COMPARISON OF DOG TEAMS AND POLYGRAPH
IN DETECTING "GUILT"

by

CARMEN ELIZABETH RAMIREZ MONZON
B.A., Pontificia Universidad Javeriana, 1972

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS
THE FACULTY OF GRADUATE STUDIES
in
Department of Psychology

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
August, 1977
© Carmen Elizabeth Ramirez M., 1977
In presenting this thesis in partial fulfilment of the requirements for
an advanced degree at the University of British Columbia, I agree that
the Library shall make it freely available for reference and study.
I further agree that permission for extensive copying of this thesis
for scholarly purposes may be granted by the Head of my Department or
by his representatives. It is understood that copying or publication
of this thesis for financial gain shall not be allowed without my
written permission.

Department of Psychology

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date Jan 10/78
ABSTRACT

A study was conducted to evaluate the ability of police dog teams to identify "guilty" subjects in a simulated crime situation and to compare their accuracy with that of a polygraph examination. Research on the olfactory acuity of dogs, and on the role of olfactory cues such as pheromones in social communication, implies that the detection of guilt by experienced police dogs could occur as reliably as police dog handlers believe. The literature on polygraph investigations shows high reliability in detecting guilt. This was one of the reasons for using the polygraph as the comparison technique. Three experienced dog teams from the Vancouver Police Dog Squad and two expert polygraph field examiners were used.

The subjects were 64 male volunteers, all university or college students. Subjects randomly assigned to the "guilty" condition were instructed to "steal" and conceal a $10 bill that had been left in an empty office, and to deny throughout the rest of the experiment that they had done so. Volunteers in the "not guilty" condition were told nothing about the "crime". Both groups were told that police dog teams and polygraph operators would try to find out whether they were guilty. They were promised $5.00 for participating plus a bonus of $10 if they succeeded in establishing that they were innocent.

Police dog team performance was about chance level, while the polygraph examination was significantly more accurate than chance and than the dog teams. No individual difference was found among the dog teams. The failure of the dog teams could be attributed either to the impossibility of detecting guilty by smell cues or to some aspect of the simulation.
procedure. Further research should be directed at developing more realistic field studies.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>The Use of Police Dog Teams</td>
<td>2</td>
</tr>
<tr>
<td>Research on Canine Olfactions</td>
<td>6</td>
</tr>
<tr>
<td>Social Olfaction</td>
<td>9</td>
</tr>
<tr>
<td>Pheromones in Human Beings</td>
<td>12</td>
</tr>
<tr>
<td>Canine Detection of Guilt: The Hypothesis</td>
<td>14</td>
</tr>
<tr>
<td>The Polygraph</td>
<td>16</td>
</tr>
<tr>
<td>Summary</td>
<td>22</td>
</tr>
<tr>
<td>METHOD</td>
<td>23</td>
</tr>
<tr>
<td>Participants</td>
<td>23</td>
</tr>
<tr>
<td>Subjects</td>
<td>23</td>
</tr>
<tr>
<td>Dog Teams</td>
<td>23</td>
</tr>
<tr>
<td>Polygraph Operators</td>
<td>23</td>
</tr>
<tr>
<td>Procedure</td>
<td>24</td>
</tr>
<tr>
<td>General Orientation</td>
<td>24</td>
</tr>
<tr>
<td>Conditions</td>
<td>25</td>
</tr>
<tr>
<td>Design</td>
<td>26</td>
</tr>
</tbody>
</table>
Table of Contents (cont'd)

Physical Facilities .... 26
Dog Test Procedure .... 27
Polygraph Test Procedure .... 27
Order of Events .... 28
Summary of Variables .... 29

RESULTS .... 30
DISCUSSION .... 34
BIBLIOGRAPHY .... 42
APPENDIX A Instructions to Dogmasters .... 49
APPENDIX B Polygraph Questions .... 51
APPENDIX C Instructions to the "Guilty" Subjects .... 54
APPENDIX D Instructions to the "Not Guilty" Subjects .... 58
APPENDIX E Polygraph Examination Technique .... 61
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 1.</td>
<td>Performance Accuracy (Pairs of Subjects)</td>
<td>32</td>
</tr>
<tr>
<td>TABLE 2.</td>
<td>Performance Accuracy (Individual Subjects; N = 64)</td>
<td>33</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

There are many people without whose assistance this investigation could not have been carried out, and to whom the author wishes to express her gratitude.

Dr. Robert H. Wright was a constant source of help and information throughout the entire course of the research. His expertise in the area of olfaction, and his experience with police dogs, were crucial in designing and conducting the work. He also greatly facilitated the participation of the police dog teams.

Ms. Margaret Byrne provided important assistance in the actual running of the experiments, helping to recruit and schedule subjects, prepare the site, and administer the project. Her participation during many weekends throughout the course of the study is greatly appreciated.

Sergeant Pat Laughy was the officer in charge of the Vancouver Police Dog Squad during the period of this work. His willingness to sacrifice many hours of his free time, and to encourage his colleagues to volunteer as well, were vital in making the research possible. Special thanks go to the three dog teams that were involved: Sgt. Laughy and Duke, Const. Adrian de Jong and Spike, and Const. John LeVoie and Rex IV.

Similarly, the two polygraph examiners, Mr. John Weller and Mr. Harold Stevenson, devoted a great deal of time and effort to this study. Working without compensation, they too were centrally involved in the carrying out of the research.

The author also expresses her thanks to the members of the thesis committee. The chairman of the committee, Dr. Peter Suedfeld, provided invaluable
assistance and guidance during all phases of the work. Drs. Charlan J. Nemeth and Lawrence M. Ward gave suggestions concerning the design and procedure of the study. Dr. Ward also helped with the analysis of the results, and contributed extensive editorial improvements to the thesis. Dr. Robert E. Knox and Dr. Roderick Wong, joining the committee at a later stage, made valuable comments on the final draft.

Besides the committee members, the author thanks Dr. A. Ralph Hakstian for his assistance in designing the analyses and interpretation of the results, and Dr. Demetrios Papageorgis for allowing his office to be used as the scene of the crime.
INTRODUCTION

This research is concerned with testing the validity of two ways in which law enforcement agencies attempt to distinguish between guilty and innocent suspects when a crime has been committed. Clearly society must be able to identify the true perpetrators of crime in order to apply appropriate sanctions, rehabilitative techniques, protective segregation from the public, and so on. Our system of criminal justice should be directed toward accomplishing this goal as accurately, efficiently and fairly as possible.

The initial contact between the law and the suspected criminal, and the one which is frequently crucial, occurs when the police identify and apprehend a suspect. While this is to a great extent a matter of routine collection of evidence and of pursuit, the police do make a preliminary judgement as to the probable guilt of the apprehended person. This judgement is not binding in the legal sense, but often it determines the subsequent course of investigation and prosecution of the accused.

When the identity of the guilty person is not obvious—e.g., he was not caught in the act of lawbreaking, or there are several suspects—modern police departments use a variety of techniques to reduce uncertainty. Some of these are related to the physical gathering and evaluation of evidence, and some to interrogation of suspects and of witnesses.

Two of the systematic methods currently employed to further this work are "lie-detector" (polygraph) examinations and the use of trained police dogs and their masters. In North America, neither of these methods can be used as direct evidence in criminal trials but police do use both to indicate
in which direction further investigation should go. In this way the results of such methods, frequently determine the fate of suspects. In brief, the research that will be described here was designed to evaluate the efficacy of trained dog and master teams in detecting guilt in a simulated crime situation, and to compare the validity of the technique with that of polygraph examinations. To the maximal degree possible, the personnel involved and the procedures used were the same as those commonly found in field investigations of actual crime.

The more versatile, but less publicized and much less researched, use of police dog teams will be discussed first. The use of the polygraph, which is fairly straightforward, restricted, and relatively well known, will be reviewed later in this section.

In order to provide the relevant background we must look at the literature bearing upon the training and employment of police dogs; research data related to the olfactory acuity of dogs and the importance of social olfaction among animals, including the role of pheromones in this context. These topics are basic to the consideration of the possibility that dogs may detect some special odor associated with feelings of guilt.

The literature related to polygraph techniques will also be discussed, since it is to these procedures, that the dogs performance will be compared.

The Use of Police Dog Teams

According to a summary of operations prepared by the Vancouver Police Dog Squad (Campbell, 1969), the use of dogs has been a definite asset to police efficacy. The use of dogs has increased the number of arrests in such situations as building searches, tracking and crowd control (Campbell, 1969).
It has been claimed that the use of dog teams can prevent or defuse some potentially dangerous situations. For example, it has been reported (Campbell, 1969) that because the public believes that dogs are truly effective, the combination of a police officer and a dog reduces the likelihood of criminal activity in crowds. Campbell (1969) estimated that one well trained dog team can replace ten policemen in crowd control work at parades, demonstrations, street festivals, and the like. The economic advantages of dog use may be seen in the fact that the maintenance of a police dog cost only about $200 per year in 1969. The 1977 estimate (Campbell, personal communication, August 1977) is $350, still a great saving. In cases where an arrest is being made once a suspect has been trapped, it has been found that the criminal usually surrenders more peacefully when confronted by a dog. As a result, the use of police dogs by the Vancouver Police Department alone increased by 12.6% between 1968 and 1969, and their use in arrest situations by 64% between 1967 and 1968. These figures were collected in the first formal evaluation of the Dog Squad (Vancouver Police Department, 1969). This is the last year for which data were available (Campbell, personal communication, 1977).

With respect to odor identification the RCMP favours the use of dogs in airports and post offices particularly for the detection of explosives, narcotics and restricted agricultural goods (Marshal, 1976). This use of dogs is common in several Western countries, including West Germany, Sweden, and the United States (Schimscaccl1976;awide1976; Linde & McEathron, 1976). Dogs assigned to this type of detection are trained to recognize the particular odor to be sought among many different odors as in the identifi-
cation of suspects, there have been many reported instances of accurate detection; but there are no systematic assessments of overall performance effectiveness.

With human odors, police have suggested that the dog learns to identify "something more than just the scent of the person" (Campbell, 1969). In fact it is sometimes claimed that trained dogs will not capture a non-guilty individual, since the dog is said to respond to a unique "guilt" odor. However, there is no scientific evidence as to the dogs' alleged guilt-discriminating capabilities. At present, the only information available about the use of dogs in criminal circumstances is to be found in police records.

According to Campbell (1969), many police dogmasters believe that there is a chemical component, a detectable substance emitted when a guilty individual confronts the police, which differs from normal human scent. This factor enables experienced dogs to identify criminal offenders even in the presence of other individual odors. A police dog with some months or years of experience with criminals and criminal suspects can supposedly recognize this smell emanating from persons who are in an emotional state that combines elements of fear, hostility and resentment. It is said to be this combination of emotions, whose olfactory concomitant is referred to as the odor of guilt, that the dog has learned as a critical stimulus. While there is no direct biochemical evidence on this point, it is conceivable that a person in this special complex emotional state may emit a particular combination of chemical molecules that the dog can learn to recognize and identify, and that he will indicate to his master.
According to Wright (1964) as well as Campbell (1969), an experienced police dog can distinguish the "odor of guilt" from normal anxiety caused simply by being a bystander in a police dog tracking or search situation. To date, however, no studies have been done to test the validity of this belief.

In one common situation, dogs are used to detect people unlawfully concealed in a building. Police dogs are trained to detect scents from different locations, classified as ground scent, contact scent and windborne scent. In this particular condition, the dog relies primarily on currents of air, which carry the fresh scent of the hidden person throughout the area. For this task it is necessary to have a thoroughly experienced dog who has participated in several successful apprehensions; novice dogs are easily confused by the situation (Campbell, 1969).

The dogs are trained in such a way that they can find persons in different areas. When an individual is located, the dog faces the location and barks until the handler arrives. The dog is trained to do anything to prevent escape, although he will bite only if the person attacks him or the dogmaster. When searching a building, the dog is under the control and guidance of his handler but is unleashed until the suspect is found.

One obvious question concerns the transmittal of information from the dog to his master. Dog masters find it very difficult to describe or explain the level of communication that they establish with their dogs. In general they say that the system combines the animal's sounds and gestures to express a complete message, and that this code system is developed and learned throughout the entire training process. Eventually, the handler is able to interpret the animal's sounds and gestures as expressing
reactions ranging from injury to enjoyment. The main sounds are categorized as bark, growl, whine, and whimper. These sounds are accompanied by physical movements, stances and facial expressions, which also have specific meanings (Campbell, 1969).

All of these hypotheses lead to many questions: Are police dogs really accurate in identifying criminals? What cues are involved in the procedure that allows the dog to pick the guilty person and the handler to respond accordingly? How accurate is the technique? Is this method better than other methods in identifying guilty suspects? The present study attempts to answer these questions.

Research on Canine Olfaction

According to our preliminary discussions with the dogmasters of the Vancouver Police Dog Squad, the most important cue used by the dog in pursuit is the smell emanating from the suspect or criminal. The suspect's bodily movements provide a second, much less important, source of information. Although this is consensual among the dogmasters, there is surprisingly little literature in the area (see Wright, 1964). The belief is plausible, given that the dog has an extremely well developed olfactory system. For example, the number of cilia per olfactory receptor is one crucial factor in olfactory acuity. In the dog, the number varies from 100 to 150, contrasted to ranges of 6 to 8 in human beings, 15 to 20 in the rat and approximately 40 in the domestic cat (Brown, 1975). It has been shown that dogs can detect the aliphatic acids present in skin secretions at concentrations less than one millionth of the olfactory threshold for human smell (Neuhaus, 1957). Police dogs are even more likely to
have good smelling capability: German shepherds, used exclusively in the Vancouver Squad, are distinguished from other breeds by their superior olfactory acuity, among other desirable characteristics (Campbell, 1969).

In view of the hypothesized fundamental role of smell in the work of the police dog, it is important to examine the results of experiments related to canine performance in detecting and discriminating odors.

It has been shown that dogs have no difficulty in identifying the owner of personal possessions, such as wallets or handkerchiefs. They can do this even with objects that have been deodorized and then handled very briefly (Kalmus, 1955; Löhner, 1926). Police dogs trained to follow human tracks and show-dogs trained to retrieve objects previously handled by people can distinguish between the body odors of different individuals (Kalmus, 1955). This is true even when the persons involved are all members of a single family. Furthermore, individual characteristics of body odor are identified by the dog regardless of the bodily region from which the smell emanates (e.g., palm, armpit, sole) although these regional odors appear quite different to the human nose. Personal odors are discriminated by the dog even when a particular individual's scents are mixed with that of another person or with various strong smelling substances (Kalmus, 1955).

In retrieval situations, dogs cannot discriminate between the body odors of identical twins if the scents are encountered one after the other (Kalmus, 1955). However, in tracking situations, where the odors of two identical twins are presented simultaneously in mixed form, the scents are consistently distinguished by the dogs. Interestingly, when the dog
is familiarized with the odor of one twin, and then is presented in a test with that of the other and of unrelated individuals, he can pick out the scent of the twin even though he had not previously encountered it. Thus, it seems that under some conditions dogs can distinguish between identical twin partners, although their odors are evidently more similar to each other than those of any pair of less closely related people.

The work of Lubow, Kahn, and Frommer (1973, 1976) approached the study of the dog as an information processing system in a discrimination situation. In particular it was of interest to find out whether there were special characteristics of the dog in the discrimination process as a result of its well-developed olfactory system. Lubow et al. found a definite learning process: the dogs showed a high level of retention of these discriminations even after two months. Early learning was retrieved better than later learning (Lubow et al., 1973).

Later studies (Lubow et al., 1976) found that dogs used a position habit (left-position preference) as part of their sampling strategy. This tactic is similar to those found in rats (Hall, 1974; Mandler, 1966; Siegel, 1967). The dogs exhibited this behavior regardless of the location of the previously reinforced stimulus. The position habit seemed to be important as the basis for more complex and successful approaches, since it was modified rather than abandoned in later trials as the sampling strategy became increasingly efficient.

As the review of the literature indicates, there are both physiological and behavioral data to support the contention that dogs have high acuity in detecting, following, and identifying human scent. Consequently,
there is no reason so far to reject the hypothesis that they could distinguish a smell related to particular emotional or biochemical conditions. If such specific odors do in fact exist, were the emission of such smells demonstrated, the occurrence and detection of a guilt odor would seem to be acceptable as a hypothesis for research.

Social Olfaction

The actual role of scents in social communication has been the topic of several investigations. In particular, researchers have looked at the communicative consequences of pheromones produced in specific emotional states. While most of this work has been done with infrahuman species, there is some evidence of similar mechanisms operating at the human level as well.

Cheal and Sprott (1971) have argued that an important contribution of olfaction research is the possibility of helping to understand animal communication. Since the senses of vision and audition are, in general superior to the tactile and olfactory senses in man, human communication tends to rely upon the first two channels. But many other species, including the largely nocturnal rodents, are quite sensitive to smell stimuli and may therefore rely upon an olfaction to a greater extent. For example, albino mice use only olfactory cues to discriminate among conspecifics, either of the same or opposite sex (Kalkowski, 1967, 1968). The disadvantage of less specificity in localization may be offset by longer persistence of the stimulus and its detectability over fairly long distances. The existence of chemical communication through pheromones has been shown in a variety of
animals, including some invertebrates, and has been hypothesized to be the first type of communication in protozoa (Cheal and Sprott, 1971).

Pheromones, which are frequently mentioned as agents in olfactory communication, have been defined as "Chemical odors emitted by one individual and perceived by a second, and in the latter, acting on the hypothalamus and associated limbic structures, causing hormonal effects" (Schneider, 1974, p. 220). The term pheromone is derived from the Greek "pherein" meaning to carry and "horman" meaning to excite or to stimulate (Cheal & Sprott, 1971). Although pheromones may also act on the recipient by ingestion and absorption, only olfactory pheromones have so far been identified in mammals. Workers have divided pheromones into two categories, primers and releasers. The first of these types has the effect of initiating a physiological reaction, which may be in the form of endocrine, morphogenic, or metabolic changes. The second type triggers a behavioral response. Releasers have also been called signalling pheromones, to avoid the implication of a reflexive or automatic S-R "release". From this point of view, the pheromone is interpreted as providing information, which the recipient may—but does not necessarily have to—use as a response-evoking stimulus (Cheal & Sprott, 1971).

Among major areas in which pheromones have been shown to affect behavioural and/or physiological reactions are reproduction and fear or stress reactions. For example, when female mice that had been exposed to male mice are subsequently housed with strange males, they show blocking of pregnancy; this effect disappears when the females' olfactory bulbs are removed, suggesting that the sensing of a pheromone was a critical factor
(Bruce, 1959, 1961). On the other hand, females housed with other females occasionally show a prolonged diestrous phase (pseudopregnancy) interrupting the regular estrous cycle, just as do females mated with sterile males (van der Lee & Boor, 1955, 1956). Cheal and Sprott (1971) argue that this disruption is due to the action of a pheromone. Sexual attraction may also involve a pheromone signalling system (Sienger, Agosta, O'Connell, Pfaffman, Bowen & Field, 1976).

Pheromones emitted by stressed animals are apparently also discriminated by conspecifics. For example, mice prefer the odor of male mice who had been victorious in an aggressive interaction in preference to that of defeated fighters, and also prefer the odor of isolated (non-fighting) individuals over that of losers (Carr, Martorano & Krames, 1970).

In a passive avoidance study, mice extinguished more slowly when exposed to the odor of subjects that were still receiving shock as compared to the odor of other animals on an extinction schedule (Sprott, 1969). They are also aversive to the odor of a stressed conspecific (Mueller-Velton, 1966), even though they are attracted to that of a nonstressed one (Rottman & Snowdown, 1973). Interestingly, this effect is unrelated to defecation and urination by the stressed animal. Rottman and Snowdown also found that removal of the olfactory mucosa abolished the aversion. While this clearly implies the olfactory action of the pheromone, environmental variables were also relevant; animals that were socially isolated for 12 weeks after weaning did not respond as did normals, even though they did emit the same odors (i.e., evoked the same response in normal animals). Thus, the hypothesis that pheromone-transmitted information is not necessarily acted upon was supported. With stressed rats, Lore
Blanc and Suedfeld (1971) showed that animals that had observed conspecifics which were learning an escape response, learned that response more rapidly. Again, the emission of a pheromone was proposed as a probable mediating variable. Stress may lead to the emission of a general signalling pheromone which indicates danger and may result in avoidance (Valenta & Rigby, 1968).

Pheromone research with a number of species and a variety of response patterns has further documented the importance of this communication channel. A detailed review was recently published by Thiesien and Rice (1976), covering research with mammals, and a newly published book (Shorey, 1976) also deals with the phenomenon. Since our interest here is primarily in cues that may be given off by human beings and detectable by dogs, the literature on human pheromones warrants a closer look at this point.

Pheromones in Human Beings

In addition to the repeated demonstration of the presence of pheromones in lower animals, further research shows the possibility that pheromones also could be present in human beings. For example, Comfort (1971) suggested that the source of pheromones in man generally appears to be the skin, but the axillary and pubic apocrine glands and the smegma are particularly important sites in both sexes. The axillary secretion appears to be the source of a social pheromone.

McClingtock (1971) reported a study with 135 female residents of a college dormitory. She found that there was a significant increase in synchronization of the onset of menstruation among roommates, and also that the extent of social association between the female subjects and male students affected the length of the menstrual cycle. McClingtock's conclusion
was that these effects were due to biochemical changes initiated by the action of primer pheromones.

Several authors have suggested that sexual processes in human beings, as well as in other animals, are communicated by smell. "The evidence appears quite convincing that odors are capable of activating hypothalamic-releasing factors which in turn allow pituitary release of trophic hormones affecting the hormonal secretions of the gonads and the adrenal cortex" (Schneider, 1974, p. 222).

A study concerning human social olfaction was reported by McBurnay, Levine and Cavanaugh (1977). A group of graduate students provided samples of body odor (sweated shirts). They were then told to rate the odor of each stimulus for pleasantness or unpleasantness, using a standard magnitude estimation procedure, and to evaluate the odor donors using bipolar adjective scales. They also attempted to identify their own odor. Results showed a high positive correlation between the rated unpleasantness of an odor and the social undesirability of the traits ascribed to the source of the odor. Subjects typically rated their own odor as the most pleasant, even though they demonstrated only marginal ability to identify it as their own.

Not enough research has been done in this area, and there are still many questions to answer about the possibility of human pheromones and their possible effects on behavior. Although there is some evidence of the presence of pheromones associated with particular emotions such as aggressiveness and fear in animals, to my knowledge there is no such research on human beings.
Canine Detection of Guilt: The Hypothesis

In summary, experienced field investigators in one of the world's outstanding police dog squads (Wright, 1976) strongly believe that the odor of human guilt reaction can be identified by the police dog. This odor is supposedly the product of biochemical processes associated with the emotions of fear, hostility and resentment, the emotions that guilty persons are purported to experience when confronting the police.

We have seen that the olfactory acuity of dogs is quite high. Analysis of their sensory mechanisms and behavioural experiments both indicate impressive sensitivity in distinguishing specific odors—including those emanating from human beings. They also have an impressive ability to follow and locate the sources of the scent, as shown by controlled experiments as well as police records.

It is also clear that in various species of animals olfactory cues are emitted as signs of specific emotional and biochemical states. The states associated with such secretions include fear and aggressiveness, both of which are related to the hypothetical odor of guilt as described above. While the existence of these specific pheromones in human beings has not yet been demonstrated, there is some evidence for human social olfaction and no reason to exclude the possibility that smell cues for fear and aggressiveness may also be emitted.

Thus, it appears worth while to test the hypothesis that police dogs can detect a special scent emanating from guilty individuals, identify the source, and communicate that identity to their handlers. The current study was designed to test that hypothesis. Of necessity, the "guilt" had to be produced in a simulation situation; but it was considered crucial
to use dog teams trained and experienced in the actual detection of criminal suspects (see Wright, 1964). As a minor point, comparability of performance across dog teams was included, since even if the general phenomenon were demonstrated it would be interesting to evaluate the existence and degree of individual differences in this context.

The absolute level of accuracy of the dog teams was, of course, one of the major dependent variables. But even supposing that the teams were not completely accurate—and we did not expect errorless performance—it would be important to see how their ability to detect guilt compared to other methods for achieving that goal. For this reason, the design of the experiment included another frequently used, controversial technique in the police repertoire: the polygraph examination.
The Polygraph

Polygraph examination was chosen as the second technique in this study for several reasons. Polygraphs are widely used by both police and private investigation organizations in North America, and in some ways their status resembles that of police dogs: that is, individuals using them are firmly convinced of their accuracy, but they have only tenuous legal standing and their employment may arouse public controversy. There is considerably more research on polygraphic lie detection, however, and the evidence attesting to the general high proportion of valid identifications using this technique is quite consistent. This makes the method an appropriate one against which to check the validity of police dog identifications. In addition, it was desirable to avoid approaches such as intensive interrogation, which would be stressful to the subjects. As a last point, even though there is a sizeable scientific literature testing polygraph validity, most simulation experiments have not used field investigators, equipment, and procedures. Thus, the current study could add to the relevant data on polygraph use besides utilizing the technique as a comparison baseline for the performance of the dog teams.

As has been mentioned, "lie-detection" by means of the polygraph is a very widely used technique. Many police departments use it as an adjunct to the interrogation of suspects and witnesses, and its use by private agencies in hiring personnel and interrogating employees has been spreading throughout North America. While in general polygraph evidence has had no legal standing in the criminal justice system, a recent
decision in a Vancouver court admitted polygraph results as interpreted by an expert psychologist to stand as evidence for consideration by the jury. This may be a precedent that will be followed by other courts; if so, the use and respectability of the "lie detector" will be likely to increase even further.

The validity of polygraph data must be judged on the basis of two kinds of evidence. One is represented by publications describing the results of practitioners in the field, while the other stems from systematic laboratory experiments. Our review will focus on these materials, and will omit as irrelevant the voluminous literature on such technical issues as instrumentation, details of scoring GSR records, and so on.

In his historical review, Horvath (1976) describes several antecedents of the modern polygraph. These include medieval "ordeals" such as the touching of hot irons to the tongue of the suspect. The guilty party was supposed to show burns, while the innocent remained unscathed. It may be argued that the actual perpetrator would experience anxiety, leading to diminished salivation, and thus to reduced protection from the heated object; and that the innocent suspects would have sufficient moisture in the mouth to minimize tissue damage. This of course, presupposes that they had so much faith in the efficacy of the technique that they felt no anxiety—an hypothesis that may also be relevant to the effects of present day lie detection.

The beginning of technological lie detecting was probably the use in 1895 of a hydromphymographby Lombroso and his student Mosso (Horvath, 1976). A combination of objective physiological measures and observation was used to detect lying. Munsterberg (1908) discussed the changes in
breathing, blood flow, skin conductivity and involuntary bodily movements that accompany attempts to deceive.

In 1914, Benussi and Moston began a series of laboratory experiments on the effects of lying on breathing and blood pressure. Even though those studies omitted relevant controls, they were crucial in pioneering this approach to the development of lie detection (Larson, 1969). From this beginning, Larson (1921) developed modified instrumentation and procedures to make possible the continuous recording of blood pressure, pulse rate, and respiration. Even today, most field examiners agree that the measurement of cardiovascular and respiratory activity is the crucial minimum of physiological recording in detecting deception (e.g., Ansley, 1972).

Most of the systematic research on the polygraph has been conducted by experimental psychologists rather than by practitioners. The most commonly used procedure involves the commission of a simulated crime, with half of the subjects in the guilty and half in the non-guilty groups. The crucial test of polygraph accuracy is the correct identification of the group to which each subject belongs by a blind operator (Horvath, 1976). Aside from the obvious question as to the external validity of the simulation situation, the research has many other dubious characteristics. Horvath's (1976) critical review lists among these the use of college students as the subject population, exclusive reliance on the galvanic skin response as the sole physiological measure used, testing by laboratory-trained research staff, and data analysis on the basis of specified objective criteria. While these factors increase the degree of control and statistical rigour, they make the generalizability of the results to
actual field procedures somewhat tenuous.

Summers (1939), one of the first researchers to use this approach, claimed that his technique was correct 98% of the time in separating "guilty," "innocent," and "accomplice" subjects in a crime situation. In a recent study, Barland and Raskin (1975) assigned 36 college students each to the guilty and innocent conditions in a simulated theft experiment. They were told that the "stolen" money ($10.00) would be theirs to keep if they convinced the polygraph examiner of their innocence. Subject expectancy was also controlled: 12 subjects in each condition were led to believe in the effectiveness of the polygraph, 12 in its lack of effectiveness, and 12 were given no expectancy induction. Both a strictly quantitative and a more general field method of scoring showed high accuracy, with 81% of the "conclusive" category assignments being correct. There was, however, a high proportion of "inconclusive" ratings. Five other examiners, who scored the protocols blind, obtained a mean inter-examiner correlation of .86 on total scores. Discriminative ability was statistically significant on all measures.

With abnormal subjects, the technique has shown mixed validity. Abrams and Weinstein (1974) found that they could detect truthfulness, but not lying, among borderline retardates (IQ 65 to 79). A high level of inconsistent reactivity was noted. The polygraph was also ineffective with a schizophrenic sample (Abrams, 1974).

By contrast, and with more relevance to field investigations, the polygraph appears to have utility with psychopathic subjects. Imprisoned psychopaths were put through the standard simulated theft situation by
Raskin & Hare (1977). With a trained psychologist serving as an evaluator, 88% of the subjects were correctly categorized as "guilty or innocent" (96% if "inconclusives" are omitted). Accuracy was about the same as for a sample of nonpsychopathic prisoners, contradicting the general belief that psychopaths either do not feel guilty to the same extent as, or can suppress the signs of guilt more thoroughly than, normal individuals. In another prison study, Lieblich, Ben Shankhar and Kugelmass (1976) used only the GSR as contrasted to Raskin's utilization of GSR, cardiovascular, and respiration measures. The ability of the rater to match GSR changes with the responses of individual prisoners to previously administered questionnaire items was significant at the .01 level.

It may be argued that the simulated crime paradigm is too artificial, although this may be interpreted as enhancing the impact of the experimental results since it demonstrates polygraph accuracy in even a low-stress situation. At any rate, two studies in which the need to deceive the lie detector was more pressing had compatible outcomes. In one, police trainees were led to believe that successful deception was important to their evaluation (Kugelmass & Lieblich, 1966); in another the tests were given to actual criminal suspects in the course of the investigation (Kugelmass, Lieblich, Ben Ishai, Opatowski and Kaplan, 1968). GSR responses, but not heart rate, were found to reveal attempts at deception.

The accuracy of practitioners in the field has also been evaluated, with positive results. Horvath and Reid (1971) obtained polygraph recordings collected in field investigations conducted by Horvath.
records were used, of which 20 were verified as innocent by the confession of another suspect. Ten polygraph examiners employed by a commercial agency were instructed to separate guilty and innocent suspects on the basis of the physiological records. Overall, the categorizations were correct in 88% of the cases. Accuracy ranged from 79% for relatively inexperienced operators (4 to 6 months of training) to 91% for the more experienced (at least one year of experience after having completed the training program). Similarly, Hunter and Ash (1973) found 86% accuracy with seven operators rating truthful and deceptive polygraph records.

Bersh (1969) collected polygraph and other evidence from military court-martial cases, and had the non-polygraph evidence evaluated independently by four experienced lawyers. These evaluators were asked to indicate whether the suspect was guilty or not on the basis of evidence, ignoring technicalities. In over 92% of those cases in which the legal judgement was unanimous, it agreed with the polygraph results; when the lawyers had only a majority determination, agreement was 75%. Polygraph agreement with the combined unanimous and majority judgements was 88%.

In a replication and extension, Barland (1975) checked his own polygraph decisions with the opinion of the panel of five lawyers and the judicial outcome; there was agreement in 90% of the cases; all disagreements were in instances where the court had acquitted the suspect. With the majority of the legal panel, polygraph agreement reached 79%; in cases where at least four of five lawyers agreed, consistency with polygraph was 87%.

To investigate the reliability of the technique separately from validity, Rouke (1941) ran a simulated crime situation involving 80 delinquent and 90 non-delinquent boys. GSR data showed high reliability.
over repeated scorings by the same operator as well as high interjudge reliability (approximately 90% agreement). In another reliability test, Barland (1972) calculated correlations between all possible pairs among independent judges rating GSR measures according to a numerical scoring system proposed by Backster (1969). On the records of 72 subjects in a simulated crime situation, the correlations ranged from .78 to .95, with a mean of .86. Thus, laboratory-trained scorers have shown high reliability in evaluating records.

These findings support the use of the polygraph in this study. There was a need for a relatively valid and frequently used guilty-detection technique against which dog team performance could be compared. The literature on polygraph accuracy demonstrates that this method fulfills the requirements.

**Summary**

In view of the demonstrated olfactory acuity of dogs, the likelihood that specific emotional states in human beings produce different patterns of body odor, and the widespread use of police canine dog corps, the belief of police officers that trained dogs can distinguish between guilty and innocent individuals calls for an empirical test. No such research has been available to evaluate the accuracy of the police claim. The experiment described here was meant to provide such an evaluation assessing the ability of experienced police dog teams to identify guilty suspects in a simulation situation more accurately than chance and comparing that ability with that of the more firmly validated polygraph examination.
METHOD

Participants

Subjects

Advertisements were posted and placed in the university newspaper, offering $5.00 for participation in an experiment lasting about one and a half hours. The payment for participation was indicated as the minimum that would be earned, but subjects could win an additional $10.00 depending upon their performance. In order to reduce extraneous sources of variability, and compatible with the predominance of males among arrested suspects (Laughy, Personal communication, October, 1976) only male volunteers were invited. A total of 64 respondents constituted the final sample.

Dog Teams

Contact was made with the Vancouver Police Dog Squad through a scientist interested in olfaction who has been a consultant to the Squad. After prolonged discussion and several meetings, the officer in charge designated three dog teams, including himself and his dog, to participate in the study. These were among the most experienced teams available. Each team consisted of a police dogmaster and a male German Shepherd dog, fully trained and with a record of success in extensive field work. Participation was voluntary for these dogmasters. Only one team was active at any particular time. Instructions to dogmasters are reproduced in Appendix A.

Polygraph Operators

Of several experienced polygraph operators who were contacted, two
were available and willing to participate. Each of them has his own polygraph testing service after long experience as a polygraph operator for the RCMP. The equipment used consisted of portable Stolting polygraph machines, each with four channels; two pneumographs, one galvanic skin response and an electronically enhanced cardiograph. Three charts were completed for each subject (for more details see Appendix B, p. 1). Both polygraph operators were used simultaneously.

Procedure

General Orientation

Subjects were scheduled so as to appear at the laboratory two at a time. Upon reporting to the laboratory, the pair of subjects was met by a female assistant who, after making them comfortable, explained that the experiment was concerned with the way in which police personnel made decisions about suspects. They were told that a team from the Police Dog Squad and a polygraph examination would be involved, but that the dog would be leashed and there would be no physical contact between the subject and the "investigators". Furthermore, there would be no harassment, intensive interrogation or any other sort of unpleasant or demeaning interaction. Subjects were also told that they could decide to end their participation in the experiment at any time (including during the orientation), would receive their $5.00, and would not be persuaded to continue.

At this point, the subjects were given a typed summary of the foregoing information and a consent slip. All subjects agreed to continue. The rest of the orientation was given individually since it varied as a function of experimental condition. Neither subject was permitted to
hear the specific instructions given to the other, nor to interact with him until after the end of the experimental session.

**Conditions**

One of the subjects in each pair was randomly assigned to the Guilty and the other to the Not Guilty condition (see below).

*Simulated Crime Instructions:* Subjects in the Guilty condition, after being briefed for the experiment in general, were told that one of their tasks was to commit a simulated crime. This consisted of going to the second floor of the building where the experiment was conducted, entering a designated office that had been left unlocked and taking an envelope containing a $10.00 bill that had been left in a location specified by the experimenter. The subject was to hide the envelope and money on his person through the remainder of the study.

Guilty subjects were informed that there would be two attempts to identify them as having committed the "crime;" one using a polygraph and one using a police dog team. The subject's task would be to convince both investigators of his innocence; he was to continue claiming to be innocent throughout the rest of the experiment regardless of what the polygraph operator and the dogmaster said or did. If both investigators accepted this claim, the subject would get to keep the $10.00 in the envelope. (See Appendix C for verbatim protocol).

Not Guilty subjects were told only that a police dog team and a polygraph operator would try to find out if they had committed a misdeed. They were to maintain their innocence and would win $10.00 extra if the investigators accepted their claim (verbatim instructions appear in Appendix D).
Design

Of the 64 subjects, half were randomly assigned to each of the two experimental conditions. Each of those two groups was in turn randomly divided into three subgroups, with \( n = 10, 10 \) and 12. One subgroup was run by each of the three dog teams, a procedure recommended by the officer in charge of the Dog Squad to avoid habituation on the part of the subject or the dog teams, to prevent excessive reliance on the talents of any one team and to facilitate the scheduling attempts of the Dog Squad to provide teams when needed.

Every subject underwent two "investigations;" one by the dog team and one by the polygraph operator. The order of these two events was counterbalanced so that equal numbers of Guilty and Not Guilty subjects encountered each investigation first.

Physical Facilities

To avoid habituating the dogs, the experiment was run in similar facilities in three locations on the campus of the University of British Columbia. The areas used were as follows:

1. Two offices for briefing the subjects.
2. One office in which the simulated crime was committed.
3. Two small seminar rooms in which the polygraph examiners performed their tests.
4. A stretch of corridor between rooms. The rooms were used as places of concealment by the subjects and the dog team walked along the corridor during its phase of the procedure. This area was around the corner of an L-shaped hall so that neither the dogmaster nor the polygraph investigator could tell which subject was judged guilty by the other technique.
Dog Test Procedure

In the dog test the subject was shown the stretch of corridor mentioned previously and was told to hide in one of the rooms after the experimenter had left so that no one knew exactly where he was. He was instructed to hide behind the curtain, out of sight from the door, and to remain there until either the dog team or the experimenter came to get him. A delay of approximately 15 minutes was imposed before the dog team went into action to allow the subject's odor to reach the hallway. At that time, the leash was removed and the dogmaster gave the command to begin searching. When he decided that the dog had located a "suspect" he replaced the leash and entered the classroom indicated. If no subject was hidden there, the search resumed; if a subject was found, the dogmaster spoke and acted toward him in a friendly way so that the dog calmed down. The subject was then escorted either to the briefing room or to the polygraph examination room by the experimenter and the procedure was repeated with the second subject. At the end of this search the dogmaster indicated to the experimenter which subject was "guilty".

Polygraph Test Procedure

Each polygraph examination consisted of three components: these were a pre-test procedure, the testing, and the post-test. The pre-test was designed to relax the subject, explain what the instrument was about, and develop information that enabled the examiner to assess the individual's physical and mental health, motivation and general suitability for examination. The subject was given the opportunity to confess that he was guilty; however, he did not undergo any intense interrogation. During a
discussion, the questions that were going to be used in the examination were phrased in the language of the examinee, and were ascertained to be relevant to the issue or the technique and not ambiguous. Once the subject agreed to continue, the questions were rehearsed twice before the examination proper was conducted.

At that time, the examinee was attached to the polygraph and the list of questions agreed upon was repeated three times with each series recorded on a separate chart. The charts were then analyzed and an opinion was formed by the examiner.

When the examiner felt that the subject was not guilty, there was a brief post-test discussion. When the subject was suspected of deception, he was informed of the examiner's opinion and invited to offer an explanation.

After the post-test, the subject departed. The examiner then went over his charts and notes, and informed the experimenter whether in his opinion that subject had been in the Guilty or Not Guilty condition.

A detailed description of diverse varieties of the polygraph procedure and scoring techniques are given in Appendix F. Specific examples of the types of questions used in the test are given in Appendix B.

Order of Events

After the original briefing, the Guilty subject was escorted upstairs where he committed the simulated crime. In the meantime, the Not Guilty subject remained in the briefing area until this was over, and was then escorted to begin either the polygraph examination or the dog test. After taking the money, the Guilty subject was also escorted to the appropriate testing area.
At the end of testing, both subjects switched to the other test procedure.

When that was completed, the subjects were escorted to the briefing area and debriefed individually. They were told whether they had won the $10. offered for convincing both investigators of innocence in the matter of the stolen money, and were given either $5 or $15. Any questions were answered, they were assured that they had done well, and they were thanked for their help. Their address was recorded if they were interested in receiving information about the results of the study.

Summary of Variables

A. Independent variables:

In the present study the following were considered as independent variables:

1. Simulated crime situation: Participants who took the money (Guilty) vs those who had no knowledge of the crime situation (Not Guilty).


B. Dependent variables:

Success or failure of the dog teams and the polygraph tester (number of errors in identification).
RESULTS

Data analyses were concerned with three aspects of the results. The first of these was the level of accuracy, relative to chance, of the two techniques in identifying Guilty and Not Guilty subjects. Next, it was important to evaluate the performance of the dog teams compared to that of the polygraph operators. And last was the issue of whether there were individual differences in accuracy among the three dog teams.

The nature of the procedure and of the dependent measures imposed some restrictions on the analysis. Because the data were in the form of assignments to the Guilty or the Not Guilty category, it was not possible to use analysis of variance nor similar parametric tests; instead, tests of proportions (Glass & Stanley, 1970, pp. 326-328) were used. Another problem arose from the fact that subjects were run individually in the polygraph examination part of the study, but in pairs during the dog team phase. In the latter procedure, if the dogmaster identified one subject as Guilty, he automatically classified the other as Not Guilty. This fact reduces the degrees of freedom in the dog team condition, calling for an analysis by pairs. This was not the case for the polygraph operators, who had a choice of Guilty or Not Guilty, for each individual examination. The data analyses reflect this difference.

The proportion of correct identifications made by the dog team was .47. Compared to the .50 to be expected by chance alone, this yields a $z$ of -.35, indicating by a test of independent proportions that the accuracy was at about chance level. The proportion of correct identifications for the polygraph examinations was .91, $z = 4.59$, $p < .01$, indicating per-
formance significantly above chance. This latter result is also obviously of practical significance.

A comparison of the two techniques with each other showed 15 correct answers out of 32 pairs of subjects for the dog teams and 28 correct for the polygraph operators. This difference was statistically significant by a test of dependent proportions, \( z = 3.36, p < .01 \) (see Table 1). The proportion of correct identifications was not affected by whether the subject had actually been in the Guilty or Not Guilty condition.

Looking at categorizations of individual subjects rather than pairs (Table 2) we find that the polygraph operators were correct in 91% of the cases. Again, the same number of errors occurred with actually Guilty as with Not Guilty subjects. As noted above, the dog teams—with their slightly different procedure that required only one real decision for each two subjects (but leading to two incorrect assignments if that one was wrong)—scored 47% correct.

As to individual differences among the dog teams, their percentages of correct identifications were 40%, 45%, and 54%. The differences among them did not reach statistical significance by Marascuillo's (1966) test \( \chi^2 (2) = 0.46, \text{NS.} \)
TABLE 1

Performance Accuracy (Pairs of Subjects)

DOGS

<table>
<thead>
<tr>
<th>POLYGRAPH</th>
<th>Not Correct</th>
<th>Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Not-Correct</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

17          15
<table>
<thead>
<tr>
<th>Method</th>
<th>Correct</th>
<th>Incorrect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dogs</td>
<td>30</td>
<td>34</td>
</tr>
<tr>
<td>Polygraph</td>
<td>58</td>
<td>6</td>
</tr>
</tbody>
</table>

TABLE 2

Performance Accuracy (Individual Subjects, N = 64)
DISCUSSION

The most striking results of this study were the failure of the police dog teams, and the success of the polygraph operators, in reliably identifying the Guilty and the Not Guilty subjects. Dog teams performed at only chance level, making correct identification in approximately one half of the cases. Since each pair of subjects was known to include one person in each category, flipping a coin would presumably have produced equally good results.

Obviously, this finding fails to confirm the belief that trained police dogs can recognize the "odor of guilt" and that they can communicate to their masters the identity of individuals emitting that odor. The police officers who worked with us on the study, including the sergeant in charge of the Dog Squad, attributed this failure to the artificiality of the experimental situation. Several other explanations for the failure are also feasible.

First, of course, is the possibility that dogs cannot in fact detect any distinctive guilt odor, or at least cannot communicate any such detection to their handlers. There is no question that police dogs are demonstrably effective in other situations, including tracking and building searches. In this particular study, all of the concealed subjects were located in a fairly brief period of time, on the average 20 min. for each pair. But these are different tasks from picking out the guilty one of two subjects; in tracking, the dog follows only one spoor, as in building searches he locates one person. In each case, the one individual thus
identified is usually accepted by the master as a valid suspect. The dog may learn to follow a smell and treat its origin as his quarry, without the hypothesized "fear, resentment and hostility" pattern characterizing the scent. Thus, the usefulness of police dogs in the usual field situation is not incompatible with their failure in the experiment.

An alternative explanation, favoured by our police participants, is that the situation failed to simulate the real-life situation in an adequate manner. Guilt odor may occur when a real criminal is confronted by the possibility that, due to the activity of the dog team, he will be located, identified, and in due course punished. The guilty subject in the current study, on the other hand, knew that his act had not been criminal: it took place in an experimental setting, the money was left deliberately, and he had been informed of its "hiding place" and instructed to take it. Furthermore, it was clear that he would not be punished even if he were identified as guilty. True, he stood to lose the $10 bonus, but he would still receive $5, and the knowledge that he had contributed to the completion of the study. The situation was hardly equivalent to that of a criminal facing arrest. The subject's motivational state might include interest, excitement, competitiveness, avarice and hope, but it was unlikely to be characterized by much fear or anger. If, as the police contend, the guilt smell represents an amalgam of the latter two emotions, it may be unreasonable to expect to find such a smell in the simulated crime setting.

One must remember, though, that this analysis of the guilt odor is a purely suppositional one on the part of the police dogmasters, and that there is no empirical evidence to support the significance of the
hostility-fear-resentment triad. It may be that guilty conscience, the knowledge of having committed a criminal or forbidden act, is sufficient to elicit the odor. Even in this case, the experimental situation may have been too artificial in view of the characteristics outlined above. Thus, the failure of the dogs to detect a guilt odor may merely reflect the inability of the situation to activate such an odor regardless of what the components of the scent may be.

One counterargument to this explanation relates to the polygraph results, which will themselves be discussed later. Clearly the simulation procedure was strong enough to produce in the guilty subjects physiological reactions that were identifiable by the polygrapher. It is feasible, however, that a situation that produced detectable changes in respiration and blood pressure may not have triggered the hormonal changes leading to a smell cue.

Another possible explanation is that the artificiality of the situation affected the dog team, as well as (or rather than) the subject. While the dogmasters were very cooperative, there may be something about their behavior on a real investigation—gestures, stance, voice, facial expression, and even smell—that communicates urgency, expectation, seriousness, danger and so on to the dog, and to which the dog responds. These cues may have been absent during the experimental runs, and their absence may have affected the dog's reactivity to the situation and to whatever olfactory stimuli were emanating from the subjects. The dogmasters reported that they felt more as though this were a training situation than a field situation, and this attitude could have been sensed by the dogs.
This "training" reaction would have been compatible with the physical conditions of the experiment. In real crimes, the dog team searches different buildings, locations, and areas from incident to incident; repeated runs over the same ground occur mostly during training. Because of the need for comparable experimental environments, the study was conducted in only three places on campus, so that each dog team repeatedly searched in each locale. This would reinforce the simulation artificiality of the experiment.

Some other differences between our procedure and the real field investigation procedure also appeared. To safeguard our subjects, we asked that the dog be leashed in the hallway as soon as he identified the door of the room in which a subject was hidden. In an actual building search, the dog is let loose to find and immobilize the suspect, and is typically free to move around the building for quite some time. The leash is put on only after the suspect has been taken into custody, searched, possibly handcuffed, etc. The restraint imposed by the leash was unusual, marking this particular search as abnormal. Not only that, but the restraint made the dogs more aggressive when they did locate a suspect, thus possibly misleading the police officers.

These differences in dog handling procedure could explain an anomaly noted in the course of the study. Beforehand, the dogmasters said that they knew when the dog had found a guilty person by the intensity of the response: lunging, barking and the like. In pilot runs, the masters did assess the intensity of aggressive behavior before coming to a decision about the guilt of the subject. But during the experiment itself, this
criterion was tacitly de-emphasized. Instead, the dogmasters almost always picked as guilty whichever subject the dog found first. In response to question, the handlers argued that the Guilty subject probably gave off a stronger odor, leading to earlier detection. This is contrary to their earlier description of how to read the dogs' reaction to guilt. Furthermore, it seems to abandon the hypothesis that the guilty smell is qualitatively different from—not merely more intense than—the odors emitted by innocent bystanders.

Interestingly, on seven occasions the members of the research team who were observing the dog judged that his reaction to the Guilty subject was in fact more intense—only to have the police dogmaster, following the primary criterion, make the wrong choice. Since this was completely unforeseen (given the verbal adherence of dogmasters to reactivity as the relevant factor) no systematic observations were made or recorded. In any future study, it would be desirable to have intensity of the dogs' reaction rated independently by members of the research team and perhaps by a nonparticipating dogmaster or dog expert.

One other possible factor in the low accuracy of the dog teams may have been the result of our interest in the ability of the dog to detect guilt by smell. In real-life situations, there are auditory and visual cues to which the police dogs may also respond. Some facial expressions, physical stances and gestures, movements and sounds may be used to identify a guilty suspect (see Lorenz, 1952). These were largely eliminated in this study. Here the suspects were behind a closed door and drawn curtain, so that the differentiation between guilty and not guilty subjects was made essentially on the basis of the smell alone. This was of course,
appropriate given the focus of the study, but it did decrease both
the realism of the procedure and the information base on which it was
made.

In contrast to the dogs, the performance of the polygraph operators
was extremely accurate. Correct identification in over 90% of the trials
supported the usefulness of this technique even in a simulated crime
situation. The implication is that polygraph accuracy with real suspects,
where physiological reactivity may be even stronger, is likely to be of
at least a comparable level, attesting to the usefulness of the approach.
It should be noted that in this study, the polygraph examination was not
administered, scored, or interpreted by the sophisticated scientific per-
sonnel and techniques typically used in simulation tests (e.g., Barland & Raskin
1975 & Raskin & Hare, 1977). Rather the testers were actual practitioners from
investigative organizations, using their standard procedures as normally
applied in the field. This meets one of the criticisms offered by some
reviewers (e.g., Horvath, 1976), that laboratory attempts to validate the
polygraph technique involve personnel and procedures that are so far
removed from actual use conditions as to make external validity highly
doubtful. Some unrealistic features still existed, of course. These in-
cluded the brief time lag between the "crime" and the examination, the
knowledge that half of the subjects were guilty, and the fact that all
suspects were university students.

The pre-test interview, the demonstration to convince subjects of
the effectiveness of the procedure, the running of the three charts per
subject and the post-test interview, took approximately 1 hour and 45
minutes per case. Evaluation of the data by the Backster number scoring procedure took varying periods depending upon the consistency of the pattern, but in general 2-1/2 hours covered the total expenditure on each subject. This seems to show that the polygraph is economical in time, which, in view of its accuracy, makes it a fairly effective investigative technique.

Future research may take several directions. One of these is related to the performance of the dog teams. It is difficult to avoid some artificiality in the test situation. For example, research ethics make it doubtful that an unleashed search procedure can be instituted since it is known that in real search situations, while it is not common, an excited and unleashed dog sometimes bites or jumps upon a suspect. More realistic role-playing by the dogmasters would help but may be difficult to elicit. The restoration of visual cues would make the situation more realistic, but would not provide an adequate test of the hypotheses related to olfaction.

Also, the whole scenario could be made more realistic. For example, the dog could first track the suspect, perhaps by being exposed to some object that the subject had handled; the area within which the person was hiding could be enlarged; only one suspect could be hidden at a time, without the pairing of one guilty and one not-guilty person each time. In fact, the original design of the current study involved a series of steps, starting with the scent marking of objects by the two "suspects;" followed by the tracking and then the detection phase. However, this plan was abandoned, on the suggestion of the Dog Squad Commander who felt that it would involve a great deal of time without commensurate usefulness.

Obviously, it would be useful if the police made tentative identifi-
cations of guilt or innocence among the subjects located by the dogs and compared this polygraph rating and the outcome of trials. Given the established validity of polygraph examinations, they could be used as a criterion measure against which other sources of judgment can be compared. Legal and ethical problems may arise, however, and would have to be handled carefully.

There seems to be little empirical reason to doubt the power of the polygraph technique, even though philosophical and ideological objections to its use may still be raised. As to the use of police dogs, this study was not designed to assess their reported efficacy for crowd control, drug detection, tracking, or apprehension of suspects. But their ability to identify guilt by odor alone, and/or to communicate such an identification to the dogmaster, was not confirmed. Some skepticism in this area is therefore justified at least until more supportive data are obtained.
References


Campbell, P. *Personal communication*, August 8, 1977.


Larson, J.A. Modification of the Marston Deception test. *Journal of the American Institute of Criminal Law and Criminology*, 1921, 12, 390-399.


a variety of animal behaviors. Psychological Reports, 1971, 29, 195-243 (Monograph Supplement, 1, V. 29).


INSTRUCTIONS TO DOGMASTERS

Only one dog and his master should be in action at any given time.

Your task is to find the two volunteer subjects who are hiding in different rooms on this floor. You should not at any time release your dog from the leash except in the corridor.

When you find each subject, he will be standing behind a curtain. Have your dog leashed as you enter the room; do not allow him to come into contact with the subject; and do not ask the subject to come out from behind the curtain. Also, please do not interrogate the subject at all, nor speak to him about anything.

When both subjects have been located, we will ask you to indicate on the basis of your dog's behaviour which one is guilty of a simulated crime that occurred before they hid in the rooms. Remember that in each pair of subjects one is guilty and one is completely innocent.
APPENDIX B
Polygraph Questions

Eight questions were used during the experiment, each with a specific purpose.

1. **Irrelevant**: Designed not to have any emotional content. It is used as the first question to introduce the examination and may be repeated at the discretion of the examiner should there be any anomalies in the charts.
   Examples of the type of question: a) Is today Wednesday? (the day of the examination). b) Is your name David? (the name of the subject).

2. **Weak Relevant**: Designed to absorb the emotional responses resulting from the first accusatory relevant questions. Type of question: Regarding the $10. taken from room #_____ this afternoon, do you intend to answer truthfully each question about that?

3. **Symptomatic**: To determine whether the examiner has developed a bond of confidence with the examinee.
   Type of question: Are you completely convinced that I will not ask you a question on this test that has not already been reviewed?

4. **Control**: Designed to have a similar emotional impact on the non-deceptive examinee as the relevant questions would have on the deceptive examinee.
   Type of questions: #4. Between the ages of 17 and 20 did you ever take something that you knew you shouldn't have?
   #6. During the first 16 years of your life did you ever take some-
Appendix B (cont'd)

5. **Relevant.** Designed to be relevant to the issue, a crucial question posed about the facts.

   Question #5. Did you steal that money?
   
   Question #7. Did you steal that money from room #____?

8. **Symptomatic:** Designed to establish whether there is an outside issue which would have an emotional impact that would be more powerful than either the relevant or the control questions and therefore invalidate the examination. Type of question: Is there something else you are afraid I will ask you about, even though I have told you I will not?
A. Instructions to the "Guilty" subjects.

As we told you, the study has to do with police behaviour. Basically, we are interested in how effective the police are in identifying suspects who are actually guilty of some kind of misdeed. So what we are doing is setting up a simulated crime situation, and then letting some real policemen have a chance at identifying whether a particular person committed it or not. What we would like you to do is to go into Dr. _______ office, Room #_____. You will be left alone there for a while. There is an envelope behind one of the books on the 6th shelf from the top in the left-hand bookcase. This envelope contains a $10 bill which when you find it, you should hide on your person. Do this when nobody is looking at you, and hide the money as well as you can. Your task will be to prevent being identified as the guilty person who "stole" the money for the rest of the study. There will be two attempts to identify you as guilty; neither will involve searching you, giving you intensive interrogation, or any other unpleasant procedures. If you succeed in deceiving the people who are trying to detect your guilt, you will get to keep the $10 you took in addition to the $5 you are getting for participating in the study. Remember that it is very important that you continue to appear innocent throughout the entire study. Don't confess under any circumstances, and try as hard as you can to persuade the interrogators that you in fact did not take the money. Keep on trying to convince them of this even if you think that they don't believe you and that they have identified you as the guilty person. Remember that if you were a real criminal you would continue to deny your guilt in order to
try to escape punishment; in this case if you manage to convince them that you are not guilty, you stand to gain $10."

"Your payment depends on how well you do in the task. If in both of the two attempts to establish your guilt you are able to fool the police (that is, to make them think that you are in fact not guilty), you will get to keep the $10 you took in addition to the $5 you will be paid for participating in the study. I want to remind you again that you can't possibly win the extra $10 if you don't continue to maintain your innocence as convincingly as you can. If either of the attempts identify you as guilty, you will have lost the money. Even if you think that one or both of the people that you are dealing with think that you are a guilty person, if you act innocent well enough they might change their minds and you might get away with it."

B. Dog Condition:

"After you have taken the money and hidden it on your person, we would like you to go down the corridor, turn left, and hide in any of the open rooms that you see there. After you enter the room, make sure that you close the door firmly. The police will then attempt to find you, so draw the curtain across the window and hide behind the curtain. Do not come out until either the police or I come to get you. Remember that the police will not touch you, nor hurt you in any way, although they will try hard to find you. Once you are hidden, stay there and do not try to change your hiding place. The police may use dogs or other equipment to try to find you; if a dog is used, he will be on a leash and controlled by his dog master at all times, so that you don't have to worry about being
Appendix C (cont'd)

attacked or hurt. However, if you have a very high fear of dogs, please let us know so that we can make sure that this procedure is not used with you."

C. Lie detection Condition

Subjects will go to a lie detector room and be interrogated.
APPENDIX D
A. Instructions to the Not Guilty subjects.

As we told you, the study has to do with police behaviour. Basically, we are interested in how effective the police are in identifying suspects who are actually guilty of some kind of misdeed. So what we are doing is setting up a simulated crime situation, and then letting some real policemen have a chance at identifying whether a particular person committed it or not; you will be innocent of the particular act.

Your payment depends on how you do in the experiment. The police will make two attempts to discover whether you have committed a particular act. If you are able in both cases to convince them of the fact that you did not commit the act, you will get $10 bonus in addition to the $5 you will be paid for participating in the study.

B. Dog Condition

We would like you to go to the corridor, turn left, and hide in any of the open rooms that you see there. After you enter the room, make sure that you close the door firmly. The police will then attempt to find you, so draw the curtain across the window and hide behind the curtain. Do not come out from behind the curtain until either the police or experimenter come to get you. Remember that the police will not touch you, nor hurt you in any way, although they will try hard to find you. Once you are hidden, stay there and do not try to change your hiding place. The police may use dogs or other equipment to try to find you: if a dog is used, he will be on a leash and controlled by his dog master at all times, so that you don't have to worry about being attacked or hurt. However, if you have a very high fear of dogs, please let us know so that we can make
Appendix D (cont'd)

sure this procedure is not used with you.

C. Lie Detection Condition

Subjects go to the lie detector room to be interrogated.
Polygraph Examination Technique

With the increasing development and use of the polygraph different methods of conducting the examination have been applied according to Horvath (1976). Among the most frequently used have been:

1. Relevant-Irrelevant: During a preliminary interview, the examiner assesses the emotional accessibility of the subject and decides on the length of the test. During the polygraph examination, he asks a series of questions that combines items relating to the crime with others that are irrelevant or non-critical. Advocates of this technique argue that truthful subjects will respond identically to relevant and irrelevant questions, while deceptive respondents will react differentially. This technique has been criticized on the grounds that it does not provide adequate controls against which to compare the "critical" reaction.

2. Control Question: To the preliminary interview and polygraph test some examiners add a post-test interrogation. The major difference between this and the previous technique lies in the kinds of questions asked during the test and in the method of data analysis. Here, irrelevant questions are used to establish patterns associated with given truthful information, e.g., the subject's name. Relevant questions are related to the investigation. A third category, control questions, deal with sensitivity issues not directly bearing upon the investigation (e.g., "Have you ever stolen anything?"). The purpose of the control questions is to focus the attention of honest subjects away from relevant questions, while lying subjects maintain a psychological set centered on relevant questions and perceive
Appendix E (cont'd)

the control items as relating to them. This procedure permits each subject to act as his own control, and has increased the validity of the polygraph examination while reducing the proportion of inconclusive results (Horvath, 1976). Since its inception (Reid, 1947) it has become the most common technique used by field practitioners.

Variations of these techniques, used in more restricted situations include the Mixed Questions procedure, where the questions on an original Control Question list are asked repeatedly in different order (Reid & Inbau, 1966), and the Silent Answer test, in which the subject gives no overt verbal answer (Horvath & Reid, 1972).

More extensive differences from the first two major techniques characterize the following procedures:

3. Affirmation: Here, the subject is given a second administration of the original list, with instructions to answer "Yes" to all questions including those to which he had previously answered "No". The point of this is to discover whether subject is deliberately attempting to distort the polygraph measures. Of course, the answers on the second examination are not interpreted in the same way as those on the original run-through (Reid and Inbau, 1966).

4. Peak of Tension (POT): An alternative procedure frequently used by critics of the Relevant-Irrelevant technique and the Control Question technique. One version is the searching POT in which the examiner asks sets of similar questions with each set concentrating on a particular point such as the location of the weapon involved in a crime until he feels that that particular point is resolved (Arther, 1967). In the known solution
Appendix E (cont'd)

POT, the questions include some to which the answer is already known to the examiner, but would not be known to an innocent subject. Reactivity to supposedly unknown items of information is interpreted as a sign of deception (Arther, 1968). In general, POT testers consider that the polygraph tracing will "peak" at critical items if the subject has guilty knowledge or is attempting to deceive the examiner (Arther, 1967; Reid and Inbau, 1966).

Since the Control Question technique is the most common in field examination, and was used in the study described here, it is useful to describe the procedure in more detail.

The first step in the session is the pre-test interview. The examiner discusses with the respondent the nature of the examination, explains the equipment, and attempts to establish a state of rapport. He also identifies the matters to be investigated and develops a list of questions that the respondent fully understands, accepts, and helps to phrase. Some examiners also ask questions about the subject's personal background, medical history etc. The exact approach differs among examiners (Barland & Raskin, 1973; Horvath, 1973). At the end of this preliminary interview the examiner impresses the subject with the power of the polygraph by identifying through the instrument a number on a card that the subject is thinking about but attempts to conceal (Barland & Raskin, 1975; Reid & Inbau, 1966).

The second component is the polygraph examination itself. There is a list of ten to eleven questions previously discussed with the subject and including relevant, irrelevant and control items as described above. The interim interval is approximately 20 seconds.
A second run-through of the original questions is then performed, with such changes in phraseology or emphasis as the subject may suggest. A third administration is then performed. At its conclusion, the examiner looks over the charts, and decides whether any further testing is needed. If it is, one or more of the more specialized procedures described previously may be used. The judgement of the examiner will dictate his choice of procedures.

Analysis of Polygraph Records

Texts and training manuals are used as guidelines for interpreting the changes in the skin conductance, breathing, and blood flow recorded on the charts. The magnitude and duration of reactivity to each question on each chart are evaluated (Backster, 1969). Field investigators realise that deception reactions differ from case to case, so that there is no specific pattern to identify but rather deviations from baseline levels for the particular individual. The important criterion is that the changes occur consistently as compared to other changes in the record (Horvath, 1976).

The study reported here used Backster's (1969) quantitative scoring system which the differences in reactions to paired relevant and control questions are assigned a score $-3$ to $+3$. The examiner then makes his decision and magnitude of these scores. This procedure was used in the current study.