Who Can Catch a Liar?

Paul Ekman — *University of California, San Francisco*
Maureen O'Sullivan — *University of San Francisco*

The ability to detect lying was evaluated in 509 people including law-enforcement personnel, such as members of the U.S. Secret Service, Central Intelligence Agency, Federal Bureau of Investigation, National Security Agency, Drug Enforcement Agency, California police and judges, as well as psychiatrists, college students, and working adults. A videotape showed 10 people who were either lying or telling the truth in describing their feelings. Only the Secret Service performed better than chance, and they were significantly more accurate than all of the other groups. When occupational group was disregarded, it was found that those who were accurate apparently used different behavioral clues and had different skills than those who were inaccurate.

Lies occur in many arenas of life, including the home, school, and workplace, as well as such special contexts as in police interrogations and courtroom testimony. In low-stake lies, the liar suffers no more than embarrassment if caught, but in a high-stake lie, the consequences for success or failure may be enormous for both the liar and the liar's target. Examples of such high-stake lies include those between heads of state during crises, spousal lies about infidelity, the betrayal of secrets through espionage, and the range of lies involved in perpetrating various crimes.

Lies fail for many reasons. The lie may be exposed by facts that contradict the lie or by a third party who betrays the liar's confidence. Sometimes, such outside information is not available or is ambiguous. Then the lie succeeds or fails solely, or primarily, on the basis of the liar's behavior, which the legal profession terms *demeanor* (see Ekman, 1985, for a discussion of different forms of lying, the role of stakes in the detection of deceit, and why lies fail or succeed).

Two types of errors may occur when truthfulness based on demeanor is judged: In a false negative, a liar is incorrectly judged to be truthful; in a false positive, a truthful person is incorrectly judged to be lying. In a high-stake lie, either type of mistake can have serious consequences. In dealing with such situations it would be important—for the clinician, the jurist, the businessman, the counterintelligence agent, and so on—to know how much confidence should be placed in judgments based on demeanor, by layman or expert, about whether someone is lying or telling the truth.

The answer from 20 years of research is "not much." In every study reported, people have not been very accurate in judging when someone is lying. In the usual study, observers are given video or audiotapes and are asked to judge whether each of a number of people is lying or telling the truth. Average accuracy in detecting deceit has rarely been above 60% (with chance being 50%), and some groups have done worse than chance (see reviews by DePaulo, Stone, & Lassiter, 1985; Kraut, 1980; Zuckerman, DePaulo, & Rosenthal, 1981). Most of these studies examined college students, who may not have had any special reason to learn how to tell when someone is lying. Perhaps professional lie catchers, those whose work requires them to detect lying, would be more accurate.

Surprisingly, three studies of professional lie catchers did not find this to be so. Kraut and Poe (1980) found that customs officials were no more accurate than college students in detecting deceit in mock customs examinations. DePaulo and Pfeifer (1986) found no difference between federal law enforcement officers, regardless of experience, and college students. Kohnken (1987) found police officers did no better than chance when they judged videotapes of college students who had lied or been truthful in an experiment.

It is difficult to draw any conclusions from these three studies of professional lie catchers because in none of them was there any evidence that the observers were exposed to behavior that differed when the people who were judged lied or were truthful. DePaulo and Pfeifer's (1986) study is the only exception, as they used materials that had earlier been shown to be significantly different in observer-rated deceptiveness (DePaulo, Lanier, & Davis, 1983). However, the differences were small, and the data were not analyzed in a way that would indicate how many liars could actually be differentiated on the basis of their observable behavior. Perhaps accuracy has been meager and no advantage found for professional lie catchers because there just was not much information in the videotapes that would allow very good discrimination when people lied.

Ours is the first study to use behavioral samples drawn from a set of videotaped interviews that prior behavioral measurement showed differed when subjects lied or told the truth. Facial muscular movements measured with the Facial Action Coding System (Ekman & Friesen, 1976, 1978) included more masking smiles when the subjects lied and more enjoyment smiles when they told

Paul Ekman's work is supported by Research Scientist Award MH 06092 from the National Institute of Mental Health.

Correspondence concerning this article should be addressed to Paul Ekman, Human Interaction Laboratory, University of California, 401 Parnassus Avenue, San Francisco, CA 94143.

 0003-066X/91/$2.00
Vol. 46, No. 9, 913-920

the truth about their feelings (Ekman, Friesen, & O'Sullivan, 1988). Vocal measurement also distinguished the lying and truthful interviews. There was an increase in fundamental pitch when the subjects lied. When both the vocal measure and the two facial measures were combined, it was possible to classify 86% of the subjects correctly as either lying or being truthful (Ekman, O'Sullivan, Friesen, & Scherer, 1991). Because there were known behavioral differences between the honest and deceptive samples, our study could focus on the question of how well observers can detect deception.

We took advantage of opportunities to test a number of different groups in the criminal justice and intelligence communities to determine whether those who have a specialized interest, and presumably more experience, in detecting deceit would do better than the usual college student observer. We made no hypotheses about the relative proficiency of the professional lie-catcher groups, although we hoped that at least one of them might do better than chance.

In addition to analyzing average accuracy on a group-by-group basis, as is usually done in studies of deception detection, we also planned to examine accuracy on an individual basis. The mean accuracy of a group of observers might be only at chance, but individual observers might reach either very high or very low levels of accuracy. Of course, it is also possible that all observers might perform at or close to chance, but that cannot be known from the usual method of examining only mean accuracy.

Because most of our observers were professional lie catchers, we were interested in their thoughts and opinions about lie catching in general, and in their own lie-catching ability, in particular. DePaulo and Pfeifer (1986) reported that confidence in one's ability to detect lying was unrelated to actual accuracy, although the federal law enforcement officers they studied were more confident than were college students about their ability to detect deception. They also reported that amount of experience in law enforcement was not correlated with accuracy. Kohnken (1987) also found no relationship between confidence in one's ability to detect lying and actual accuracy. Unlike DePaulo and Pfeifer, however, he found a significant negative correlation between experience and accuracy when age was partialed out. We sought to replicate these findings, and so we asked our professional lie catchers about their confidence in their ability to detect deception before and after taking our test, as well as the amount of time they had been in their present job.

We also asked our professional lie catchers to describe the behavioral clues they relied on in making their judgments. We had two reasons for being interested in this matter. First, we wanted to test our hypothesis that those who make accurate judgments, regardless of their occupational group, would describe different behavioral clues than those who make inaccurate judgments. On the basis of our prior analyses of observers' judgments of these videotapes when they were exposed to either the verbal, nonverbal, or combined verbal and nonverbal behaviors (O'Sullivan, Ekman, Friesen, & Scherer, 1991), and our findings that there were clues to deceit in the nonverbal but not in most of the verbal behaviors we measured (Ekman & Friesen, 1976; Ekman et al., 1988; Ekman et al., 1991). Hypothesis 1 predicted that accurate observers would report using nonverbal clues more than would the inaccurate observers.

Our second reason for asking the observers to describe the behavioral clues they relied on in making their judgments was to have data that would be relevant to resolving the question of whether any individual differences in accuracy we obtained were due to chance. If those who were highly accurate gave different reasons for their judgments than those who were inaccurate, it would argue against the possibility that individual differences in accuracy were simply chance variations.

In one of our groups, we were also able to test how well subjects could recognize microexpressions, facial expressions that last no more than 1/25 of a second. On the basis of Ekman and Friesen's (1969) proposal that microexpressions are an important source of behavioral clues to deceit, Hypothesis 2 predicted a positive correlation between accuracy in detecting deceit and accuracy in recognizing microexpressions of emotion.

Method

Observers

Following the publication of his book *Telling Lies,* Ekman (1985) was asked by a variety of groups that had a professional interest in lying to conduct a workshop on behavioral clues to deceit. At the start of each such workshop, the participants were given a test of their ability to detect deception, which provided the data for this study. None of the observers had read Ekman's book prior to being tested. The following groups were tested:

1. U.S. Secret Service. All members of the Forensic Services Division of the Secret Service who were available in Washington, DC, when the workshop on lying was given were tested.

2. Federal polygraphers. All participants in a Federal Inter-Agency Polygraph Seminar organized by the Central Intelligence Agency (CIA) held in the Federal Bureau of Investigation (FBI) Academy at Quantico, Virginia, were tested. This included 10 CIA, 10 FBI, 5 National Security Agency (NSA), 21 Army, Air Force, or Marine personnel, and 14 polygraphers employed by other federal agencies.

3. Judges. All of the participants in three courses on fact finding offered in the midcareer college for municipal and superior court judges organized by the California Center for Judicial Education and Research, and judges taking a similar course in Oregon, were tested.

4. Police. Those attending the annual meeting of the California Robbery Investigators Association, which included city, county, state, and federal law enforcement officers who specialize in dealing with robbery were tested.

5. Psychiatrists. Texas psychiatrists attending an annual professional meeting, as well as psychiatrists at-

tending staff training sessions in Texas and San Francisco, were tested.

6. Special interest group. People who enrolled for a day-long University of California Extension course on deceit were tested. This group included businessmen, lawyers, accountants, police officers, housewives, social workers, psychologists, and nurses.

7. Students. For comparison with prior studies, we also used a sample of undergraduate psychology students at the University of San Francisco.

Table 1 shows that these groups differed in age and in sex.

Detecting Deception Measure

The detecting deception measure consisted of 10 one-minute samples taken from 10 videotaped interviews, preceded by a practice item. The videotapes showed a black-and-white, head-on view of the full face and body of each subject.

The observers were told that they would see 10 college-age women, about one half of whom would be lying to an interviewer as she answered questions about how she felt about a film she was watching. Each subject would describe positive feelings she would claim to be feeling as she watched what she said were nature films. Some subjects were actually watching such films and would honestly be describing their feelings. Other subjects would really be watching a terribly gruesome film that was very upsetting to them, and they would be lying when they claimed to be having positive feelings about a nature film. The observers were told that all of the subjects were highly motivated to succeed and believed that success in their deception was relevant to their chosen career. After seeing each interview, the observers were allowed 30 seconds to record their choice as to whether the subject was honest or deceptive (more information about the deception scenario is provided in Ekman & Friesen, 1974).

The 10 subjects who were shown to the observers were selected from a group of 31 subjects who had participated in a study of deceit. Behavioral measurements (described earlier) on all 31 subjects had found that both facial and vocal measures differentiated the honest from the deceptive interviews. Most of those findings had not been published when the test was given (Ekman et al., 1988, 1991).

Prior to seeing the videotape, all of the observers except the college students and the special interest group responded to the following questions: "How good do you think you are in being able to tell if another person is lying? Check one of the following: very poor, poor, average, good, very good." "What evidence or clues do you use in deciding that another person is lying or telling the truth?" (Three lines were given for the observers' handwritten responses.)

The videotape was then shown. After seeing each of the 10 persons, the observers recorded their judgment by circling either the word *honest* or the word *deceptive*. Following the second and the eighth videotape samples, the observers were also asked to indicate briefly their reasons for deciding that the interview was honest or deceptive. These two items were selected because pilot data indicated that they differed markedly in difficulty level, although the subjects in both items were lying. The handwritten responses were categorized using a coding system developed in previous research (O'Sullivan & Morrison, 1985) for categorizing observers' descriptions of their reasons for believing that subjects were lying or telling the truth. This system classified responses into 20 categories with interrater agreements ranging from 87% to 94%.

After judging all 10 people, the observers were asked the following questions: "How well do you think you did in telling who was lying?—very poorly, poorly, average, well, very well." "If your job required it, could you lie and conceal a strong emotional reaction?—yes, probably, maybe, probably not, no." Additional questions were asked in some of the groups prior to seeing the videotape. These questions, as well as another experimental test that was given to the special interest group, will be described next.

Results

Which Group Is Most Accurate?

The observers had judged whether each of the 10 persons they saw was lying or telling the truth. The observers' accuracy scores could range from 0 to 10 correct. Because

Table 1
Total Sample Size, Sex, Age, and Job Experience in Observer Groups

Observer group	*N*	Women (%)	Age (in years)		Job experience (in years)	
			M	*SD*	*M*	*SD*
Secret Service	34	3	34.79	5.96	9.12	6.69
Federal polygraphers	60	8	39.42	6.76	6.54	6.19
Robbery investigators	126	2	39.21	8.26	14.77	7.15
Judges	110	11	52.64	9.37	11.50	7.77
Psychiatrists	67	3	54.24	10.28	23.63	10.28
Special interest	73	53	43.33	13.44	10.76	9.89
College students	39	64	19.90	1.74	—	—

exactly one half of the 10 persons they judged were lying, the observer would obtain only a chance total accuracy score if an observer were to judge everyone to be lying or to be telling the truth.

A one-way analysis of variance (ANOVA) on the total accuracy scores for the seven groups was computed. There was a significant between-groups effect, $F(6) = 2.07$, $p < .05$. A Duncan procedure showed that the Secret Service differed from each of the other six groups at the .05 level and none of the other groups were significantly different from one another; Table 2 gives these means and standard deviations.

Another way to consider these data are in terms of the numbers of individuals in each group who scored very high or very low. Table 3 shows three levels of accuracy scores for observers in the seven groups. The first column includes scores from 0 to 30%; these observers can be termed *inaccurate*. The middle column for scores from 40% to 60% includes the mean accuracy levels reported in prior research on either college students or specialized occupational groups, and is close to or at chance. The last column (70% to 100%) represents higher accuracy than has been reported before. None of the Secret Service observers performed below chance (i.e., at or below 30%), and 53% of them scored at or above 70% accuracy. Their superior performance is more markedly shown by considering just those who achieved accuracy scores of 80% or more. Nearly one third (29%) of the Secret Service sample reached this very high level of accuracy. The next closest was the psychiatrist group, in which only 12% reached this high level of accuracy in detecting deception. We also computed binomial tests for each group separately, to ascertain whether any of the groups' accuracy was significantly different from chance (defined as 50%). Only the Secret Service group had a better than chance distribution ($p < .03$).

What Variables Are Related to Accuracy in Detection Deception?

Demographic characteristics. Although there were few women in the Secret Service, federal polygrapher, and police groups, sex differences in lie-detection accuracy were examined in the remaining groups. There was no significant correlation between accuracy in detecting deception and sex in the judge, psychiatrist, special interest, or student groups, or across all of these groups combined.

Table 2
Deception Accuracy Means and Standard Deviations in Observer Groups

Observer group	*M*	*SD*
Secret Service	64.12	14.80
Federal polygraphers	55.67	13.32
Robbery investigators	55.79	14.93
Judges	56.73	14.72
Psychiatrists	57.61	14.57
Special interest	55.34	15.82
College students	52.82	17.31

Table 3
Percentage of Observers In Each Group Achieving Different Lie-Detection Accuracy Levels

Observer group	0–30	40–60	70–100
Secret Service	0	47	53
Federal polygraphers	5	73	22
Robbery investigators	8	66	26
Judges	9	57	34
Psychiatrists	5	63	32
Special interest	10	59	31
College students	15	59	26

Note. Each column heading denotes percentage correct.

Correlations were also computed across all groups, between age, years of job experience (this measure was not relevant to the college sample), and accuracy in detecting deception. None of these correlations approached significance. This was not so, however, when these correlations were computed separately in each occupational group. Age was negatively correlated with accuracy for both the Secret Service ($r = -.347$, $p < .03$) and the federal polygraphers ($r = -.343$, $p < .005$). The scatterplots for these correlations suggested that the relationship with age was strongest at the 80% or better accuracy level. In both the Secret Service and federal polygrapher groups, all of the observers who scored 80% or higher in lie-detection accuracy were less than 40 years old.

Years of job experience were also negatively correlated with accuracy in the Secret Service group ($r = -.376$, $p < .02$), but this correlation was not significant for the federal polygraphers ($r = -.102$, $p < .23$). Age and experience were very strongly correlated in the Secret Service group ($r = .88$, $p < .000$), so that when the influence of age was removed from the correlation between accuracy and experience, and the influence of experience was removed from the correlation between accuracy and age, the partial correlations were not significant. Age and experience were less strong correlated among the federal polygraphers ($r = .35$, $p < .005$). When accuracy was correlated with age, removing the influence of experience, the partial correlation was still statistically significant ($r_{(12.3)} = -.330$, $p < .05$). The correlation between experience and accuracy, controlling for age, was nonsignificant.

Confidence in their ability to detect deception. All groups except the special interest and college students were asked twice to estimate their ability to tell when other people are lying. Before seeing the videotape they were asked about their general ability to detect lies. After the videotape, they were asked specifically how they thought they had done on that measure. When computed, disregarding occupational group, neither observers' general predictions about their ability to tell when other people are lying ($r = .03$, $p < .282$) or their more specific

postdiction of how well they had done in detecting deceit in the videotape they had just viewed ($r = .02$, $p < .358$) were significantly correlated with their actual accuracy. When the correlations were computed separately for each group, there were two exceptions. The federal polygraphers' initial ratings of their general ability to tell when someone is lying was correlated with their actual accuracy ($r = .217$, $p = .05$.), and the Secret Service ratings of how well they had done after viewing the videotape was negatively correlated with actual accuracy ($r = -.31$, $p < .035$). This negative correlation can be attributed to five subjects who rated their ability very high, although their accuracy scores were among the lowest. Examination of the group-by-group scatterplots showed that the failure to find a correlation between prediction and performance in the other groups was not due to outliers or other idiosyncrasies in the distribution of the responses.

Table 4 reports the means and standard deviations of the pre- and posttest ratings of confidence in ability to detect deception. A one-way ANOVA showed that the average ratings of the five groups were significantly different before they saw the videotape, $F(4, 351) = 16.66$, $p < .000$. A Duncan post hoc procedure found that the psychiatrists rated their ability to detect lies in general significantly lower than did all of the other groups and the judges rated themselves lower than did the other law-enforcement groups. A one-way ANOVA of the ratings made after viewing the videotape about how well the observers thought they had done in detecting deceit, was also significant, $F(4, 356) = 7.54$, $p < .000$. The Duncan test showed that the psychiatrists and judges were not different from each other, but that the ratings of both groups were significantly lower than those of Secret Service, police, or federal polygraphers.

The special interest group had been asked a similar question (i.e., "I am very good at telling when another person is lying"), using a nine-point rather than a five-point scale. Special interest observers who rated themselves as good lie detectors scored high in detecting deception in the videotape ($r = .322$, $p = .007$).

Willingness to lie. Observers in three of the occupational groups (Secret Service, federal polygraphers, and psychiatrists) had been asked how well they could conceal an emotional reaction if their job required it. Responses to this question were not significantly correlated with accuracy, either within each of the three groups or across all groups. A one-way ANOVA across the three groups was, however, significant, $F(2, 142) = 7.05$, $p < .0012$. Duncan's post hoc procedure showed that the psychiatrists were significantly different ($M = 2.85$) than federal polygraphers ($M = 2.16$) in their belief that they were able to conceal an emotion if their job required it.

Other self-ratings. To determine whether clinical orientation might be related to accuracy, one of the samples in the psychiatric group had been asked whether they worked from a psychodynamic or a behavioral perspective. To determine whether psychiatrists who had courtroom experience might be better in detecting deception, we also asked whether they did forensic work. Neither of these variables was significantly correlated with lie-detection accuracy in the psychiatrist group.

Recognizing microexpressions. A measure of the ability to recognize emotional facial expressions presented at very brief exposures (1/25 s) was given only to the special interest group. It consisted of 30 black-and-white slides of facial expressions of six prototypic emotions (happiness, sadness, fear, anger, surprise, and disgust), preceded by three practice slides. The 30 items were scored to yield a total accuracy score. The correlation between accurately recognizing the emotions displayed in this test and accuracy in judging which subjects were lying was significant ($r = .270$, $p = .02$), supporting Hypothesis 2.

Table 4
Confidence in Lie-Detection Ability Before and After the Deception Detection Measure (DDM)

	Before DDM		After DDM	
Observer group	*M*	*SD*	*M*	*SD*
Secret Service	3.76	0.61	3.26	0.57
Federal polygraphers	3.56	0.57	3.14	0.61
Robbery investigators	3.53	0.58	3.26	0.81
Judges	3.34	0.59	2.77	0.74
Psychiatrists	2.86	0.75	2.86	0.75

Observers' Descriptions of Behavioral Clues to Deceit

The observers in all of the groups, except for special interest and students, gave open-ended descriptions of behavioral clues they used in judging whether someone was lying on three occasions: prior to seeing the videotape, after judging the second person shown on the videotape, and after judging the eighth person shown on the videotape. The handwritten answers varied in the amount of detail given and in the number of verbal and behavioral clues mentioned. To test Hypothesis 1, a simple three-way classification was performed: all responses that referred only to *speech* clues (e.g., "answers too slowly," "evasive," "talks too much," "contradicts herself"), responses that referred only to *nonverbal behaviors* (e.g., "voice strained," "avoids eye contact," "phony smile," "body language"), or responses that mentioned both speech and nonverbal behaviors. Across both items and all observer groups, 37% of observers reported using speech clues alone, 29% reported nonverbal clues alone, and 25% reported using both verbal and nonverbal clues. The coder performing the classification did not know the occupational group nor the accuracy scores.

Chi-square analyses showed that there were no significant differences among the five occupational groups in the types of clues mentioned prior to their viewing the videotape. Disregarding occupational group, those whose accuracy scores were 80% or more were compared with those whose accuracy scores were 30% or less. Again,

there was no difference in the types of deception clues listed prior to seeing the videotape.

The observers had also been asked to describe the behavioral clues they had relied on immediately after making their judgments of 2 of the 10 subjects. The second person they saw had been expected to be difficult to judge accurately, and indeed only 44% of the observers accurately identified her as deceptive. Also as expected, the eighth person they judged was easy to detect; 84% accurately identified her as deceptive. In support of Hypothesis 1, Table 5 shows that for both of these items, more of the accurate observers described using nonverbal or nonverbal plus speech clues to arrive at their correct choice than did the inaccurate observers, who listed speech clues alone as the basis for making their judgment.

Discussion

Why Did We Find Differences in Accuracy When Others Did Not?

Our results directly contradict those reported by Kraut and Poe (1980), DePaulo and Pfeifer (1986), and Kohnken (1987), all of whom found that occupational groups with a special interest in deception did no better than chance or no better than college students did in detecting deceit. Three reasons may explain why we found differences in accuracy among occupational groups and between the occupational groups interested in deceit and college students, whereas they did not. First, they did not examine the Secret Service, and we did. If we had not examined this group, our results would have replicated theirs. Second, we performed a subject-by-subject analysis, which revealed that there were both highly accurate and inaccurate observers. There may have been significant numbers of such observers in their samples, but they did not report examining their data to determine that. Third, we used samples of honest and deceptive behavior that we knew through prior behavioral measurement did differ. Previous investigators did not establish that their samples of honest and deceptive behavior actually differed, or if they did, that they permitted accurate classification of most of the subjects, and so their observers might not have had much of a chance to detect deception.

Table 5
Percentage of Accurate and Inaccurate Observers Describing Each Type of Behavioral Clue to Deceit

Judgments	Speech only	Nonverbal only	Nonverbal plus speech
Subject 2			
Accurate observers	22	54	22
Inaccurate observers	52	28	20
χ^2 (2) = 45.5, $p < .001$			
Subject 8			
Accurate observers	43	27	30
Inaccurate observers	67	24	9
χ^2 (2) = 10.96, $p < .01$			

Did We Actually Measure the Ability To Detect Deceit?

Although our lie-detection measure contained different behaviors in honest and deceptive samples, some critics might argue that we measured the ability to distinguish positive from negative emotion rather than the ability to detect deception. Such an argument would point out that in the deception samples the subjects were experiencing negative affect from two sources: the negative emotions aroused by the gruesome films they were watching, and negative affect aroused by the need to conceal their feelings, including but not limited to the fear of being caught.

Measurements of the behaviors shown in the videotapes suggest that they do *not* differ in affect valence. There were *no* negative emotional facial expressions in the deceptive interviews, except for those masked by a smile. Also, the generally poor performance by most of our observers—as compared with the very high levels of accuracy obtained when observers are shown samples in which positive and negative affect is elicited without any deliberate attempt to deceive (Ekman, Friesen, & Ellsworth, 1972)—suggests that this was not simply a test of positive versus negative affect discrimination.

Granting that we measured the ability to detect deception, questions can be raised about whether the deception shown in the test is relevant to the interests and experience of the observers who were tested. Clearly, none of the observers were familiar with the particular deception scenario they encountered on our detecting deception measure. Nor do they typically have to make judgments based on one-minute samples of unfamiliar people shown on videotape. Nevertheless, the results showed that this measure did discriminate among observers in predicted ways (Hypotheses 1 and 2). For some of our professional lie catchers (Secret Service, psychiatrists, and police), observing people in an interview format in order to evaluate them is not too dissimilar from their everyday work. The professional experience of federal polygraphers and judges, however, occurs in far more ritualized surroundings—either the stylized administration of the polygraph or the mannered choreography of the courtroom.

This question, about the relevance of this or any other type of deception to the interests and experience of any particular group of observers, raises a theoretical issue about whether behavioral clues to deceit are situation specific or generalizable across situations, regardless of the type of deceit that occurs. Ekman's (1985) analysis of why lies fail, suggests that the behavioral clues to deceit that are generated by emotions are likely to generalize across situations. Not every deceit, of course, involves concealing emotions, but even those that do not may still generate emotion-based behavioral clues to deceit if the liar has strong emotional reactions about engaging in the lie, such as being fearful of being caught, guilty about lying, or excited by the challenge. When lies do not involve

strong emotional reactions, more cognitively based clues, which are more specific to the lie being told, may be more useful. This reasoning is consistent with the findings of DePaulo and her colleagues (DePaulo, Kirkendol, Tang, & O'Brien, 1988) that there are more behavioral clues to deceit when the liar is more motivated to succeed in the lie. We are currently doing research that examines the generality versus situation specificity of behavioral clues to deceit.

It is also important to note the special nature of the deception scenario we showed in this study, one that is not directly relevant to *any* of the occupational groups we studied. Frank and Ekman (1991) have begun to test different occupational groups, using two other deception scenarios: lying about the theft of money and lying about one's opinion. Although their study is not complete, preliminary findings replicate some results reported here, such as the negative correlation between accuracy and age and the lack of correlation between confidence and accuracy. Their work directly addresses the question of whether people who are accurate in detecting one type of lie are also accurate when judging a different type of lie.

Why Are Some People More Accurate in Detecting Deceit?

Across our total sample, we found no relationship between accuracy in detecting deception and age, sex, or job experience. Within two groups, the Secret Service and federal polygraphers, age was negatively correlated with accuracy, with the most highly accurate observers (accuracy 80% or greater) all being under 40 years of age. Although Kohnken (1987) found a positive relationship between accuracy and age for his truth detection task, when he partialed experience out of the relationship between age and accuracy it became significantly negative for his overall task. We found a sizable correlation between experience and accuracy only for the Secret Service. Experience and age were very highly correlated in this group. When age was partialed out of the experience–accuracy relationship, it dropped to insignificance. Kohnken's sample was both younger and more homogeneous than any of our professional groups. Within a younger group, experience may contribute to lie-detection accuracy up to a point of diminishing and even negative returns. More experienced professionals typically are less involved in face-to-face interrogation and more involved with administrative duties, which may result in a decline in their skill in detection deception.

Like our colleagues (DePaulo & Pfeifer, 1986; Kohnken, 1987), we found that observers' confidence in their overall lie-detection ability bore little relationship to their measured accuracy. Ratings of overall lie-detection ability were not significantly related to detection accuracy for either the total sample or any of the separate occupational groups. Most observers' ratings of how well they thought they did, even after they viewed the videotape, were unrelated to actual accuracy.

Our findings suggest that accurate lie catchers used different information than did the inaccurate ones. They listed different and more varied behaviors, emphasizing nonverbal more than verbal ones, and also mentioned using both verbal and nonverbal, rather than relying on verbal behavior alone. This is consistent with the findings from Knapp's (1989) study of what clues military interrogators report they rely on, namely, "subtle cues and nonverbals." Interestingly, our accurate and inaccurate observers did *not* describe different behavioral clues when answering general questions about how they make their decisions before they saw the videotape; they differed only when they described the basis of their decision about a specific person they had just seen.

Also, our finding that accuracy in identifying microexpressions was correlated with accuracy suggests that in addition to informational differences, accurate observers may possess superior skills in spotting and decoding emotional information displayed on the face. One way to test this explanation would be to provide information to observers based on behavioral measurements and train them in recognizing microexpressions. Prior attempts to train observers to detect deceit have yielded contradictory results. Zuckerman, Koestner, and Alton (1984) found benefits only in judging the person on whom training was given; Kohnken (1987) reported no benefits of training, but Zuckerman, Koestner, and Colella (1985) reported benefits of training beyond just the person on whom training was given. However, the training in these studies was simply to tell the observers the correct answers, not to provide information about specific behavioral clues nor to train specific perceptual skills.

Why Is the Secret Service Better Than Other Occupational Groups?

There are a number of possible explanations. Many of the members of this group had done protection work, guarding important government officials from potential attack. Such work may force reliance on nonverbal cues (e.g., scanning crowds), and that experience may result in greater attention to nonverbal behavior in our test. Also, there may be a difference in the focus of their interrogations. The members of the Forensic Services Division of the Secret Service whom we tested spend part of their time interrogating people who threaten to harm government officials. Secret Service officials told us that most of these people are telling the truth when they claim that their threat was braggadocio, not serious. It is only the rare individual who is lying in his or her denial and actually intends to carry out such a threat. Members of the criminal justice community told an opposite story; they believe that everyone lies to them. Thus, the Secret Service deals with a much lower base rate of lying and may be more focused on signs of deceit, whereas the criminal justice groups, with a higher base rate of lying, may focus more on obtaining evidence, not detecting lies. There is no way to test such speculations, although experimental studies could try to manipulate these variables.

In discussing our findings with members of the other occupational groups, they suggested other explanations

for their poor performance. The polygraphers claim to focus on the polygraph exam itself, the preparation of the questions to be asked, and the reading of the charts. Many of them specifically disavow attending to nonverbal behaviors, which in our test were the measurably discernible source of clues to deceit. Judges told us that they usually are seated in a position that prevents them from seeing the faces of those who testify, and are often focused on taking notes rather than attending to the nuances of behavior. They tend to pay most attention to the words the witnesses say, rather than to their behavior. Many psychiatrists claim not to be interested in lying, saying that patients will eventually reveal the truth to them. This is not so for those who do forensic work, and thus it was surprising that we found no difference between them and psychiatrists who do not do forensic work.

Conclusion

Our study demonstrated that some lie catchers (viz., the Secret Service) can catch liars, that more accurate lie catchers report using nonverbal as well as verbal clues to deceit, that they are better able to interpret subtle facial expressions, and that in some occupational groups, accurate lie catchers are younger rather than older. Accurate lie catchers cannot be identified by either sex or their confidence in their lie-catching ability. Some caution about these findings must be maintained, however, because they are based on judgments of only one kind of deception, the concealment of strong negative emotions.

We are developing a psychometrically sound test of the ability to detect deceit that includes a number of forms of deception, which could be used to identify those individuals who are very good and very poor at this task. In different aspects of the criminal justice process, in certain business settings, and in some clinical situations, it may be useful to have such information.

REFERENCES

DePaulo, B. M., Kirkendol, S. E., Tang, J., & O'Brien, T. (1988). The motivational impairment effect in the communication of deception: Replications and extensions. *Journal of Nonverbal Behavior, 12,* 177–202.

DePaulo, B. M., Lanier, K., & Davis, T. (1983). Detecting the deceit of the motivated liar. *Journal of Personality and Social Psychology, 45,* 1096–1103.

DePaulo, B. M., & Pfeifer, R. L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology, 16,* 249–267.

DePaulo, B. M., Stone, J. I., & Lassiter, G. D. (1985). Deceiving and detecting deceit. In B. R. Schlenker (Ed.), *The self and social life* (pp. 323–370). New York: McGraw-Hill.

Ekman, P. (1985). *Telling lies: Clues to deceit in the marketplace, marriage, and politics.* New York: Norton. (Paperback ed., Berkeley Books, New York, 1986)

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32,* 88–105.

Ekman, P., & Friesen, W. V. (1974). Detecting deception from body or face. *Journal of Personality and Social Psychology, 29,* 288–298.

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior, 1*(1), 56–75.

Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System: A technique for the measurement of facial movement.* Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the human face: Guidelines for research and an integration of findings.* New York: Pergamon Press.

Ekman, P., Friesen, W. V., & O'Sullivan, M. (1988). Smiles when lying. *Journal of Personality and Social Psychology, 54,* 414–420.

Ekman, P., O'Sullivan, M., Friesen, W. V., & Scherer, K. R. (1991). *Face, voice and body in detecting deceit.* Manuscript submitted for publication.

Frank, M., & Ekman, P. (1991). [Success in detecting deception across situations]. Unpublished raw data.

Kohnken, G. (1987). Training police officers to detect deceptive eyewitness statements: Does it work? *Social Behaviour, 2,* 1–17.

Knapp, B. G. (1989). *Characteristics of successful military intelligence interrogators* (USAICS Tech. Rep. No. 89-01). Fort Huachuca, AR: U.S. Army.

Kraut, R. E. (1980). Humans as lie detectors: Some second thoughts. *Journal of Communication, 30,* 209–216.

Kraut, R. E., & Poe, D. (1980). On the line: The deception judgments of customs inspectors and laymen. *Journal of Personality and Social Psychology, 39,* 784–798.

O'Sullivan, M., Ekman, P., Friesen, W. V., & Scherer, K. R. (1991). *Judging honest and deceptive behavior.* Manuscript submitted for publication.

O'Sullivan, M., & Morrison, S. (1985, April). *How can I tell you're lying? Let me count the ways.* Paper presented at the meeting of the Western Psychological Association, Burlingame, CA.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 1–59). San Diego, CA: Academic Press.

Zuckerman, M., Koestner, R., & Alton, A. O. (1984). Learning to detect deception. *Journal of Personality and Social Psychology, 46,* 519–528.

Zuckerman, M., Koestner, R., & Colella, M. J. (1985). Learning to detect deception from three communication channels. *Journal of Nonverbal Behavior, 9,* 188–194.