Journal of Personality and Social Psychology
1997, Vol. 72, No. 6, 1429–1439

0022-3514/97/$3.00

The Ability to Detect Deceit Generalizes Across Different Types of High-Stake Lies

Mark G. Frank
Rutgers—The State University of New Jersey

Paul Ekman
University of California, San Francisco

The authors investigated whether accuracy in identifying deception from demeanor in high-stake lies is specific to those lies or generalizes to other high-stake lies. In Experiment 1, 48 observers judged whether 2 different groups of men were telling lies about a mock theft (crime scenario) or about their opinion (opinion scenario). The authors found that observers' accuracy in judging deception in the crime scenario was positively correlated with their accuracy in judging deception in the opinion scenario. Experiment 2 replicated the results of Experiment 1, as well as P. Ekman and M. O'Sullivan's (1991) finding of a positive correlation between the ability to detect deceit and the ability to identify micromomentary facial expressions of emotion. These results show that the ability to detect high-stake lies generalizes across high-stake situations and is most likely due to the presence of emotional clues that betray deception in high-stake lies.

Professional lie catchers recognize that some lie catchers are consistently better judges of interpersonal deception than others. For example, Australian customs agents have noted that the same group of officers seems to top their "contraband recovered" rankings each month (S. Van Der Kooy, Australian customs officer, personal communication, June 18, 1993; see also Kraut & Poe, 1980). Moreover, some American police organizations ask their consistently good lie catchers to train other interrogators (J. J. Newberry, Alcohol, Tobacco, and Firearms agent, personal communication, April 1992). Even observations of professional poker players have suggested that the best players are characterized by their ability to recognize deception across a variety of opponents and situations (Hayano, 1980). Although this anecdotal evidence concerns detection abilities within a given occupation, it suggests that a person's ability to detect deceit is not random or limited to specific people or situations but may be a skill that generalizes across different people and different kinds of lies.

These observations, however, run contrary to years of psychological research that has suggested that the ability to detect deceit is not general but rather situation or person specific. For example, within the context of a single deception situation, studies have found no relationship between a lie catcher's ability to judge accurately the truthfulness of one stimulus individual and his or her ability to judge accurately that of a different stimulus individual (Kraut, 1978, 1980). Moreover, research has found no relationship between accuracy of judgment and gender of stimulus individual (DePaulo & Rosenthal, 1979b) or between accuracy of judgment and culture of stimulus individual (i.e., same culture as judge vs. different; Bond, Omar, Mahmoud, & Bonser, 1990). Taken together, these studies demonstrated that the skill levels and characteristics of the liars seemed to override any potential individual differences in lie catchers' skill. These researchers reasoned that if there is no relationship between accuracy of judgment and liar characteristics within a given type of lie, there should be no relationship between accuracy of judgment and liar characteristics across different types of lies (DePaulo, 1994; DePaulo, Zuckerman, & Rosenthal, 1980). However, this same research showed a strong relationship between ability to detect different emotions and gender of stimulus individual (DePaulo & Rosenthal, 1979b).

This failure to find generality in the ability to detect deceit extends to training individuals to detect deceit. Using a paradigm in which stimulus individuals either lied or told the truth about someone they liked and someone they disliked, researchers found that when judges were presented with information about deception prior to the item or were given feedback on their performance after seeing an item, they improved their ability to detect the deceptions of a given deceiver, but they did not improve their ability to detect the deceptions of different deceivers (Zuckerman, Koestner, & Alton, 1984). However, using a paradigm in which stimulus individuals were shown slides of landscapes and burn victims and were asked both to lie and to tell the truth about how they felt, researchers found evidence that training improved deception detection ability across different deceivers (deTurck & Miller, 1990). It seems that the main difference between these experiments is that the second was more likely to induce strong emotions, such as disgust and fear. Signs of these emotions are produced and recognized across

This work was originally supported by National Institute of Mental Health National Research Service Award MH09827 and later by a research grant from the Australian Research Council and Research Scientist Award, MH06092.

We would like to thank David Matsumoto and his emotion research group at San Francisco State University for their assistance in conducting Experiment 1 and Colette D'Abreo Read for FACS coding the stimulus sample.

Correspondence concerning this article should be addressed to Mark G. Frank, School of Communication, Information, and Library Studies, Rutgers—The State University of New Jersey, 4 Huntington Street, New Brunswick, New Jersey 08901-1071, or to Paul Ekman, Department of Psychiatry, University of California, 401 Parnassus Avenue, San Francisco, California 94143. Electronic mail may be sent via the Internet to mgfrank@scils.rutgers.edu or ekmansf@itsa.ucsf.edu.

cultures and situations and thus would provide context- or situation-independent evidence for deception (Ekman, 1989, 1992, 1994). However, this evidence for generality in detecting deceit was limited in deTurck and Miller (1990), because the deceivers in the study were selected on the basis of their self-monitoring scores, and more important, there was no independent evidence for the existence of perceivable emotion on the part of the deceivers.

A different approach to this question of generality provides another piece of evidence that is consistent with the situational specificity conclusion. This approach examined the relationships between stable individual difference measures, such as self-monitoring (Snyder, 1974), and the ability to detect deceit. The rationale guiding these studies was that if the ability to detect deceit was correlated with these temporally and situationally stable individual difference measures, then the ability to detect deceit must also be temporally and situationally stable. The actual results of these studies have been contradictory. Some have reported positive relationships between accuracy at detecting deceit and variables such as social participation, perceived complexity of human nature, social anxiety, and self-monitoring (DePaulo & Rosenthal, 1979a; Geizer, Rarick, & Soldow, 1977), whereas others have reported no relationship between accuracy at detecting deceit and self-monitoring, CPI scores, or other personality variables (e.g., Kraut & Poe, 1980; O'Sullivan, Ekman, Friesen, & Scherer, 1992). Overall, these equivocal results suggest that the ability to detect deceit would not generalize across situations or lies.

The disparity between the anecdotal observations suggesting generality and the psychological findings suggesting specificity might be accounted for by the differences in the structural features of the deception situations used to arrive at their respective conclusions (Frank, 1992). The anecdotal evidence is derived from very high-stake, real-world deception situations—police interrogations, customs inspections, and high-stake poker games—in which liars and truth tellers have much to gain or lose by being judged deceptive. In contrast, the psychological findings are derived from mostly low-stake, real-world deception situations—white lies, day-to-day polite lies (DePaulo et al., 1980)—where the liars and truth tellers either have little to gain or lose by being judged deceptive or have little fear of being caught telling these lies because they are required by polite society. According to Ekman (1985), this distinction between high and low stakes for successful or unsuccessful deception is critical, because the presence of high stakes is central to a liar feeling strong emotions when lying. It is the presence of these emotions, such as guilt, fear of being caught, and disgust, that can betray the liar's deception when they are leaked through nonverbal behaviors such as facial expression (Ekman, Friesen, & O'Sullivan, 1988) or voice tone (Ekman, Friesen, & Scherer, 1976). Given the finding that emotions expressed in the face are universally recognized (Ekman, 1989, 1992, 1994), Ekman (1985) has further argued that the extent to which the stakes elicit emotions that provide clues to deceit in the expressive behavior of the liar, a lie detector who attends to these behavioral clues would not need to know the specifics of the situation being evaluated in order to accurately judge deception. On the basis of this reasoning, he predicted that the ability to detect deceit would generalize across different types of high-stake lies.

To date, no experiment has tested this idea directly by showing the same group of observers two different samples of liars culled from two different high-stake deception situations.[1] However, two studies have provided indirect evidence consistent with Ekman's (1985) generality hypothesis. In the first study, undergraduate judges were shown videotapes of both high- and low-motivated, and hence high- and low-aroused, stimulus participants who lied and told the truth about their opinions, attitudes, and feelings on a variety of topics (DePaulo, Lanier, & Davis, 1983). The experimenters motivated the stimulus participants by telling them either (a) that their success in deceiving was related to career success and that their behavior would be carefully scrutinized (high motivation) or (b) that their deceptions were simply a little game and that their behavior would not be scrutinized (low motivation). The results showed that the high-motivation stimulus participants were more accurately judged from their nonverbal behavior, whereas the low-motivation participants were more accurately judged from their verbal behavior. However, there was no independent measure of expressed emotion in this study, and a manipulation check showed that the high- and low-motivation participants did not differ on their self-reports of tension while responding.

In the second study, professional lie catchers such as agents of the Secret Service, federal polygraphers, judges, and psychiatrists, as well as students, were shown videotapes of highly motivated nursing students who were attempting to convince an interviewer that the films they were watching made them feel pleasant, when one of the films was pleasant and the other one featured gory scenes of an amputation and scenes of burn victims (Ekman & O'Sullivan, 1991). The experimenters motivated the nurses by telling them that their success in convincing an interviewer that the gory films they were watching were actually pleasant would be related to their success in their nursing careers (see Ekman & Friesen, 1974, for more details). In this high-stake, emotionally arousing deception situation, not only could many of the observers detect deceit significantly better than chance but those observers who were most accurate reported using a deception detection strategy that was based on examining the nonverbal behavior of the stimulus participants, whereas those observers who were least accurate reported using a deception detection strategy that was based on examining only the verbal behavior of the stimulus participants. The highly accurate observers' strategy was successful in this deception situation because they judged individuals who—because they faced very high stakes for successful and unsuccessful deception—displayed facial and vocal signs of emotion when lying (Ekman & O'Sullivan, 1991; Ekman, O'Sullivan, Friesen, & Scherer, 1991).

Taken together, these studies show that high-stake deceptions are more likely to produce clues to deception in a person's

[1] Although one could argue that the faking positive and faking negative lies used by DePaulo and Rosenthal (1979b) are different types of lies, they are both derived from the same deception situation: describing people. Our study featured two different situations: a lie about an action one just took and a lie about opinions. However, we acknowledge that this can be a slippery definitional issue.

nonverbal behavior, which, research has shown, is more likely than verbal behavior to contain information about the emotional state of the person (Ekman & Friesen, 1969; Zuckerman, DePaulo, & Rosenthal, 1981). Moreover, the Ekman and O'Sullivan (1991) study showed that professionals were able to judge deceit accurately in a high-stake deception situation that they were not at all familiar with, which is consistent with Ekman's (1985) proposal that accuracy when detecting different kinds of lies is not contingent on knowledge of the details of a situation if the stakes are sufficiently high to arouse emotions.

However, none of the aforementioned studies examined observers' abilities to judge deception across more than one type of lie, let alone more than one type of high-stake lie. Thus, we propose to make a parsimonious test of lie catchers' abilities to judge deception across different types of lies by comparing their abilities to judge accurately a sample of liars in one high-stake deception situation with their ability to judge accurately a different sample of liars in a different high-stake deception situation. We predicted that the ability to detect deceit would generalize across high-stake lies such that a person who scored high in judging one high-stake lie would score high when judging a different type of high-stake lie.

Experiment 1

Overview: Creating Ecologically Valid High-Stake Scenarios

There are a number of structural features of deception situations that must be recreated in the laboratory to generalize laboratory results to the real world (Frank, 1992; Podlesny & Raskin, 1977). For example, to recreate the structural features of a criminal investigation in the laboratory, one must have high stakes—that is, the situation must contain not only a reward for successful deception but also a punishment for unsuccessful deception. The liar who is believed obtains freedom as well as whatever ill-gotten gain he or she may have acquired in the commission of the crime. The liar who is disbelieved faces severe consequences—for example, he or she can go to jail or even be executed. In addition, these stakes must apply not only to the liar but also, in a slightly different way, to the truth teller. A truth teller who is believed obtains his or her freedom, but a truth teller who is disbelieved faces the same consequences as the liar who is disbelieved. Newspaper accounts of death row inmates who are released when the actual killer confesses illustrate that in many real-life situations, not only are disbelieved liars punished, but so are disbelieved truthful people.

For this experiment, we created two high-stake deception situations that we felt would be very different from each other yet still contain many of the structural features described above. The first deception situation involved the mock theft of money, and the second involved falsifying one's opinion about a current event issue. We adapted the first scenario from a mock crime paradigm used in polygraph research (e.g., Horowitz, 1989). In our version of the mock crime paradigm (what we refer to as the *crime scenario*), participants were told that they and a second participant—actually a confederate—would have the opportunity to take $50 cash from inside a briefcase. The participant and the confederate would enter the room containing the briefcase one at a time. The person who entered first could choose whether to take the money; the person who entered second would have to take the money if it was there and, of course, could not take it if it was not there.[2] We used this procedure for two reasons. First, we wanted to mirror what usually happens in an actual crime—that is, in the real world, a crime already has been committed before an investigation is begun. We wanted the participant to know that this "crime" had been committed and that it was either that participant or the other individual who had taken the money. Second, our pilot participants reported that they were very suspicious of the whole procedure because each knew that he was the only plausible suspect if the money was missing. The participants never saw the confederate, but they heard his voice over an intercom and heard him shuffling about the room with the briefcase when it was his turn to enter. No participant expressed suspicion about the true identity of the confederate. The size of the reward was based on pilot studies that suggested this amount of money was sufficient to induce about half the participants to choose to lie and sufficient to motivate the liar to want to succeed. These rewards represent those situations in life in which there is more to be gained—albeit in the short run—from lying than from being truthful.

All participants knew that after this crime had been committed, they would be interrogated concerning the theft. All participants were instructed to deny that they had taken the money. Participants (all of whom were men) were told that if they took the money and were able to convince the interrogator that they had not taken it, they could keep the $50. If a participant took the money and the interrogator judged him as lying, that participant was told that he would be punished by being forced to forfeit not only the bonus money but also his $10 per hour participation fee. Moreover, half of the participants were randomly assigned to face an additional punishment if judged lying: They were told they would have to sit on a cold, metal chair inside a cramped, darkened room labeled ominously *XXX*, where they would have to endure anywhere from 10 to 40 randomly sequenced, 110-decibel startling blasts of white noise over the course of 1 hr.[3] These participants were given a sample of this

[2] What this means is that half of the participants were allowed to choose whether to take the money and half were assigned to take it or not take it. This manipulation did not affect any of the analyses reported in this article; for example, observers were no more accurate at judging the participants who chose whether to take the money. However, one reviewer raised the point that any relationship for accuracy across scenarios might be due to observers' being able to judge which participants would choose to lie rather than due to observers' being able to judge deception from behavioral clues. We were able to rule out this explanation by finding a nonsignificant correlation between observers' accuracy (in Experiments 1 and 2 combined) for judging those who could choose in the crime scenario and for judging those who could choose in the opinion scenario ($r = .14$, $p > .10$, $n = 78$). Ironically, that was the only nonsignificant correlation for accuracy across choice and deception scenario. Therefore, the choice manipulation is not discussed further.

[3] We did not have to threaten all participants with this punishment to make the situation high stake; the gain or loss of $50 and the participation fee seemed to be a high enough stake to produce the facial signs of negative emotion. We take the advice of Levenson (1988), who argued that the criterion for study of emotional phenomena should be the independently confirmed presence of the emotion and not the attempts to elicit that emotion. There was no difference in the presence of emotion,

punishment prior to engaging in the task, but no participant actually received the punishment afterward. Most prior investigators (including us) have not punished the liar who is detected by the interrogator (Mehrabian, 1971, is an exception; he used mild electric shock as a punishment to the liar who was caught). We felt these punishments were appropriate because in most real-life situations, not only is a reward reduced or eliminated when someone is caught lying, but a punishment is administered to the unsuccessful liar. It is this punishment that is central to a liar's fear of being caught (Ekman, 1985).

We told each participant that if he told the truth and the interrogator judged him to be truthful, he would receive a $10 bonus on top of his participation fee. If a participant told the truth but the interrogator judged him as lying, he would not get the $10 bonus or the participation fee and would face the same punishment as the guilty participant who had been judged as lying. The reason we did this is that often in high-stake situations, both truthful and deceitful people evidence fear: In the truthful person it is the fear of being punished if mistakenly judged to be lying (called the "Othello error" by Ekman, 1985), and in the deceitful person it is the fear of being punished if judged correctly.

We adapted the second deception scenario (we refer to this as the *opinion scenario*) from the false opinion paradigm (e.g., Mehrabian, 1971). In our version, participants were given a questionnaire that assessed both the direction and magnitude of their opinions on social issues (e.g., "Should convicted cold-blooded murderers be executed?" and "Should smoking be banned from all enclosed public places?") on a 1 (*strongly agree*) to 7 (*definitely disagree*) rating scale. The opinion that the participant felt most strongly about—and hence was most certain about—was selected as the one to be described truthfully or falsely. A false description meant trying to convince the interrogator that the participant held an opinion on that issue exactly opposite to the participant's true opinion. The stakes for this scenario, participants were told, were identical to those of the crime scenario: Truth tellers who were believed received a $10 bonus, liars who were believed received a $50 bonus, and liars or truth tellers disbelieved lost all money and some faced the additional punishment. We administered the opinion questionnaire to participants before they received the full instructions for either of these paradigms to ensure that they would not lie about their opinions when they completed this questionnaire. All participants were instructed to claim that they were honestly describing their opinions.

Method

Participants. Participants were 20 men, age 18–28, from the San Francisco area who volunteered for a study entitled "Communication Skills Experiment." They were told that they would be paid $10 per hour for participation, with the possibility of gaining more money depending on performance.

The observers who judged the participants were 32 female and 17 male undergraduates from San Francisco State University who volunteered for a study entitled "Can You Spot a Liar?" They received course credit for their participation.

Deception procedure. All participants engaged in the crime scenario first. The participant was instructed to enter a room in the laboratory where the money-filled briefcase was; half of the time the confederate went first, and half of the time the participant went first. The participant was instructed that the confederate would be interrogated before him in the crime scenario and after him in the opinion scenario. Using this procedure allowed us to inform the participants that the interrogator knew the money was taken but did not know whether it was the participant or the confederate who took it.

After having the opportunity to take the money or not, the participant returned to the original room. From there he was led by the experimenter to the punishment room to await the arrival of the interrogator. The participant stayed in the punishment room while the interrogator ostensibly interrogated the confederate. After 6–7 min, the experimenter led the participant to the interrogation room. The interrogation room was approximately 3 × 4 m and featured two chairs—one comfortable, padded chair and one metal chair with no armrests, which was anchored to the floor. Next to the metal chair was a stand-alone lamp, and at a 120° angle was the second, more comfortable chair, where the interrogator sat. The experimenter led the participant to the metal chair and introduced him to the interrogator. The interrogator was described as an expert in judging deception. The interrogator did not know whether the participant had taken the money or whether the participant was the first or second to enter the room. The participant was then interrogated for approximately 5–7 min on a series of 20 scripted questions. After that, the interrogator told the participant that he would examine his notes and return in a few minutes for a second interrogation. The interrogator then left the room.

At this point, the experimenter reentered and gave the participant the instructions for the opinion scenario. These instructions stated that the participant was to choose whether to lie about his opinion on the social issue that was written on top of the instruction form. (Earlier, the experimenter had examined the participant's responses to the opinion questionnaire and selected the topic that the participant indicated that he felt most strongly about.) The participant was told that he would have only 2 min to prepare himself for the next interrogation. The instructions also reminded the participant of the stakes for successful and unsuccessful deception. After 2 min, the interrogator returned and asked the participant a series of 20 scripted questions parallel to those asked in the crime interrogation. The interrogator reminded the participant that he would be interrogating the "other participant" (the confederate) on his opinion after the participant had finished. When the opinion interrogation was complete, the participant was asked to fill out another questionnaire, which included questions that probed his suspicions about the experiment in general, how successful he felt he was at convincing the interrogator, and how much he feared the consequences of being judged deceptive. When the participant finished, he was told that the experiment was over and that he had been judged either innocent or guilty on the crime interrogation and judged either honest or deceptive on the opinion interrogation. Participants who were judged to be honest were paid whatever bonus was applicable, and participants who were judged to be deceptive were told that they would not receive any bonus but would not have to face any punishment. All participants, regardless of judgment, received their $10 per hour participation fee and then were debriefed.

Stakes check. We designed both deception scenarios to have high stakes, and as such they should have been able to elicit strong emotions on the part of the participants. We assessed the presence of emotion by applying the Facial Action Coding System (FACS; Ekman & Friesen, 1978) to score the participants' facial expressions during the interrogations. FACS is a comprehensive system that measures all visible facial

as assessed by the facial measurement, between those who were threatened with the noise punishment and those who were not. There was also no difference in detectability between these participants, so this difference is not discussed further.

muscle movements and not just those presumed to be related to emotion. FACS was chosen because it is nonobtrusive and has been used to verify the physiological presence of emotion in a number of studies (e.g., Ekman, Davidson, & Friesen, 1990; Ekman et al., 1988; Levenson, Ekman, & Friesen, 1990). In this high-stake situation, we expected that the deceptive participants would feel strong negative emotions, such as fear, disgust, and anger (e.g., fear of getting caught, Ekman, 1985; disgust at oneself for lying; and anger at having one's integrity challenged). We used a computerized dictionary that offered a priori predictions about which configuration of FACS scores represented various facial expressions of emotion; these predictions were based on particular combinations of facial movements that have been judged by observers across cultures to represent various emotions (see Ekman & Friesen, 1975, for more details about the facial configurations). We applied the FACS coding to the participants after we had edited the stimulus materials (see below).

Creation of stimulus materials. Each participant's mock crime and opinion interrogations were videotaped from behind a hidden panel. The participant knew beforehand that this interrogation was going to be videotaped, but this hidden camera technique kept the camera from distracting the participant. Overall, 12 out of 20 participants lied in the crime scenario (which meant that 7 out of the 10 participants who entered the room first chose to take the money), and 9 out of 20 lied in the opinion scenario. Each participant's responses to the first six questions of each interrogation were edited and assembled onto two separate videotapes;[4] one videotape consisted of the crime interrogations for 10 participants, and the second videotape consisted of the opinion interrogations for the remaining 10 participants. We assigned participants to each videotape based on three criteria. First, no participant could appear on more than one tape; otherwise, observers' deception accuracy across different types of lies would be confounded by their accuracy for different stimulus participants. Second, each videotape had to contain 5 men lying and 5 men telling the truth so that observers who simply guessed would get only 50% correct. Third, the number of opinions discussed on each videotape could be no more than three to reduce confusion, and of these three, approximately even numbers of participants had to be lying and telling the truth about both the affirmative and negative perspectives on each opinion (i.e., equal numbers lying and telling the truth about supporting the death penalty, and equal numbers lying and telling the truth about opposing the death penalty). There was only one split of the 20 participants that satisfied all three criteria. Each participant's videotape excerpts were approximately 1 min in length and showed him in facial close-up with full audio. The interrogator also could be heard—but not seen—asking his questions. The duration of the crime videotape was approximately 12 min, and the duration of the opinion videotape was approximately 14 min.

Judgment procedure. We showed the two deception detection videotapes to 49 observers, in counterbalanced order, in classroom settings ranging from 7 to 11 observers per group. For the crime video, we told observers that they would see 10 men who were being interrogated about their involvement in a theft of money. For the opinion video, we told observers that they would see 10 men who were being interrogated about their opinion on a current event topic. We then asked observers to judge whether each person they saw was lying or telling the truth by circling the word *truthful* or the word *lying* on a response form after viewing each participant's segment. We encouraged observers to record their judgments during the 10 s of silence that occurred between each segment. We told the observers that anywhere from one fourth to three fourths of the people they would see in each video were lying. We did this because our experience has shown that there are always a few observers in any group who circle *truthful* or *lying* every time without actually closely observing the videotape because they think the test is a trick in which all the people are actually truthful or all are deceptive.

We also asked observers to assess on the response form their confidence in their ability to detect deception both before and after taking the test. The pretest question was "How good do you think you are in being able to tell if another person is lying?" The posttest question was "In this video, how well do you think you did in telling who was lying?" Both questions were rated on a 1–5 scale, with 1 indicating *very poor* (or *very poorly*) and 5 indicating *very good* (or *very well*).

Design. The independent variables were type of lie (crime or opinion, a within-subject comparison), gender of observer, and order of videotape presentation (between-subjects comparisons). The dependent variable was the observer's accuracy score for each videotape. Our hypothesis was that observers who excelled at detecting deceit when judging the crime video would also excel at detecting deceit when judging the opinion videotape. We made no predictions about the overall accuracy level of the judgments. Given the past null findings on the relationship between observer confidence and accuracy (e.g., DePaulo & Pfeifer, 1986; Ekman & O'Sullivan, 1991), we predicted no relationship between accuracy and self-rated ability or confidence in detecting deceit.

Results

Stakes confirmation. A key assumption made in our paradigms was that these high-stake scenarios would be successful in producing emotion on the part of the liars. A FACS-trained scorer scored the 10 items for each tape, and a second scorer checked the reliability for approximately 20% of the sample (the scorers agreed on over 76% of the scoring). We then ran these FACS scores through a dictionary created by Ekman and Friesen to get a priori predictions of emotion from facial expression. We found that 80% of participants in the crime video and 80% of the participants in the opinion video could be successfully classified as liars or truth tellers on the basis of the presence or absence of fear or disgust (and inconsistent with meta-analytical findings such as those of Zuckerman et al., 1981, we did not find that smiling differentiated truth from deception; see Frank & Ekman, 1997, for more details). When we considered both videos, the presence of fear or disgust accurately categorized 90% of the liars, and the absence of these emotions accurately categorized 70% of the truthtellers. The percentage of accurate categorizations, which we derived solely from facial expression, compares quite favorably with the 86% accuracy rate demonstrated by Ekman et al. (1990), who used both facial expression and voice pitch to classify truth tellers and liars. Our results clearly show that the stimulus participants were feeling strong emotions when lying during both scenarios.

Preliminary analyses. One observer did not follow instructions and was removed from the analysis. We derived each observer's accuracy score for each video by counting the number of correct judgments he or she made out of the 10. For the sake

[4] The six questions for the crime interrogation were as follows: (a) Describe exactly what happened, what you saw and did, when you were in that room. (b) Describe for me what your thoughts were when you entered that room. (c) Do you know how much money was—or was supposed to be—in the envelope? (d) Did you take the money from the envelope? (e) Did you bring the money with you into this room? and (f) Are you lying to me now? The six questions for the opinion interrogation were as follows: (a) What is your position on this current event issue? (b) Why is it that you believe what you do on this issue? (c) How long have you had this opinion? (d) Is this your true opinion? (e) You didn't just make up this opinion a few minutes ago? and (f) Are you lying to me now?

of clarity, in the text and tables these numbers are converted to the percentage of correct scores, but all analyses were performed on the raw numbers. We found no main effects or interactions for order of presentation of videotapes on accuracy, all $Fs(1, 42) < 1.52$, *ns*. Likewise, we found no main or interaction effects for gender of observer on accuracy, all $Fs(1, 42) < 0.96$, *ns*, so we collapsed across these variables.

Main analysis. An observer's accuracy score could range from 0% to 100% correct in increments of 10%. Because each video featured 5 men lying and 5 men telling the truth, observers who simply guessed would average five items correct (50% accuracy). Thus, we categorized those observers who scored 60% or better on a given video as high scorers for that video and those who scored 50% or less as low-scorers.[5] We chose these cutoffs for two reasons. First, the experimental literature reported that observers rarely surpass 60% accuracy when detecting deceit (DePaulo et al., 1980). Second, a series of one-sample t tests that set μ at five items correct showed that six items correct in either scenario was significantly different from five items correct at the two-tailed $p < .05$ level.

The number of high and low scores in each scenario is presented in Table 1. This table shows that there is a strong association between being a high scorer on the crime video and being a high scorer on the opinion video, $\chi^2(1, N = 48) = 6.15$, $p < .02$. Seventy-eight percent of the high scorers on the crime video were also high scorers on the opinion video, and 70% of the high scorers on the opinion video were high scorers on the crime video. Likewise, 67% of the low scorers on the opinion video were low scorers on the crime video and 57% vice versa. Thus, it appears performance when judging one high-stake deception is related to performance when judging a different high-stake deception.

A Pearson correlation revealed that this relationship was linear; a comparison of all 48 observers' accuracy rates for detecting deception in the crime videotape with their performance in the opinion videotape showed a significant positive correlation ($r = .48$; $p < .001$, one-tailed). This means that those individuals who excelled at detecting one type of lie tended to excel in detecting a different type of lie, those who performed at chance levels for one were at chance for another, and those who did poorly for one type of lie did poorly for the other.

Subsidiary analyses. There were no differences in overall accuracy between observers' performance in the crime video ($M = 58\%$) and their performance in the opinion video ($M = 59\%$), $F(1, 42) < 1$. These levels of accuracy are at the high end of the typical range of detection accuracy reported in other studies of deception detection (see reviews by DePaulo, Stone, & Lassiter, 1985; Zuckerman & Driver, 1985), and both the crime video accuracy, $t(47) = 3.78$, $p < .01$, and the opinion video accuracy, $t(47) = 3.61$, $p < .01$, were greater than chance (50%). The accuracy scores for observers ranged from 10% to 90% in a normal distribution for the opinion video, and 10% to 80% in a normal distribution for the crime video.

Table 1
Number of Observers Who Scored High and Low on Judging Deception Across Deception Scenarios in Experiment 1

	Opinion scenario score		
Crime scenario score	High	Low	Total
High	21	6	27
Low	9	12	21
Total	30	18	48

Note. A high scorer judged deception at 60% accuracy and above. A low scorer judged deception at 50% accuracy and below. Contingency table $\chi^2(1, N = 48) = 6.15$, $p < .02$.

We also found no relationship between observers' pretest confidence in their abilities to detect lies and their actual accuracy at detecting lies for either video (for crime, $r = .10$; for opinion, $r = -.15$; both *ns*), nor did we find any relationship between observers' posttest confidence in their performance after each video and accuracy (for crime, $r = .13$; for opinion, $r = .13$; both *ns*). Thus, those who were more accurate at detecting deceit did not necessarily think they were any better than those who were less accurate, either before engaging in the task or after completing the task. However, we did find that pretest confidence and posttest confidence were significantly correlated for the crime video such that those who were confident prior to the crime video tended to be more confident after that video, and those less confident before the video remained less confident afterward ($r = .53$, $p < .001$). There was no such pre–post relationship for judging the opinion video ($r = .15$, *ns*). However, those who were more confident prior to judging the crime video were also more confident prior to judging the opinion video ($r = .30$, $p < .05$). Thus, although this finding is not related to their accuracy, observers seem to have a fairly reliable view of their abilities to detect deception.

We also found a significant decrease in observers' confidence over the course of the videos; observers' confidence scores averaged 3.11 before the tests and 2.91 afterward, $F(1, 43) = 4.13$, $p < .05$. However, we found that observers were no more confident on one test than on the other (for crime, $M = 2.99$; for opinion, $M = 3.02$), $F(1, 43) < 1$. We also found that the observers' confidence for each test did not interact with their ratings made before or after the test, $F(1, 43) = 3.01$, *ns*.

Discussion

A person's ability to detect lies in one high-stake deception situation was correlated with his or her ability to do so in a different high-stake deception situation. This is the first evidence that the ability to detect deceit generalizes across high-stake lies. We also found that detection accuracy and confidence in one's ability were not correlated; these null results replicate other findings that confidence and detection accuracy are not related (e.g., DePaulo & Pfeifer, 1986; Ekman & O'Sullivan, 1991). However, confidence before and after the video was correlated, even though it dropped significantly during the video.

Demonstrating generality across these two types of lies is especially noteworthy given that the observers made their judgments on two scenarios that differed on a number of important

[5] We did not divide observers into groups of above chance, at chance, and below chance because it would have created unanalyzably small cell sizes.

dimensions that would work against the generality hypothesis. First, they differed on the content of the lie. In the crime scenario, participants could tell the truth about everything except whether they had taken the money; in the opinion scenario, participants had to concoct a coherent false opinion and reasons for it—a considerably more cognitively complex task. Second, these scenarios differed in the amount of time available to fabricate the lie. In the crime scenario, deceptive participants had at least 8 min to create their alibis; in the opinion scenario, deceptive participants had only about 2 min to formulate a position opposite to a position that they felt extremely strongly about. Third, these scenarios differed in their order such that all participants went through the crime scenario first and the opinion second. Thus, factors that might affect the participant's performance, such as fatigue, familiarity, and other effects due to order, would be apparent only in the opinion scenario. Finally, each videotape featured an entirely different sample of men with very different abilities to lie (consistent with Ekman et al., 1991, and Kraut, 1980); thus, the finding of cross-situational generality is striking when one considers the additional variance produced by these different abilities to deceive.

There are a number of similarities across these scenarios that may have worked in favor of the hypothesis. For example, the same interrogator followed the same line of questioning in both interrogations, the participants faced the same stakes in both scenarios, and the participants were all men—who, research has shown, differ from women in their deception clues (DePaulo et al., 1985). However, these similarities might also work against the generality hypothesis because the similarities would enable the participant to practice his demeanor or to change a strategy that did not seem successful to the participant in the first interrogation. We cannot precisely measure what effect having one scenario follow another had on each participant. If we presume what is most likely—that it had different effects on different participants—this would work against the generality hypothesis by introducing more variance to participants' behavior.

A final methodological issue involves the small number of stimulus items used in both tests (10 per video). Although some may argue that this small size does not allow an adequate test of the generality hypothesis, this small sample size also works against the finding of stable individual differences because of the problem of restricted range of accuracy scores (0–10).

It appears as if the reason we were able to find evidence for generality was that we were able to create two realistic high-stake deception scenarios in which we could document the expression of negative emotion on the part of the liars. Across both videos, we were able to classify accurately 80% of the participants solely on the basis of facial expressions of emotion. Although the previous literature has used realistic paradigms, it appears as if researchers used realistic low-stake paradigms that in all likelihood did not generate the high levels of emotion that occurred in our paradigm. The presumed lack of strong emotions in those paradigms meant that it is most likely the verbal information—that is, the words participants use to describe their situations, beliefs, or actions—that betrays deception. This verbal information, which is most easily under the volitional control of participants (Ekman & Friesen, 1969; Zuckerman et al., 1981), must necessarily be tied to the features of the situation. Thus, it makes sense that these lower stake paradigms would find evidence consistent with the notion that the ability to detect deceit is situation specific or at least overwhelmed by the individual differences in the liars' abilities to deceive (DePaulo et al., 1980; Kraut, 1980). Yet we must note that those who studied low-stake lies have not to date done an experiment comparable to this one; they have not shown observers two different kinds of lie situations to determine whether accuracy is general or situation-specific in low-stake lies.

Experiment 2

In this high-stake paradigm, facial expressions of emotion betrayed deception, as Ekman (1985) predicted. This suggests that observers who are proficient at reading facial expressions of emotion should be better detectors of deceit. Ekman and O'Sullivan (1991) found in fact that observers who more accurately recognized facial expressions of emotion presented in a way to resemble micromomentary facial expressions (Ekman & Friesen, 1969; Haggard & Isaacs, 1966) were also more accurate in judging deception. However, their study featured participants who were watching films designed to elicit emotions, and so their paradigm may have been strongly weighted toward the importance of recognizing emotions. Our paradigm is a more prototypical high-stake deception situation, in which people are having emotions caused by lying, as compared with lying about emotions (Ekman & Frank, 1993). Thus, we attempted to replicate Ekman and O'Sullivan's (1991) finding by showing photographs of facial expressions of emotion at $^1/_{25}$ s and then determining whether accuracy in this task was correlated with accuracy in judging the two deception scenarios used in Experiment 1. Experiment 2 also provided an opportunity to replicate the generality findings we obtained in Experiment 1.

Method

Participants. The observers were 13 male and 17 female undergraduates from San Jose State University who received course credit for participating in a study entitled "Can You Spot a Liar?"

Materials. We used the same deception detection materials in this experiment as in Experiment 1. We assessed the ability to accurately judge microexpressions of emotion with a 40-item microexpression test videotape. This test consists of 40 slides of facial expressions of emotion—specifically, anger, contempt, disgust, fear, happiness, sadness, and surprise (taken from *Japanese and Caucasian Facial Expressions of Emotion,* Matsumoto & Ekman, 1988)—that were flashed tachistoscopically onto a screen for $^1/_{25}$ s. These tachistoscopic presentations were videotaped, and the videotape served as the microexpression test. Observers were instructed to circle the correct emotion term from a list of seven for each of the 40 items presented.

Procedure. The observers saw the same two deception videotapes, in counterbalanced order, as the observers did in Experiment 1. They were run in classroom settings in groups of 5, 5, 6, and 14. We also asked observers to rate their pre- and posttest confidence on the same measures as used in Experiment 1. Observers viewed and judged which of the 10 men in the crime video were lying and which were truthful and then did the same for the opinion video (or vice versa). After judging both deception videotapes, all observers judged the microexpression test.

We again predicted that an observer's level of accuracy in distinguishing the liars from the truth tellers in the crime video would be related to his or her performance in the opinion video. We also predicted a positive correlation between observers' performance on the deception videos and their performance on the microexpression test. Finally, we again predicted no relationship between confidence and accuracy.

Results

As in Experiment 1, we found no effect for gender of observer or for order of videotape presentation, all $Fs(1, 28) < 1$, *ns*, so we collapsed across these variables. We divided observers into high (60% or higher) and low (50% or lower) scorers using the rationale outlined in Experiment 1. A chi-square tabulating the high scorers in one or both videos and the low scorers in one or both videos shows the same significant pattern as reported in Experiment 1; that is, those who were high scorers in one video scenario were more likely to be high scorers in the other, $\chi^2(1, N = 30) = 5.46$, $p < .025$, one-tailed.

Table 2 shows that 79% of the observers who were classified as high scorers in the crime video were also classified as high scorers in the opinion video (and vice versa). Likewise, 64% of the observers who were classified as low scorers in one of the videos were classified as low scorers in the second. As in Experiment 1, we found a significant positive correlation between an observer's performance in detecting lies in the crime video and his or her performance in the opinion video ($r = .31$; $p < .05$, one-tailed).

Means and subsidiary analyses. As in Experiment 1, we found no overall mean difference in accuracy for detecting lies in the opinion video ($M = 58\%$) versus the crime video ($M = 61\%$), $F(1, 28) = 1.38$, *ns*, although accuracy scores were again greater than chance for the opinion video, $t(29) = 3.53$, $p < .01$, and the crime video, $t(29) = 4.04$, $p < .01$. Except for a significant relationship between pretest confidence and accuracy for the opinion video ($r = .39$, $p < .05$), there was generally no relationship between pre- and posttest confidence and accuracy at detecting lies in the crime video (for pretest, $r = .20$, *ns;* for posttest, $r = .05$, *ns*) and the opinion video (for posttest, $r = .03$, *ns*). Again there was a significant relationship between pre- and posttest confidence for the crime video ($r = .41$, $p < .01$) and this time also for the opinion video ($r = .31$, $p < .05$). We also found a significant relationship between confidence prior to the crime video and confidence prior to judging the opinion video ($r = .59$, $p < .001$). Unlike in Experiment 1, this time we found no change in pre- and posttest confidence and no differences in confidence between the crime and opinion videos, all $Fs(1, 29) < 3.78$, *ns*. These results for confidence replicate the pattern found in Experiment 1; that is, observers seem to have fairly reliable beliefs about their abilities to detect deception, independent of their actual ability.

Microexpression test. We scored observers' responses on the microexpression test as 1 for each correct response and 0 for each incorrect response so that an observer's accuracy score could range from 0 to 40. As predicted, the observers' scores on the microexpression test correlated significantly with their overall (combined crime and opinion video) accuracy ($r = .35$, $p < .04$, one-tailed). When observers' scores on the crime and opinion videos were correlated separately with the microexpression test, the crime video accuracy correlated significantly with the microexpression test ($r = .34$, $p < .04$, one-tailed), but the opinion video accuracy, although also a positive relationship, did not correlate significantly ($r = .20$, $p = .15$). Although this suggests that the ability to accurately judge emotion was more important in detecting deceit for the crime scenario than for the opinion scenario, an *r*-to-*z* transformation test comparing the correlation between microexpression test accuracy and crime video accuracy ($r = .34$) and between microexpression test accuracy and opinion video accuracy ($r = .20$) found that these correlations did not differ from each other ($z = .55$, *ns*). Finally, the microexpression test itself did appear to be a reliable measure of ability to detect microexpression of emotion (Cronbach's $\alpha = .82$; split-half reliability = .84).

Table 2
Number of Observers Who Scored High and Low in Judging Deception Across Deception Scenarios in Experiment 2

	Opinion scenario score		
Crime scenario score	High	Low	Total
High	15	4	19
Low	4	7	11
Total	19	11	30

Note. A high scorer judged deception at 60% accuracy and above. A low scorer judged deception at 50% accuracy and below. Contingency table $\chi^2(1, N = 30) = 5.46$, $p < .025$.

Discussion

These results replicate the finding from Experiment 1 that a person's ability to detect lies does not appear to be situationally specific but is a stable skill that generalizes across different lies told in different high-stake deception scenarios. An observer who was able to accurately detect lies in a crime scenario in which the liar denied an allegation was also able to accurately detect lies in an opinion scenario in which the liar attempted to create a coherent and defensible opinion to which he was vehemently opposed. Of course the converse is also true; those observers who were poor at detecting deceit in one scenario were poor in the other.

We also replicated Ekman and O'Sullivan's (1991) finding that the ability to accurately detect lies is related to the ability to accurately recognize micromomentary facial expressions of emotion, thus extending their findings to deception scenarios besides those that directly involve concealing negative emotions. These findings are consistent with Ekman's (1985) reasoning that high-stake situations arouse emotions that can often betray deception and that the ability to recognize those emotions will aid the lie detector.

Finally, as in Experiment 1, we found a general pattern suggesting no relationship between an observer's accuracy and his or her rated confidence in his or her ability to detect lies—either before or after completion of the lie detection task—even though observers' pre- and posttest confidence was relatively unchanged by the task. Although we did find one significant relationship between pretest confidence and accuracy for the opinion video, this correlation may have been due to chance given that we ran eight different confidence–accuracy combinations and obtained only one significant correlation. Moreover, a number of other studies also have not found this relationship (e.g., DePaulo & Pfeifer, 1986; Ekman & O'Sullivan, 1991).

General Discussion

Taken together, these two studies provide the first experimental evidence that a person's ability to detect deceit may not be situationally specific but instead may be a skill that generalizes across different high-stake lies. It also appears that related to this ability is the ability to accurately distinguish among micromomentary facial expressions of emotion.

At first glance, this generality finding seems to contradict years of psychological literature, which has concluded that the ability to detect deceit is specific to situations (e.g., DePaulo et al., 1980; Kraut, 1980). This contradiction is resolved when one notes that our study was able to show independent evidence for the existence of strong negative emotions in the facial expressions of 90% of the liars and the absence of these negative emotions in the facial expressions of 70% of the truth tellers. Thus, if most of the liars across two situations are showing signs of negative emotion, and most of the truth tellers across the two situations are not, then clearly lie catchers who observe the presence or absence of these emotions are going to be consistently better detectors of deceit across both situations than are lie catchers who do not attend to these clues. The fact that observers who were best able to distinguish among different emotions shown at tachistoscopic speed were also the best detectors of deceit in this experiment, where the liars showed facial signs of negative emotions, strongly supports this explanation. Thus, we were able to find evidence for generality in the ability to detect deceit while others have not because our paradigm successfully generated signs of emotion, whereas all but one of the other studies have reported no evidence for such visible clues. The only other study that documented visible and auditory clues that betrayed deception reported findings consistent with ours, that is, that the best detectors of deceit attended more to nonverbal rather than verbal clues to deceit (Ekman & O'Sullivan, 1991). Finally, consistent with this finding is research showing that liars who were motivated, and hence more aroused, were more accurately judged from their nonverbal behavior and that liars who were not motivated were more accurately judged from their verbal behavior (DePaulo et al., 1983).

Our study was the first to find evidence that the ability to detect the emotions involved in deceit is related to one's ability to detect deception across different high-stake situations. This is consistent with earlier research, which has shown that within a single high-stake deception situation, lie catchers who were better able to detect leaked negative emotion were better able to detect leaked positive emotion; moreover, those who were better at detecting leakages by men were better able to detect leakages by women (DePaulo & Rosenthal, 1979b).

The consistent pattern of these findings for emotion and deception, obtained in different laboratories, suggests that the generality of the ability to detect deceit would not be limited to some strange quirk in our choice of paradigms—any two paradigms that are able to generate strong emotions in the liars should show evidence for generality. We make this assertion on the basis of the large literature that has shown that across situations, people, and cultures, facial expressions of emotion not only appear similar but also are recognized at levels greater than chance (Ekman 1989, 1992, 1994). This is consistent with Ekman's (1985) argument that under high-stake deception situations, people are more likely to feel strong emotions when lying; the extent to which these signs of emotions betray deception is the extent to which a lie catcher would not have to know much about the situation in order to infer accurately the presence or absence of deception. Conversely, if strong emotions are not elicited, then one would expect deception to be betrayed mostly through "thinking" clues (words, factual descriptions, pauses, long speech latencies, speech errors, etc.; see DePaulo et al., 1985, for a review). Because thinking clues appear to be more specific to particular situations than do emotional clues, one would expect that in deception situations that do not arouse strong emotions, the ability to detect deceit should be specific to the situations.

These results highlight the fact that judging deception is a two-step process (e.g., Bond et al., 1992; Burgoon & Walther, 1990; DePaulo et al., 1980; Ekman, 1985; Ekman & Friesen, 1969; Kraut, 1978). The first step in the process is to recognize a sign, a clue, a behavior that violates expectations, or an emotion displayed by a target person that is at odds with his or her verbal line. The second step in the process is to interpret those clues accurately. For example, is the person feeling anxious because he or she is lying, or is that person truthful but afraid of being disbelieved? What a high-stake, emotion-eliciting paradigm does is to make more evident the signs of emotion, thus facilitating the recognition phase of the judgment. Our results show that people who are good at spotting these clues, as measured through the microexpression test, stand a better chance of successfully completing the interpretation phase of the judgment process, and if the situation is such that the clues to emotion are correlated with deception, then an observer who can recognize these emotion signs will outperform an observer who fails to recognize these signs. Alternatively, if these emotional signs are not present or are situation specific, then any advantage brought about by higher skill in recognition of emotion will be nullified, and one would expect situational specificity in the ability to detect deception. This is exactly what the psychological literature, which has focused on lower stake situations, has reported over the years. To use a sports analogy, one cannot win a championship without first making the playoffs; thus one cannot accurately detect deception without first noticing some quirk in the deceiver's behavior.

These results also highlight the notion that deception detection ability may not be a unitary construct, such as mathematics ability or reading skill (cf. Knapp & Comadena, 1979). Because there is no singular sign of deceit, and for some people or situations, there are no signs of deceit at all, there can be no singular strategy that would be perfectly successful in detecting deceit. In other words, there is no universal algorithm for detecting deception (Ekman, 1985; Zuckerman et al., 1981). Thus, when we discuss deception detection ability, we must take into account that the term is shorthand for a number of different skills and abilities—some related, and some not.

Our results and the results of others suggest that one such skill would be the recognition of emotion, as assessed by our microexpression test or by some other test, such as the Profile of Nonverbal Sensitivity (Rosenthal, Hall, DiMatteo, Rogers, &

Archer, 1979). This is by no means the only skill involved in detecting deceit; other possible skills are verbal comprehension, verbal reasoning, logic, and so forth. This suggests that any strategy to uncover the link between more global personality traits such as extraversion or conscientiousness and the detection of deceit is relevant only to the extent that it captures or is composed of more basic-level skills such as recognition of emotion.

If it is these basic skills, such as the ability to recognize fleeting emotions, that impact deception detection ability, then one could predict that people who are forced to thoroughly develop these skills—because of life circumstances or whatever—should outperform others when detecting deceit. We have preliminary evidence that one such group, individuals with left hemisphere brain damage, who cannot process speech and must rely exclusively upon nonverbal behavior to assess communications from others, tends to outperform normal, intact groups when detecting high-stake deceit (Etcoff, Ekman, Frank, Magee, & Torreano, 1992). We are following up this work with similarly impaired groups.

Finally, these studies have implications for improving people's abilities to detect deceit. The deception detection training studies have shown modest yet statistically significant improvements in accuracy (usually between 5% and 10% increases in accuracy; deTurck & Miller, 1990; Zuckerman et al., 1984). However, rarely do the groups receiving training surpass 70% accuracy. Yet in Experiment 1, 19% of the participants scored at or above 70% accuracy on both tests, and 6% scored at or above 80% on both (in Experiment 2, the numbers were 13% and 3%, respectively). This suggests two, possibly contradictory, approaches to improving the deception detection abilities of personnel in organizations that deal with high-stake deceptions in psychiatric, customs, and law enforcement situations. The first approach would be to select those individuals who have shown consistent ability to detect lies to be the organizations' interviewers, rather than spend the resources to train all individuals in the organization to be better detectors of deceit. The second approach would be to train these individuals to recognize not deception but emotion. The utility of each of these approaches could be determined empirically, and the exact skills, strategies, and experiences that make someone a good or poor detector of deceit could be isolated.

Finally, deception occurs in day-to-day life in many situations and under a variety of circumstances. In order to fully understand the processes of deception and detection of deceit, researchers must create realistic paradigms that cover both high- and low-stake situations. Researchers have been quite successful at documenting the interpersonal and situational interaction processes involved in day-to-day, low-stake sorts of deception situations and the important questions they address about communication and human nature (e.g., DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996; DePaulo et al., 1980). Our research is now beginning to unravel the interaction processes involved in high-stake deception situations, in which the successful detection of deceit may be critical to health and public safety.

References

Bond, C. F., Jr., Omar, A., Mahmoud, A., & Bonser, R. N. (1990). Lie detection across cultures. *Journal of Nonverbal Behavior, 14,* 189–204.

Bond, C. F., Jr., Omar, A., Pitre, U., Lashley, B. R., Skaggs, L. M., & Kirk, C. T. (1992). Fishy-looking liars: Deception judgment from expectancy violation. *Journal of Personality and Social Psychology, 63,* 969–977.

Burgoon, J. K., & Walther, J. B. (1990). Nonverbal expectancies and the evaluative consequences of violations. *Human Communication Research, 17,* 232–265.

DePaulo, B. M. (1994). Spotting lies: Can humans learn to do better? *Current Directions in Psychological Science, 3,* 83–86.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70,* 979–995.

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., & Rosenthal, R. (1979a). Ambivalence, discrepancy, and deception in nonverbal communication. In R. Rosenthal (Ed.), *Skill in nonverbal communication* (pp. 204–248). Cambridge, MA: Oelgeschlager, Gunn, & Hain.

DePaulo, B. M., & Rosenthal, R. (1979b). Telling lies. *Journal of Personality and Social Psychology, 37,* 1713–1722.

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

DePaulo, B. M., Zuckerman, M., & Rosenthal, R. (1980). Humans as lie detectors. *Journal of Communication, 30,* 129–139.

deTurck, M. A., & Miller, G. R. (1990). Training observers to detect deception: Effects of self-monitoring and rehearsal. *Human Communication Research, 16,* 603–620.

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

Ekman, P. (1989). The argument and evidence about universals in facial expressions of emotion. In H. Wagner & A. Manstead (Eds.), *Handbook of psychophysiology: The biological psychology of the emotions and social processes* (pp. 143–164). New York: Wiley.

Ekman, P. (1992). Facial expressions of emotion: New findings, new questions. *Psychological Science, 3,* 34–38.

Ekman, P. (1994). Strong evidence for universals in facial expression: A reply to Russell's mistaken critique. *Psychological Bulletin, 115,* 268–287.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology: II. *Journal of Personality and Social Psychology, 58,* 342–353.

Ekman, P., & Frank, M. G. (1993). Lies that fail. In M. Lewis & C. Saarni (Eds.), *Lying and deception in everyday life* (pp. 184–200). New York: Guilford Press.

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. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues.* Englewood Cliffs, NJ: Prentice Hall.

Ekman, P., & Friesen, W. V. (1978). *The facial action coding system.* Palo Alto, CA: Consulting Psychologists Press.

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

Ekman, P., Friesen, W. V., & Scherer, K. (1976). Body movement and voice pitch in deceptive interaction. *Semiotica, 16,* 23–27.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist, 46,* 913–920.

Ekman, P., O'Sullivan, M., Friesen, W. V., & Scherer, K. (1991). Invited article: Face, voice, and body in detecting deceit. *Journal of Nonverbal Behavior, 15,* 125–135.

Etcoff, N. L., Ekman, P., Frank, M. G., Magee, J., & Torreano, L. (1992, August). *Detecting deception: Do aphasics have an advantage?* Paper presented at the 1992 Annual Convention of the International Society for Research in Emotion, Pittsburgh, PA.

Frank, M. G. (1992). Commentary: On the structure of lies and deception situations. In S. Ceci, M. DeSimone Leichtman, & M. B. Putnick (Eds.), *Cognitive and social factors in early deception* (pp. 127–146). Hillsdale, NJ: Erlbaum.

Frank, M. G., & Ekman, P. (1997). *Facial behavior betrays deception.* Manuscript in preparation.

Geizer, R. S., Rarick, D. L., & Soldow, G. F. (1977). Deception and judgmental accuracy: A study in person perception. *Personality and Social Psychology Bulletin, 3,* 446–449.

Haggard, E. A., & Isaacs, K. S. (1966). Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy. In L. A. Gottschalk and A. H. Auerbach (Eds.), *Methods of research in psychotherapy* (pp. 154–165). New York: Appleton Century Crofts.

Hayano, D. M. (1980). Communicative competency among poker players. *Journal of Communication, 30,* 113–120.

Horowitz, S. (1989). *The role of control questions in the physiological detection of deception.* Unpublished doctoral dissertation, University of Utah, Salt Lake City.

Knapp, M. L., & Comadena, M. E. (1979). Telling it like it isn't: A review of theory and research on deceptive communications. *Human Communication Research, 5,* 270–285.

Kraut, R. E. (1978). Verbal and nonverbal cues in the perception of lying. *Journal of Personality and Social Psychology, 36,* 380–391.

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

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

Levenson, R. W. (1988). Emotion and the autonomic nervous system: A prospectus for research on autonomic specificity. In H. L. Wagner (Ed.), *Social psychophysiology and emotion: Theory and clinical applications* (pp. 17–42). New York: Wiley.

Levenson, R. W., Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology, 27,* 363–384.

Matsumoto, D., & Ekman, P. (1988). *Japanese and Caucasian facial expressions of emotion (JACFEE)* [Slide set, available from first author]. San Francisco State University.

Mehrabian, A. (1971). Nonverbal betrayal of feeling. *Journal of Experimental Research in Personality, 5,* 64–73.

O'Sullivan, M., Ekman, P., Friesen, W. V., & Scherer, K. (1992). *Judging honest and deceptive behavior.* Unpublished manuscript.

Podlesny, J. A., & Raskin, D. (1977). Physiological measures and the detection of deception. *Psychological Bulletin, 84,* 782–799.

Rosenthal, R., Hall, J. A., DiMatteo, M. R., Rogers, R. L., & Archer, D. (1979). *Sensitivity to nonverbal communication: The PONS test.* Baltimore: Johns Hopkins University Press.

Snyder, M. (1974). The self-monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30,* 526–537.

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., & Driver, R. E. (1985). Telling lies: Verbal and nonverbal correlates of deception. In W. A. Siegman & S. Feldstein (Eds.), *Multichannel integration of nonverbal behavior* (pp. 129–147). Hillsdale, NJ: Erlbaum.

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

Received April 19, 1994
Revision received July 1, 1996
Accepted July 5, 1996 ■