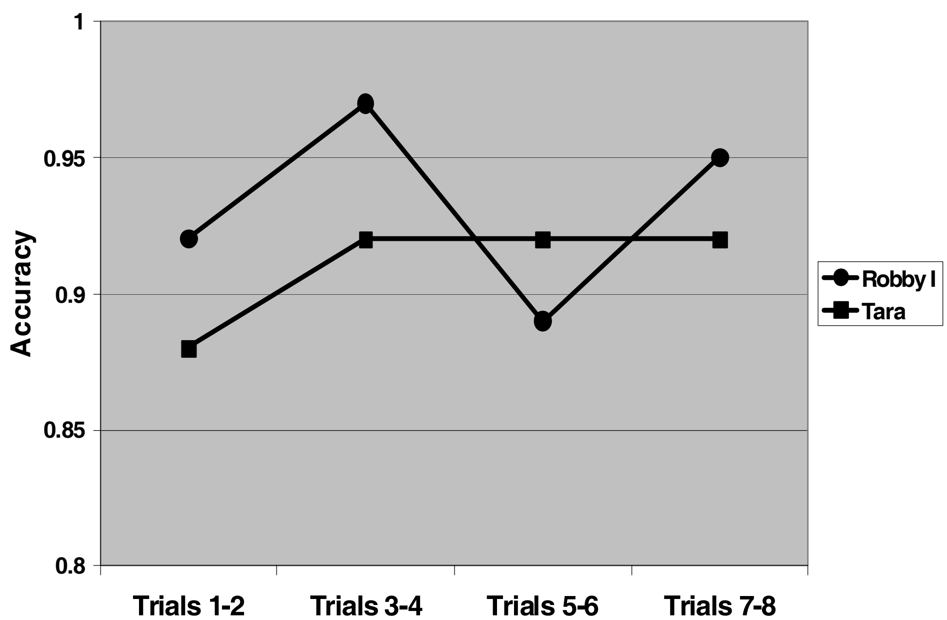

Figure 5. Experiment 2. Mean accuracy (in terms of discriminating dichlorvos boxes from distracter, vehicle, and blank boxes) for each subject over four 2-trial blocks.

In order to examine the data for performance improvements over the course of testing, each subject's mean accuracy score (Accuracy = (Hit rate + (1 – False Alarm rate)) / 2) for each two-trial block (Trials 1 –2, 3 – 4, etc.) was calculated, classifying as a False Alarm any response to distracters, vehicles, or blanks. Over the four blocks, Robby I's data exhibited no systematic improvement (see Figure 5). Tara's data showed an initial increment from Trials 1 – 2 to an asymptote over Trials 3 - 8, but the magnitude of this increment comprised less than 5% of the asymptotic level of accuracy.

Discussion

Robby I and Tara exhibited a mean of 20.5 responses to boxes containing dichlorvos. In contrast, they averaged only 0.50 of a response to distracter boxes. The reinforcement schedule was arranged so that rewards were programmed at the same rate for sits on distracters as for sits on dichlorvos (and, indeed, only one sit on a distracter took place), so there was little possibility for discrimination learning to have taken place during testing. Thus novel odors exhibited little control over responding, and we may conclude that the detection behavior exhibited throughout Experiment 1 was not the result of a nonspecific mechanism in which the dogs responded to any odor that was novel, or discrepant from or more salient than the olfactory background.

Very weak responding to vehicle boxes (M = 0.50 of a response) and blank boxes (M = 1 response) showed that vehicle aids (i.e., chromatography paper, thumbtack, and associated odors) and empty scent boxes also exerted only negligible stimulus control.

As in Experiment 1, the cardboard scent boxes in Experiment 2 were dedicated to dichlorvos, distracter, vehicle, and blank roles, and the same set of boxes were used for both days of testing. However, there was no evidence for notable improvement in performance over the course of testing, and thus no indication of a significant role for dog contamination odors in controlling detection responding during this experiment.

All of these observations are compatible with the idea that canine dichlorvos detection responding was controlled predominantly by dichlorvos odor(s).

General Discussion

The Steel Helmet project was carried out to ascertain whether dogs could be trained to detect an organophosphate CWA surrogate (dichlorvos) at low concentrations, and to estimate the canine detection threshold for this substance. This canine detection threshold was defined as the lowest concentration of dichlorvos vapor at which a majority of the subjects could exhibit accuracy of 0.95 or better.

Experiment 1 showed that two out of three detector dogs could discriminate scent boxes containing 1.4 mg dichlorvos training aids (producing about 3 ppbv of dichlorvos vapor) from those containing vehicles with accuracy greater than 0.95, even on the first day of testing when the effects of contamination of scent boxes were minimal. However, they could not do so with 0.50 mg training aids (producing about 1 ppbv). Accordingly, the canine dichlorvos detection threshold falls at or below 3 ppbv, but not as low as 1 ppbv.

Experiment 1 also showed that repeated use of dedicated scent boxes leads to differential contamination of vehicle versus dichlorvos boxes (possibly by cues arising from the greater amounts of saliva and hair which the dogs deposited on the dichlorvos boxes), and that these contamination cues can exert a degree of stimulus control (slight compared to the control by dichlorvos) over detection responding. Future detector dog investigations based on the DoD scent-box protocol should account for box contamination by employing box-reversal procedures, or limit contamination by using new boxes for each day's testing.

Experiment 2 confirmed that the Steel Helmet subjects' responding was not controlled by concomitant odors associated with the filter paper pads and the scent boxes used to deliver dichlorvos, because there was very little responding to vehicle boxes and empty boxes. Experiment 2 also provided additional evidence that detection responding was controlled specifically by dichlorvos, because there was likewise very little responding to boxes containing aids made with a variety of salient and odorous distracting chemicals.

It should be noted that these observations probably apply only to dogs with high aptitude for substance detection. Individual differences play a major role in determining the trainability and usefulness (Department of the Treasury, 1993) and the low-concentration thresholds (Johnston, 1998; Waggoner et al., 1998) of detector dogs. For example, in the course of the Steel Helmet project, attempts were made to train five additional subjects, two of them naïve dogs, and three of them rejects from the basic detector dog training programs of the Department of Defense (DoD) and the Federal Aviation Administration (FAA). All of these subjects were eliminated from the study after failing to meet criterion by substantial margins at various stages of the project (the two naïve dogs during initial dichlorvos training, and one each of the DoD/FAA rejects during 70 mg, 2.8 mg, and 1.40 mg testing). The data from these animals is not included in this report because, as rejects of either basic detector dog training programs or initial Steel Helmet training procedures, they were deemed unrepresentative of the population of operational detector dogs. However, their limited success in the course of Steel Helmet is instructive from the standpoint of showing that highly accurate performance on 3 ppbv or less of dichlorvos vapor is probably within the capability of only a select group of dogs.

Viability of the CWA Detector Dog Concept

For the sake of discussion, we will use the figure of 3 ppbv as the canine detection threshold for dichlorvos. This figure is two to three orders of magnitude greater than the *occupational* airborne exposure limits (AELs) for GB and VX-0.01 ppbv, or 0.0001 mg/m^3, and 0.001 ppbv, or 0.00001 mg/m^3, respectively (U.S. Army Soldier and Biological Chemical Command, 1996, 2001). If we consider detector dog exposure to target odor to be acute rather than occupational, we may be justified in using as a safe canine GB exposure limit a figure higher than the AEL for GB, such as the figure of 5.56 ppbv (0.05 mg/ m^3) for 10 min arrived at for canine exposure by the U.S. Navy's Office of Special Technology (OST, 1997) after an extensive review of available sources. This figure exceeds the canine dichlorvos detection threshold of 3 ppbv (0.027 mg/ m^3) by nearly a factor of two, and should provide sufficient margin of safety to permit exposure of detector dogs to GB.

However, although trained and experienced dichlorvos detector dogs are clearly capable of accurately detecting 3 ppbv (0.027 mg/ m^3) of dichlorvos, it is likely that, in order to acquire the necessary skills and odor recognition, they require much higher concentrations of target odor initially. (Recall that the Steel Helmet subjects were taught dichlorvos odor recognition at about 171 ppbv, or 1.55 mg/m^3.) Unless GB is substantially greater in salience (and hence "associability") than dichlorvos, we are doubtful that dogs can be trained to recognize only 3 ppbv (0.027 mg/ m^3) of GB vapor. Thus a gulf may exist between OST's estimate of safe

GB exposure levels, and the lowest concentration of GB that a talented dog can learn to recognize during initial training.

This gulf may be reduced if, by taking into consideration the very brief, phasic, and intermittent target odor exposure involved in initial training of a detector dog, we were to revise upward from 5.56 ppbv or 0.05 mg/ m^3 (Office of Special Technology, 1997) our estimates of the safe limits for canine exposure to GB. In this case, it may be possible to arrive at safe, practical criteria and methods by which detection dogs can be trained to recognize and localize GB vapor. In the case of substances significantly more toxic than GB, such as VX, the gulf between safe exposure levels and the vapor concentration a dog can readily learn to recognize may be wide indeed.

References

Blair, D., Dix, K. M., Hunt, P. F., Thorpe, E., Stevenson, D. E., & Walker, A. I. (1976). Dichlorvos: A 2-year inhalation carcinogenesis study. *Archives of Toxicology,* **35** , 281-294.

Cain, W. S, Mason, J. R., & Morton, T. H. (1985). *Use of animals for detection of land mines and other explosives: A review and critique of prospects.* Report prepared for U.S. Army Belvoir Research and Development Center: Fort Belvoir, VA.

Caluwé, R. de (1995). *De eerste Belgische politiehonden: 1899-1914*. Gent, Belgium: Uitgeverij KOJ.

Ceulebroeck, G. van (1983). *L'historique du Berger Belge*. Charleroi, Belgium: Presses de la Concorde.

Chao, E. T. (1977). *Olfaction in dogs: A critical review.* Unpublished doctoral dissertation, Florida State University.

Craig, D. (1989). *Detector dog training protocol*. Lackland Air Force Base, TX: Department of Defense Military Working Dog Veterinary Service.

Department of the Treasury (1993). *Law enforcement canine training.* Office of Planning and Management Analysis. Washington, DC.

Ehlers, R. C., & Barclay, W. L., Jr. (2001). *The canine detection of chemical agents study—Phase I*. San Antonio, TX: Southwest Research Institute.

Guttman, N., & Kalish, H. I. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology,* **51**, 79-88.

Hanson, H. M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology,* **58**, 321-334.

Home Office Standing Advisory Committee on Police Dogs (1996). *Police dogs: Training and care*. London, UK: Home Office Standing Advisory Committee on Police Dogs.

Johnston, J. M. (1998). *Enhanced canine explosives detection*. Fort Washington, MD: Office of Special Technology.

Johnston, J. M., & Waggoner, L. P. (1998). *Canine olfactory detection of methyl benzoate and cocaine*. Fort Washington, MD: Office of Special Technology,

Most, K. (1951). *Training dogs: A manual*. Essex, UK: Anchor Press LTD.

Nicklin, S. (2000). *Why use dogs*? Paper presented at the Gordon Conference on Illicit Substance Detection, New London, CT.

Office of Special Technology (1997). *Steel Helmet Project—Report 1: Canine detection of sarin (GB) and the effects of the inhaled toxic vapor*. Fort Washington, MD.

Spectrum Laboratories (1998). *Chemical fact sheet: Dichlorvos* (On-Line). Available: www.speclab.com.

Stephanitz, M. von (1923). *The German shepherd dog in word and picture*. Munich, Germany: Verein für Deutsche Schäferhunde.

Stevenson, H. W., & Zigler, E. F. (1958). Probability learning in children. *Journal of Experimental Psychology,* **56**, 185-192.

Terrace, H. S. (1966). Stimulus control. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application*. New York: Appleton Century Crofts.

U.S. Army Soldier and Biological Chemical Command (1996). *Material safety data sheet: Lethal nerve agent (GB).* Aberdeen Proving Ground, MD.

U.S. Army Soldier and Biological Chemical Command (2001). *Material safety data sheet: Lethal nerve agent (VX).* Aberdeen Proving Ground, MD.

Waggoner, L. P., Jones, M., Williams, M., Johnston, J. M., Edge, C., & Petrousky, J. A. (1998). Effects of extraneous odors on canine detection. *Proceedings of the SPIE Conference on Enforcement and Security Technologies*, *Vol* 3575.

Walker, A. I., Blair, D., Stevenson, D. E., & Chambers, P. L. (1972). An inhalational toxicity study with dichlorvos. *Archives of Toxicology,* **30***,* 1-7.

Williams, M., Johnston, J. M., Waggoner, P., Jackson, J., Jones, M., Boussom, T., and Hallowell, S. F. (1997). Determination of the canine detection odor signature for NG smokeless powder. *Proceedings of the second explosives detection technology symposium and aviation security technology conference*. Federal Aviation Administration, Atlantic City, NJ.

Received October 7, 2002.
Revision received March 20, 2003.
Accepted April 28, 2003.