INTERPERSONAL DECEPTION: V. ACCURACY IN DECEPTION DETECTION

JUDEE K. BURGOON, DAVID B. BULLER, AMY S. EBESU, AND PATRICIA ROCKWELL

Previous research on accuracy in deception detection has typically occurred in a noninteractive context, which has resulted in many potentially salient influences being ignored. Guided by interpersonal deception theory, the current experiment examined the influences of suspicion, deception type, question type, relational familiarity, and expertise on accuracy in detecting truth and deceit. An adult sample of novices and a second sample of experts (military intelligence instructors and related military personnel) participated in interviews with strangers or acquaintances during which interviewees gave some truthful answers and some deceptive answers, the latter being one of three types. Interviewers, half of whom were induced to be suspicious, followed a standard interview protocol that introduced different question strategies. Results showed that (a) accuracy was much higher on truth than deception, (b) novices were more accurate than experts, (c) accuracy depended on type of deception being perpetrated and whether suspicion was present or absent, (d) suspicion impaired accuracy for experts, (e) truth-biases intensified with familiar others, especially when interviewers were suspicious, and (f) question strategy ameliorated or aggravated inaccuracy.

"Lies, Lies, Lies!," screamed the October 1992 cover of *Time* (Gray, 1992). Coupled with regularly appearing articles on the erosion of business ethics and the rise of cheating among students, it served notice that the longstanding assumption of truth in communication has become battered and frayed. Even so, increasing public cynicism about the prevalence of honesty and truthfulness seems not to have translated into greater awareness of the amount of deceit being perpetrated in everyday conversations nor improved people's ability to detect such deceptions. Substantial evidence attests to poor deception detection ability; people are frequently no more accurate than they would be by chance alone (Bauchner, Kaplan, & Miller, 1980; DePaulo, Stone, & Lassiter, 1985; Kalbfleisch, 1985; Kraut, 1980; Miller & Stiff, 1993; O'Sullivan, Ekman, & Friesen, 1988). Although being oblivious to duplicity can be an adaptive response in certain contexts (Knapp & Comadena, 1979), more often, accuracy in judging truth and detecting deception is highly functional and, in occupations such as law enforcement, secret service, customs inspection, and military intelligence, is imperative. Determining what factors improve or impair accuracy therefore becomes an important research objective.

An important consideration is whether sender and receiver actually interact with one another. Given that past investigations have been conducted primarily in noninteractive settings, many factors relevant to interpersonal interaction have not been studied systematically or simultaneously. In fact, it may be

Judee Burgoon is professor and David Buller is associate professor of communication at the University of Arizona, where Amy Ebesu is a doctoral student. Patricia Rockwell is assistant professor of communication at the University of Southwestern Louisiana. *This project was supported by the U.S. Army Research Institute (Contract #MDA903-90-K-0113) and the U.S. Army Research Office (Grant #30235-RT-AAS). The views, opinions, and findings in this report are those of the authors and should not be construed as an official Department of the Army position, policy, or decision. The authors thank Dr. Beverly Knapp and Ms. Anne Scott of the U.S. Army Research Institute's Field Office at Fort Huachuca, Arizona for their support and assistance.*

premature to conclude that detection ineptitude is widespread because past findings may not generalize to interactive interpersonal communication. Additionally, some past conclusions may have been faulty due to reliance on imprecise or inappropriate measurement strategies (Ekman, O'Sullivan, Friesen, & Scherer, 1991; McCornack & Levine, 1990). Guided by an interactive theory of deception, the current investigation therefore examined detection accuracy within an interactive context, using improved measures of accuracy and judgment biases.

Accuracy in Interactive Contexts

A criticism lodged against much previous deception theorizing and research is that it has characterized deception as a psychological rather than a communication event (Knapp, Cody, & Reardon, 1987). Deception has been viewed as a unidirectional activity in which senders create deceitful messages that are passively absorbed and judged by receivers. The main foci have been the internal psychological processes governing senders' overt manifestations of deception and receivers' ability to accurately infer deceit from such cues. As a consequence, research seldom permitted deceivers and deception detectors to interact with one another (Toris & DePaulo, 1985). Much of what we know about deception therefore comes from noninterpersonal and/or noninteractive contexts and may not generalize to the common circumstances where people interact face-to-face.

As a counterpoint to this orientation, Buller and Burgoon have advanced interpersonal deception theory (IDT) (Buller & Burgoon, 1994, in press; Burgoon, 1992; Burgoon & Buller, 1994; Burgoon, Buller, Dillman, & Walther, 1994), which emphasizes the impact that unique interpersonal communication constraints and demands are likely to place on deceptive encounters. Among the key assumptions and propositions of IDT are that (a) interpersonal deception activates strategic behavior on the part of both sender and receiver—the sender, to create a credible performance, the receiver, to determine the credibility of the sender's communication; (b) as the interaction dynamic evolves, both people's behavior changes and influences one another; (c) the multiple communication functions that must be accomplished simultaneously may facilitate or hinder deception and detection success; and (d) interaction promotes expectations and familiarity that guide behavior and judgments.

Examining deception as an interpersonal, face-to-face communication event implicates a number of factors that should influence deception detection accuracy. Recognition that senders and receivers engage in strategic behavior shifts attention from static, preinteractional variables such as age, sex, and motivation to interactional variables, i.e., to what they do and think while interacting. For senders, this includes examining the type of deception they encode, which may be differentially successful. For receivers, this includes examining suspicion, which can be seen as the counterpart to sender's deceit, and the communication strategies that receivers employ to discern the truth. A communication perspective also considers the relationship between sender and receiver and the degree to which they are familiar with one another. Recognition that deception frequently occurs among acquainted people and that such familiarity results in different communication patterns underscores the importance of examining forms of familiarity and their impact on successful deception and detection.

Thus, four possible moderators of accuracy derived from IDT are deception type, suspicion, receiver probing strategies, and familiarity.

Deception Type

Deception includes a wide array of forms other than an outright lie (Hopper & Bell, 1984; Turner, Edgley, & Olmstead, 1975). Among the features differentiating them are amount and sufficiency of information, degree of truthfulness, clarity, relevance, ownership, and intent (Buller & Burgoon, 1991; McCornack, 1992). Taking these various features into consideration led us to distinguish among three relatively distinct classes of deceptions: falsification, concealment, and equivocation.

It stands to reason that these different forms of deception should be differentially successful in deceiving others based on their adherence to conversational norms, expectations, and requirements for completeness, truthfulness, relevance, and so forth (see Grice, 1989). Metts (1989) found that falsification is the most frequently used form of deception in close relationships, followed by concealment (omissions) then equivocation (distortions). One might therefore expect falsification to be the most successful form of deception because it is most prevalent and therefore most practiced. However, Bavelas, Black, Chovil, and Mullett (1990) suggested that equivocation is the safest deception because it is closest to truth and thus most people feel more comfortable equivocating than using other deception forms. Based on this premise, Burgoon, Buller, Guerrero, and Feldman (1993) hypothesized that equivocation would be more successful than concealment or falsification. Instead, they found that deceivers only thought they were more successful when they used equivocation, but observers did not perceive equivocators as more believable and successful. Burgoon, Buller, Guerrero, Afifi, and Feldman (in press) provided a possible explanation in a companion study which analyzed the ways in which information was managed in each type of deception. Deceivers and trained coders rated deceiver answers on several information manipulation dimensions related to veracity (actual or apparent truthfulness), personalization (taking ownership of or disassociating self from one's statements), informational and conversational completeness, directness/relevance, and clarity. Equivocation was rated by observers as the least complete, direct, clear, personalized, and truthful; falsification was seen as fairly complete, direct, clear, and truthful (even though deceivers reported the most dishonesty and disassociation from their message when using falsification). Based on their perceived communicative features, then, falsification would be the least readily detected form of deception, i.e., the one producing the lowest accuracy. Given the mixed empirical evidence about the normative frequency of the three types, the relative ease of producing them, and their differential adherence to conversational requirements, we posed a nondirectional hypothesis:

H1: Detection accuracy differs across deception types.

Receiver Suspicion

Very little research has examined the effects of receiver suspicion on deception detection and that which has been conducted has yielded inconclusive findings. Some studies found that suspicion failed to improve detection accuracy

and decreased receivers' confidence in their judgments (Buller, Strzyzewski, & Comstock, 1991; Toris & DePaulo, 1985). McCornack and Levine (1990) found that moderate suspicion increased accuracy, whereas Zuckerman, Spiegel, DePaulo, and Rosenthal (1982) found that suspicion decreased accuracy in decoding positive affect.

The inconsistent findings may be due to receiver suspicion affecting detection accuracy in sometimes contradictory ways. First, it may affect how receivers process information from senders. As stipulated in IDT, one of the primary assumptions in interpersonal encounters is that people are telling the truth, that is, they enter interpersonal interactions with a truth-bias (Buller, Strzyzewski, & Hunsaker, 1991; Kalbfleisch, 1992; Levine & McCornack, 1991; Riggio, Tucker, & Throckmorton, 1987).[1] There are exceptions: Some people who are suspicious by nature exhibit a lie-bias (Levine & McCornack, 1989). O'Sullivan et al. (1988) argued that a truth-bias may improve detection accuracy with honest senders, but a lie-bias may improve detection accuracy with dishonest senders. Suspicion may alter this basic equation by making people more wary and vigilant. Heightened skepticism and attention might improve accuracy in detecting deception among those who typically overestimate truth by causing them to evaluate incoming information more critically, but it would simultaneously reduce their accuracy in detecting truthful messages. Conversely, greater skepticism might exacerbate a lie-bias that already leads to underestimates of the truth but improve detection of deceptions, although suspicion could as easily impair accuracy by encouraging receivers to seek information confirming their lie-bias. Thus, suspicion might help or hinder accuracy, depending on one's basic predisposition to judge others as truthful or dishonest and whether the information being judged is truthful or untruthful. At the same time, research by Buller and associates (Buller & Hunsaker, in press; Buller, Strzyzewski, & Hunsaker, 1991) found that participants in an interaction were more inclined to think a sender was truthful than did observers, even when informed that deception had been manipulated, indicating that a truth-bias might be intensified in interactive contexts, counteracting any benefit gained from increased suspicion.

Second, IDT postulates that suspicious receivers may alter their communication patterns in ways that alert senders to their doubts. Research by Silverman, Rivera, and Tedeschi (1979) found that participants in a suspicion feedback condition changed their behavior. Recent experiments also have shown that receivers leak their suspicions nonverbally and that senders are aware of suspicion when it is present (Buller, Strzyzewski, & Comstock, 1991; Burgoon, Buller, Dillman, & Walther, 1994). This feedback may lead senders to adjust their behavior, whether they are lying or not. However, efforts to suppress leakage cues and to create a credible performance often backfire, perhaps because self-conscious attempts to improve performance increase arousal, cognitive load, and unnaturalness (Buller, Strzyzewski, & Comstock, 1991; Burgoon & Buller, 1994). Ironically, even truthtellers may look dishonest, causing receivers to misjudge their truthful performances as deceptive. At the same time, socially skilled communicators may take advantage of the feedback to craft believable deceptions (Burgoon, Buller, Guerrero, & Feldman, 1993; DePaulo,

Blank, Swaim, & Hairfield, 1992), causing receivers to misjudge senders' deceptive performances as truthful. Clearly, suspicion is tricky business. There may be only a small range in which being suspicious is productive, which led us to hypothesize that suspicion largely jeopardizes accuracy.

H2: Suspicious receivers are less accurate in detecting deception and truthtelling than nonsuspicious receivers.

Type of Questioning

As active agents during interpersonal exchanges, receivers may use a variety of strategies to discern a sender's truthfulness, including use of different types of questions to probe for more information. Several studies that have investigated the use of probing (follow-up) questions (Buller, Comstock, Aune, & Strzyzewski, 1989; Buller, Strzyzewski, & Comstock, 1991; Stiff & Miller, 1986) have failed to confirm that probing per se improves deception detection accuracy, in part because deceivers altered their performances accordingly (Kalbfleisch, 1992). But qualitative differences in question strategies may make a difference. Stiff and Miller (1986) compared positive and negative questions and found that deceivers and truthtellers who were questioned using negative probes were seen as more truthful by observers than when positive probes were used.

What may be even more relevant than question valence is whether questions catch deceivers off guard and cause them to divulge information or leak nonverbal deception cues before they are able to monitor and manage their behavior. Two tactics for doing so are to ask senders unexpected questions and to return to previous questions later in a questioning sequence (McBurney & Comadena, 1992; O'Hair, Cody, Wang, & Chao, 1990). Support for the idea that question type may influence detection accuracy is implicit in the greater success found for planned, as opposed to spontaneous, lies (deTurck & Miller, 1990; Miller, deTurck, & Kalbfleisch, 1983). The ability to plan and rehearse lies seems to enhance deceiver success because deceivers exhibit more control over their nonverbal behaviors with planning time (O'Hair, Cody, & McLaughlin, 1981; Zuckerman & Driver, 1985). Thus, evoking spontaneous lies through the use of unexpected questions may instead aid receivers in detecting deception. By the same token, returning to a previously asked question at a later point in a conversation, after a sender presumably has answered it completely, may elicit revealing information or nonverbal leakage because coordinating deceptive answers with previous ones is a difficult task, especially if the need to do so was not anticipated. Also, the tendency to be reticent and nonspecific when deceiving may make repetitive answers appear overly redundant and therefore less credible. Alternatively, in interactive situations, spontaneous answers may appear more sincere because such answers may include nonverbal qualities characteristic of spontaneity itself (e.g., surprised laughter, nonfluencies) or may employ a different type of deception such as falsification (Hopper & Bell, 1984; Metts & Chronis, 1986) that is more difficult to detect. Similarly, answers to a "repeat" question may allow a deceiver to repair previous answers and concoct a more plausible reply. This uncertainty led us to pose a research question:

RQ1: How do unexpected and repeat questions affect detection accuracy?

Familiarity

A final potential moderator of detection accuracy is familiarity. Familiarity can take at least two forms: relational or behavioral. Relational familiarity, which refers to the degree to which interactants are acquainted with one another, combines personal knowledge of senders' background and habits with first-hand experience with their particular interaction style, which should enable receivers to utilize verbal and nonverbal information effectively to make accurate judgments of truthfulness. Indeed, several studies (Comadena, 1982; Kalbfleisch, 1985; McBurney & Comadena, 1992; McCornack & Levine, 1990; Miller et al., 1983) found this to be the case. However, other research has shown that the opposite occurs, possibly because relational familiarity creates a positivity bias: People are more tolerant of, or favorable toward, people with whom they have a personal relationship. Supporting a truth-bias are one study that found that truth-bias increased with greater familiarity (Stiff, Kim, & Ramesh, 1992) and three other studies that found that receivers judged friends as more trustworthy than strangers and interpreted friends' behaviors more favorably (Buller, 1987; Buller, Strzyzewski, & Comstock, 1991; Burgoon, 1992). Because it is unclear whether relational familiarity in the context of actual interaction improves accuracy through the greater knowledge one has of partner behaviors or impairs accuracy by exacerbating the truth-bias, we posed the following question:

RQ2: How does relational familiarity affect detection accuracy?

A second kind of familiarity is behavioral familiarity. This label can actually refer to prior exposure to a sample of an interactant's behaviors (such as seeing baseline behavior before judging an individual's deceptiveness), training in cues associated with deception, or experience with deception behaviors in general (as in the case of professionals in investigatory agencies). Research on the first two types of behavioral familiarity has typically shown some gain in deception detection accuracy with baseline exposure and training (Brandt, Miller, & Hocking, 1980, 1982; deTurck, Harszlak, Bodhorn, & Texter, 1990; Manstead, Wagner, & MacDonald, 1986; Mongeau, 1988; Zuckerman, Koestner, & Colella, 1985). However, when behavioral familiarity takes the form of expertise, in which training is coupled with experience, it can become counterproductive. Research on expert accuracy has confirmed that experts such as law enforcement officers and customs agents, like laypeople, are often no more accurate than chance and may even be less accurate than nonexperts (DePaulo & Pfeifer, 1986; Ekman & Friesen, 1974; Ekman & O'Sullivan, 1991; Kohnken, 1987; Kraut & Poe, 1980).

Why might this be? For one, expertise often entails other factors such as age, high motivation, and access to other sources of information that accompany an expert's occupation (Kalbfleisch, 1992). DePaulo et al. (1985) also speculated that experts gain little self-correcting feedback when they commit false-negatives (i.e., believing a dishonest person is honest). We believe during actual communication, they also seldom gain feedback when they commit false-positives (believing an honest person is dishonest) and so may become overly confident in their detection abilities. Together, these factors may induce chronic

suspicion resulting in a lie-bias that impairs judgment. Because we suspected that interaction alone would be insufficient to overcome experts' chronic suspicion and resultant lie-bias, we hypothesized,

H3: Behavioral familiarity in the form of expertise impairs accuracy.

Measurement of Accuracy

We alluded earlier to the fact that past approaches to measuring accuracy may have led to misestimates of it. One common approach has been to measure accuracy with dichotomous truth-lie judgments (Kalbfleisch, 1985). This is problematic because the truthfulness of any given utterance often varies in degree rather than being an on-off variable. Respondents forced to choose truth or lie are not permitted to display such variability and so may opt in the direction of leniency, producing results that look like a truth-bias solely as a measurement artifact. Dichotomous scales also cannot detect subtle variations in type of deception. For example, deceptions that include equivocations or concealments cannot be assessed accurately by scales that only measure truth or falsity.

In other cases, continuous measures of deceptiveness or honesty, which improve measurement sensitivity, are used but judgments are made at the conclusion of an entire message or interaction (e.g., Buller, Strzyzewski, & Hunsaker, 1991; DePaulo, LeMay, & Epstein, 1991). Like their dichotomous counterparts, such judgments may be inappropriate when messages are longer than single utterances. Extended discourse entails multiple statements, some of which may be truthful and some of which may be deceptive. Single judgments may be insensitive to variability in the amount of deceit present in a complete message or interaction.

Third, most studies have relied on observer judgments as the measure of accuracy (e.g., Bauchner, Brandt, & Miller, 1977; deTurck & Goldhaber, 1988; Ekman & Friesen, 1969; Fay & Middleton, 1941; Littlepage & Pineault, 1979; Maier & Thurber, 1968; McCornack & Parks, 1986), while only a handful of recent studies have used participant judgments (e.g., Buller & Aune, 1987; Buller et al., 1989; Buller, Strzyzewski, & Comstock, 1991; McCornack & Parks, 1990). The inference often drawn is that the two are interchangeable, yet abundant research documents differences between participant and observer perspectives (Zebrowitz, 1990). Within the deception arena specifically, Comadena (1982) argued that participants' accuracy ratings are different from observers. Buller, Strzyzewski, and Hunsaker (1991) and Buller and Hunsaker (in press) found that participants had a stronger truth-bias than observers and, in one study, were less accurate than observers in detecting deception. Given the focus of IDT on interpersonal interaction, participants' judgments should be most salient.

To rectify these problems, we used continuous rather than dichotomous measures of truthfulness; we obtained truthfulness estimates for responses to each question in an experimental interview, providing multiple judgments over the course of the interaction; and we relied on participants' rather than observers' judgments. In addition to obtaining honesty attributions from receivers, we compared receiver perceptions to sender self-reports to obtain accuracy and bias measures.

METHOD

Participants

Participants (N = 132) consisted of (a) a sample of nonexperts drawn from a metropolitan southwestern community who were recruited to participate in an experiment purportedly on interviewing skills and (b) a sample of experts drawn from a military human intelligence school who were encouraged but not required to participate in this study as part of an ongoing applied research program. Thus, both samples consisted of volunteers.

Community participants (hereafter called novices), who participated in exchange for individualized communication training and feedback, consisted of 72 adults (37 males, 35 females). After completing a social skills inventory at the jury assembly room and indicating their willingness to participate in the interview portion of the study, these participants were then contacted by telephone to arrange an interview time and asked if they would be willing to bring an acquaintance or friend with them. Those who did not were paired with strangers when they arrived. This resulted in 18 acquainted and 18 stranger dyads, 17 same-sex dyads and 19 mixed-sex dyads in the novice sample. The sample of experts consisted of 60 adults (50 males, 10 females) who were paired with acquaintances, friends, or strangers according to availability of common times for completing the interviews. This resulted in 11 acquainted dyads and 19 stranger dyads, 20 same-sex dyads and 10 mixed-sex dyads. The unequal cell size was due to availability problems during the scheduled experiment times and rotations of personnel to overseas or other assignments.

The novices ranged in age from 19 to 75 years (M = 39.9 years). The median education level was two-year college graduate and the mode was some college or vocational education. Ethnically, 88% were Caucasian, 1% were Black, 7% were Hispanic, and 4% were Native American or Alaskan Native. The demographic characteristics of experts were similar to novices. They ranged in age from 25 to 56 years (M = 36.8 years); the median for education level was two-year college graduate and the mode was some college or vocational education; and their ethnic distribution was 81% Caucasian, 9% Black, 7% Hispanic, 2% Native American or Alaskan Native, and 2% Asian or Pacific Islander.

All experts assigned to the role of interviewer (ER) and approximately two fifths of those assigned to the role of interviewee (EE) were instructors at the military intelligence school. Earlier in their careers they had received training in conducting interviews and interrogation (how to ask questions and probe for additional information of prisoners, field agents, and other information targets). This training also consisted of feedback on their ability to identify accurate information and informal advice on how to detect nontruthful or incomplete messages. Experts subsequently had several years of experience in one or more of the following areas: (a) tactical interrogation (M = 6.74 years for the 23 ERs who had such experience; M = 5.33 years for the 12 EEs who had such experience), (b) strategic interrogation (M = 3.38 years for the 16 ERs, M = 6.00 for the 10 EEs), (c) tactical interviewing (M = 7.36 years for the 11 ERs, M = 7.44 for the 9 EEs), and (d) strategic interviewing (M = 3.83 years for the 12 ERs, M = 10.10 for the 10 EEs). These positions required making decisions about the truthfulness and accuracy of information. Thus, these experts had

much more extensive knowledge and experience in conducting interviews and evaluating interviewee behavior than did novices. The remaining interviewees consisted of military and civilian personnel assigned to the post.

For the novices, 50% of the dyads were strangers and 50% were acquainted, i.e., they mutually reported their relationship as being acquaintances, casual friends, good friends, best friends, or family members. Among acquainted dyads, the modal response (47%) was "good friend." In addition, the acquainted dyads reported having known each other for an average of 9.0 years, interacting with each other at least once a day (94%), and on average rated their relational closeness a 5.03 on a 7-point scale which asked them to describe their relationship in terms of whether or not they shared personal and important information, spent time together, engaged in a variety of activities, and expressed their attitudes and opinions to each other. Due to time constraints, the military sample recorded only whether their partner was a stranger, acquaintance, or friend. Over half (56%) were unacquainted; among acquainted dyads, 37% were friends. The comparative distributions suggest that "familiars" within the novice sample had closer relationships on average with their interaction partner than did experts, which should have given them greater behavioral familiarity.

Procedures

Participants from the community were tested at the Communication Research Laboratory (a comfortable apartment-like setting equipped with a one-way mirror). Experts were tested at a nearby military installation in a facility used for interrogation training. The testing rooms were smaller than those used for the community sample, and experts were observed by remote video from a centralized location.

When participants arrived at the testing site, they were greeted by two research assistants and told that they would be participating in a videotaped interview. One participant was randomly assigned the role of interviewer (ER) and the other the role of interviewee (EE). They were told that the interview would be based on a series of questions that EEs would answer in advance. Participants were separated for training and induction.

Truth/deception, deception type, and question type inductions. After completing pretest measures, EEs received the deception manipulations from an assistant who explained that it is not always in one's best interest to tell the whole truth and that they were to practice their skill at adapting to such situations. The assistant showed the EEs a list of 15 questions to be asked in the interview. Truth/deception was manipulated by instructing the EEs to answer truthfully to the first three questions in the upcoming interview, after which all answers were to be deceptive.[2] This truthful baseline was intended to give EEs an opportunity to acclimate to the task and setting as well as to provide a control comparison for the different types of deception. We chose to minimize the number of truthful questions because the previous pilot experiment had already compared a full truthful interview to a full deceptive one and the focus in this investigation was on accuracy in detecting deception. It should be noted that there are numerous ways to order question type and deception replies. For example, one could randomize when truth or deception were to occur or alternate between truthful and deceptive answers in some prearranged fashion. However, this is problem-

atic for analyses because it confounds question content with truth/deception. Moreover, the pilot experiment had indicated that the task of creating a particular type of deception on demand was difficult in itself; requiring EEs to alternate between truthful and deceptive answers might have exceeded their abilities and resulted in many mistakes in answering. While we recognize that asking for two truthful questions initially may have created the expectation of honesty, during the analysis of this data, the true questions and the false questions were either analyzed separately (with true and false questions treated as a within subjects factor in the design), or only the deception questions were analyzed. Also, it must be kept in mind that our intent was not to obtain a precise estimate of absolute accuracy, but rather, estimates of relative accuracy under different forms of deception and the presence or absence of suspicion.

Deception type was manipulated by instructing EEs to alter their answers by giving (a) completely untrue answers (falsification; $n = 23$), (b) vague, indirect, unclear, or ambiguous answers (equivocation; $n = 21$), or (c) answers that withheld, omitted, or avoided relevant information (concealment; $n = 22$). EEs were given samples of each type of deception, and assistants practiced asking them potential questions until it was apparent that the EEs could enact the assigned deceptive behavior.

The interview questions, derived or modified from a pilot experiment, asked for emotional, factual, and opinion information. They were chosen for their ability to elicit longer answers and to cover a wide range of content (unlike many previous experiments that have utilized only factual or emotional content). Question type was varied by adding a spontaneous question during the actual interview that did not appear in the list of questions shown to the EE prior to the interview (hereafter referred to as an unexpected question) and by adding a reprise question which referred to and probed for information reported in an earlier question (hereafter referred to as a repeat question).[3]

Suspicion induction. Simultaneously in another room, ERs completed appropriate pretest measures. Then, a second research assistant prepared the ERs by having them review the interview protocol and plan possible follow-up questions. Immediately preceding the start of the interview, the suspicion induction was introduced. Half of the ERs were told that, "in our research so far, we have discovered that some people are far less candid and truthful when answering these questions in a face-to-face conversation than when they are completing an anonymous questionnaire. Also, we find if they are not truthful on one question, they are likely to be less than 100% honest on others as well." They were informed that during the interview, the assistant would follow the EE's questionnaire answers and if it became apparent that the EE had deceived on one question, the assistant would surreptitiously signal the ER by walking past the interaction room after which the ER should "pay careful attention to your partner's answers to see if you can determine if he/she is lying or telling the truth on the remaining questions." A similar procedure was used at the military installation except that assistants signalled ERs with a light tap on a window only visible to the ER. The suspicion signal always coincided with the onset of deception, occurring after the first deception question.

Interview. When both ER and EE were trained, they were brought to the main room where the interview was conducted and videotaped. Interviews lasted

approximately 15 minutes. In the community sample, interviews were stopped after all the questions were asked or after 15 minutes were reached. In the expert sample, interviews continued until all questions were asked. Following the interview, participants were separated again, returned to their respective training rooms, given posttest measures and debriefed. Additionally, novice participants received communication skills training, and military participants engaged in a more extensive open-ended debriefing.

Dependent Measures

To measure honesty attributions and accuracy, we created three dependent measures. To measure absolute honesty attributions, ERs estimated EEs' truthfulness on each of the 15 questions on a scale from 0 (completely untruthful) to 10 (completely truthful). To measure accuracy, we modified a method introduced by McCornack and Parks (1990) to assess "leakage accuracy." They calculated discrepancies between a sender's actual attitude score and a judge's estimate of that attitude, and then averaged the overall differences. We refined this measurement by calculating separate accuracy scores for truthful answers and for deceptive answers. Accuracy was then computed as the absolute discrepancy between ER's estimate and EE's report. To measure biases, we used the signed discrepancy scores. A positive score means that the ER's estimate is higher than the EE's reported truthfulness and therefore reflects a truth-bias; a negative score means that the ER underestimated truthfulness and therefore reflects a lie-bias. This approach differs from one used by McBurney and Comadena (1992) to measure bias in which all scores are averaged and higher scores toward the truthful end of the scale are regarded as reflecting a truth-bias. In the absence of knowing how truthful or deceptive the sender's answer was, one does not really know if the estimate was actually "biased" or fairly accurate. One caveat when measuring accuracy on an interval-level scale is the ceiling effect. If EEs report being highly "truthful," truth cannot be overestimated by ERs because EEs' actual scores will be close to the upper limit of the truth scale, leaving ERs little leeway to overestimate it. To compensate for this, we also calculated separate bias scores for the truth and deception conditions.

As a manipulation check, ERs completed a suspicion measure (Buller, Strzyzewski, & Comstock; 1991; Burgoon, Buller, Dillman, & Walther, 1993) consisting of three subscales measuring belief that EE was deceptive (α reliability = .82), vigilance (α = .70), and motivation to detect deceit (α = .52).[4]

RESULTS

Manipulation Checks

Deception. A mixed model 2 (truth versus deception phase) × 3 (deception type) × 2 (suspicion) × 2 (relationship) × 2 (expertise) analysis of variance confirmed that the deception manipulation was successful and very potent. The analysis produced a main effect for truth/deception, $F(1,51) = 475.93$, $p < .0001$, $\eta^2 = .90$, an interaction between deception type and truth/deception, $F(2,51) = 17.80$, $p < .0001$, $\eta^2 = .41$, and a main effect for deception type, $F(2,51) = 22.61$, $p < .0001$, $\eta^2 = .47$. EE answers while deceiving were far less truthful ($M = 3.43$) than when truthtelling ($M = 9.59$), and degree of dishon-

esty differed across deception types: Falsification was the most deceptive ($M = 1.25$), followed by equivocation ($M = 4.12$) and concealment ($M = 5.15$).

Suspicion. The effects of suspicion on ERs' vigilance, motivation to detect, and belief that EE was deceptive were tested in a multivariate analysis of variance with the three suspicion measures as the dependent variables and deception type, suspicion, relationship, and expertise as the independent variables. Although a significant main effect for suspicion failed to obtain, $F(3,52) = .42, p > .10$, there were near-significant deception type by suspicion interactions on vigilance, $F(2,54) = 2.83, p = .067, \eta^2 = .10$, and motivation, $F(2,54) = 2.78, p = .071, \eta^2 = .09$, and the pattern of means was as hypothesized (suspicion belief $M = 3.83$, vigilance $M = 5.36$, motivation $M = 4.91$; nonsuspicion belief $M = 3.52$, vigilance $M = 5.12$, motivation $M = 4.72$). A follow-up analysis conducted on the single item, "I expected my partner to tell the truth," also produced a significant main effect for suspicion, $F(1,61) = 2.82, p < .05$ one-tailed, $\eta^2 = .05$, providing further weak support for the suspicion manipulation. Coupled with the fact that some suspicion effects did emerge in the hypotheses tests, these findings imply that suspicion was manipulated.

Other significant effects complicate interpretation of the suspicion manipulation. Main effects emerged for expertise, $F(3,52) = 6.24, p = .001, \eta^2 = .27$, and deception type, $F(6,104) = 2.89, p = .012, \eta^2 = .14$. Experts were more suspicious ($M = 4.44$), vigilant ($M = 5.62$), and motivated to detect deceit ($M = 5.05$) than novices (belief $M = 3.04$, vigilance $M = 4.93$, motivation $M = 4.62$). Thus, the results that follow have suspicion embedded within the expertise factor. Concealment also attenuated suspicious beliefs ($M = 2.74$) and motivation ($M = 4.69$), while equivocation elevated vigilance ($M = 5.64$) and motivation ($M = 5.17$). The deception type and deception by suspicion interactions mean either that deception type moderated suspicion during the interaction or confounded the postinteraction reporting of suspicion. This indicates that suspicion is fluid, changing as a sender's encoding changes.

Hypothesis 1: Deception Type

The primary analysis model used to test hypotheses and research questions was a reduced, mixed model repeated measures analysis of variance including *deception type*, *suspicion*, *relationship*, and *expertise* as between subjects factors and *role* (interviewer/interviewee) as a within subjects factor. Depending on the analysis, an additional within subjects factor was included. In some analyses, this factor was *truth/deception* (truthful answers/deceptive answers), which compared pooled answers for the first two "truthful" questions to pooled answers for the remaining "deceptive" questions; in others, the factor was *question type* (truthful/rehearsed/unexpected/repeated), which permitted comparisons among the specific types of questions asked. (Reduced models were obtained by pooling all terms from the saturated models that had F-values less than 1.0 in the error term.) Omnibus analyses were followed by appropriate planned or a posteriori contrasts. Variations on this basic analysis plan are noted where appropriate.

Hypothesis 1, that detection accuracy differs across deception types, was tested on the three dependent measures of honesty judgments, accuracy, and bias. The honesty judgments analysis produced a main effect for deception type, $F(2,59) = 9.80, p < .0001, \eta^2 = .26$; an interaction between deception type and

truth/deception, $F(2,59) = 18.91$, $p < .0001$, $\eta^2 = .40$; and an interaction among truth/deception, deception type, and role, $F(2,58) = 6.67$, $p = .003$, $\eta^2 = .20$. ERs saw the truthful answers as far more honest than the deceptive ones and among deceptive answers, concealment was seen as the most truthful and equivocation as the least truthful (see Table 1 for all means). Although this might imply that ERs were least accurate in judging concealment, the EEs reported being more truthful when concealing (see manipulation check). This underscores the need to measure EE's own reported deception and the discrepancy between the two. Based on the comparison of the aggregate means, the biggest discrepancy was between ER's honesty judgment of falsification and EE's report, i.e., ERs were least accurate judging this type of deception.

The analysis on accuracy scores differs from the preceding analysis in that it was based on a calculated discrepancy score for each dyad. Because the hypothesis concerns different forms of deception, the analysis omitted replies during truthtelling. A near-significant deception type by suspicion interaction emerged, $F(2,56) = 3.08$, $p = .076$, $\eta^2 = .11$. Although this analysis was less powerful than the preceding one, the moderate effect size warrants attention. The means again revealed that ERs were most accurate when judging equivocation *if they were suspicious* ($M = 2.98$; nonsuspicion $M = 4.18$) or judging concealment *if they were not suspicious* ($M = 3.45$; suspicion $M = 4.54$) and least accurate when judging falsification (suspicion $M = 4.28$, nonsuspicion $M = 4.86$).

The analysis on bias (which again omitted answers during truthtelling) adds further information by identifying the direction of errors (toward a truth-bias or toward a lie-bias). Results showed a main effect for deception type, $F(2,54) = 7.89$, $p = .002$, $\eta^2 = .25$. All biases were in the direction of a truth-bias (see Table 1). As might be expected, this was most pronounced for falsification inasmuch as falsification departs the most from the truth.

Overall, results strongly confirmed Hypothesis 1. Falsification produced the greatest inaccuracy. Concealment also produced inaccuracy, but only if the ER was suspicious. Equivocation generally produced the greatest accuracy, especially if the ER was suspicious, and produced the least bias. Most inaccuracies were in the direction of a truth-bias.

Hypothesis 2: Suspicion

Hypothesis 2, that suspicious receivers are less accurate in judging truth and detecting deception than nonsuspicious receivers, was tested on the same three dependent measures as Hypothesis 1, except that replies during truthtelling questions were also included. The three analyses failed to produce a significant main effect for suspicion: honesty judgment $F(1,59) = .07$, $p > .10$; accuracy $F(1,56) = .11$, $p > .10$; bias $F(1,56) = .07$, $p > .05$. However, a significant main effect did obtain on total accuracy within the expert sample, $F(1,21) = 3.50$, $p < .05$ one-tailed, $\eta^2 = .15$.[5] Experts were less accurate overall when suspicious than when not. Additionally, the honesty judgments analysis produced a suspicion by relationship interaction, $F(1,59) = 4.74$, $p = .040$, $\eta^2 = .08$. Suspicion made interviewers ascribe less honesty to strangers and more honesty to acquaintances, the net result being underestimates of strangers' truth and acquaintances' dishonesty (compared to interviewers in the nonsuspicion condition) (see Table 1). Thus Hypothesis 2 received qualified support. Suspicion

TABLE 1

MEANS FOR INTERVIEWER HONESTY JUDGMENTS, ACCURACY DISCREPANCIES, AND BIAS ON TRUTHFUL AND DECEPTIVE ANSWERS

Dependent Measure	Suspicion	No Suspicion	Falsification	Equivocation	Concealment
HONESTY JUDGMENT					
TRUTHFUL answers	7.88	8.39	8.24	7.79	8.16
With strangers	7.44	8.52			
With acquaintances	8.34	8.17			
DECEPTIVE answers	3.32	3.55	5.50	4.87	7.18
With strangers	5.23	6.72			
With acquaintances	6.08	5.05			
ACCURACY					
TRUTHFUL answers	2.08	1.80	1.96	1.90	1.95
With strangers	2.50	1.62			
With acquaintances	1.62	2.12			
DECEPTIVE answers	3.93	4.18	4.59	3.55	3.99
With strangers	3.83	4.35			
With acquaintances	4.04	3.88			
TOTAL ACCURACY	3.55	3.72	4.00	3.31	3.58
BIAS					
TRUTHFUL answers	−1.83	−1.08	−1.30	−1.62	−1.45
DECEPTIVE answers	2.35	2.51	4.31	.79	2.02

Note. Accuracy scores represent the absolute discrepancy between EE reports and ER estimates. Higher scores reflect greater discrepancies (i.e., less accuracy). Total accuracy combines the truthful and deceptive answers. Bias scores represent the signed difference between EE reports and ER estimates. Positive scores, which can be interpreted as a truth-bias, represent ERs overestimating the truthfulness of EE answers; negative scores, which can be interpreted as a lie-bias, represent ERs underestimating the truthfulness of EE answers (compared to what EEs reported).

impaired accuracy for experts and, for experts and novices alike, intensified predilections to judge acquaintances more favorably than strangers. As noted for Hypothesis 1, suspicion also interacted with deception type so that it had a deleterious effect when judging concealment but a beneficial effect when judging equivocation.

Research Question 1: Question Type

Research Question 1 addressed the differences in accuracy resulting from different question strategies by comparing answers to four different questions, one during truthtelling, one for which the answer had been planned and rehearsed, one that was unexpected, and one that repeated and probed an earlier question. Question type produced a significant main effect, $F(3,156) = 9.79$, $p < .0001$, $\eta^2 = .18$, but also interacted with expertise and with relationship by expertise, $F(3,156) = 3.66$, $p = .02$, $\eta^2 = .08$. Follow-up contrasts confirmed that: (a) ERs were more accurate judging answers to the truthful question than the three deceptive ones, (b) novices were more accurate than experts except with repeat questions (especially when judging acquaintances), (c) both novices and experts were least accurate when employing repeat questions with strangers, (d) among novices, greatest accuracy in detecting deception was achieved when using unexpected questions, and (e) among experts, accuracy was far worse judging strangers than acquaintances when employing repeat questions or questions eliciting rehearsed answers (see Table 2 for means).

TABLE 2

ACCURACY CELL MEANS FOR QUESTION TYPE BY RELATIONSHIP BY EXPERTISE INTERACTION

	Expert	Novice
TRUTHFUL question		
With strangers	2.82	.89
With acquaintances	3.09	.71
REHEARSED question		
With strangers	4.71	3.89
With acquaintances	4.09	3.64
UNEXPECTED question		
With strangers	4.17	3.17
With acquaintances	4.45	2.36
REPEAT question		
With strangers	5.29	5.83
With acquaintances	3.27	5.07

Note. Higher scores represent greater discrepancy between interviewee reports and interviewer estimates of truth (i.e., they are *less* accurate).

Hypothesis 3 and Research Question 2: Familiarity

Hypothesis 3, that relational familiarity impairs accuracy, received partial support in the preceding H2 and RQ1 analyses showing that relational familiarity interacted with suspicion on honesty judgments and with expertise on accuracy. ERs overestimated the honesty of acquaintances, as hypothesized. Suspicious novice ERs especially underestimated honesty of truthful strangers and overestimated it for deceptive acquaintances. However, when they used unexpected questions, novice ERs were able to achieve greater accuracy in judging acquaintances than in judging strangers. Conversely, experts were best able to achieve accuracy in judging acquaintances when they repeated and probed an earlier question, but this strategy garnered them the worst accuracy with strangers. Thus, relational familiarity generally impaired accuracy but could be ameliorated by question type.

Research Question 2 addressed the role of behavioral familiarity in the form of expertise. Results yielded significant effects consistently. On honesty judgments, expertise produced a main effect, $F(1,59) = 7.60$, $p = .009$, $\eta^2 = .12$. Expert ERs judged both truthful and deceptive answers as less truthful than did novices. These estimates resulted in a decrement in accuracy. The accuracy analysis produced a main effect for expertise, $F(1,56) = 6.27$, $p = .020$, $\eta^2 = .12$, and an expertise by truth/deception interaction, $F(1,57) = 4.05$, $p = .049$, $\eta^2 = .08$. Specifically, experts were more inaccurate than novices judging truthful answers but not deceptive ones. In the bias analysis, there was likewise a main effect for expertise, $F(1,56) = 11.03$, $p = .003$, $\eta^2 = .18$. Experts were more biased (and therefore less accurate) when judging the truth (expert $M = -2.50$, novice $M = -.58$) but less biased (and therefore more accurate) when judging deception (expert $M = 1.65$, novice $M = 3.08$) (see Table 2).

Although the negative scores for experts on truth might imply a lie-bias, whereas the positive scores on deceptive answers might imply a truth-bias, a better way to characterize experts' behavior is as exemplifying greater skepticism but also lesser extremity in judgment. This becomes apparent if scores are calculated in a manner analogous to McBurney and Comadena's (1992) bias score. In their experiment, stimulus people told the truth on half the questions

and lied on half, so that respondents who chose truth over half the time would be considered as having a truth-bias. In our case, answers varied in degree of truthfulness, so the standard against which ER judgments needed to be assessed was EEs' average truthfulness across all questions. The weighted average score across the truthful and deceptive answers produced an average honesty score of 4.38 for EEs. By comparison, novice ERs gave an average rating of 6.81, and expert ERs gave an average rating of 5.46. Thus, novices had a more pronounced truth-bias than did experts, but both overestimated, on average, the truthfulness of EEs. The net result is that experts' inaccuracy interfered most with their judging the truth, but novices' bias interfered most with their judging deception.

Finally, it was noted previously that these inaccuracies could be partially offset by question strategy used, but only in the case of acquaintances. Novices using unexpected questions and experts using repeat questions improved their accuracy (see Table 2).

DISCUSSION

A guiding premise of interpersonal deception theory is that the act of communicating face-to-face alters deception relative to noninteractive contexts. The current results document several ways in which it does so. First, it exacerbates already-present truth-biases: People conversing face-to-face overattribute honesty to one another, especially with familiar others, which results in inaccuracy in detecting another's deception. The current results generally confirm this depiction of interpersonal deception. Interviewers were less accurate judging deception than they were judging truth, consistent with a truth-bias. This truth-bias held even for experts, despite the more "jaundiced" view that their experience or training had inculcated.

Second, relational and behavioral familiarity are highly salient influences on judgments during interpersonal deception. Consistent with the IDT proposition that familiarity would affect honesty attributions and accuracy, relational familiarity led to greater overattribution of honesty, while behavioral familiarity in the form of expertise led to decreased accuracy in judging truth.

Third, the active role of the receiver during interpersonal deception makes receiver suspicion another highly salient factor. IDT posits that suspicion introduces additional information-processing errors by altering attentional and attributional processes. Coupled with suspectors inadvertently leaking their suspicions to deceivers who use the feedback to craft more credible deceptions, the net effect should be reduced detection accuracy. Results confirmed that suspicion led experts and novices to see acquaintances as even more honest and strangers as even less honest, and suspicious experts were less accurate than nonsuspicious ones.

Fourth, IDT holds that specific sender and receiver communication strategies employed during interpersonal deception determine detection success. Results confirmed that senders' deception type and receivers' question type influenced receiver judgments. Equivocation as a deception strategy in itself engendered receiver suspicion, while concealment allayed it. Interviewers induced to be suspicious also became more accurate judging equivocation but less accurate judging concealment. Comparatively, then, equivocation was a less successful

deception strategy than concealment because it tended to make interviewers suspicious and was easier to detect. The most successful strategy of the three, however, was falsification, because interviewers were the least accurate in judging it. Interviewers in turn could improve their detection accuracy with acquaintances by using particular questioning strategies—novices, by using unexpected questions; experts, by using repeat questions. But when interviewing strangers, repeat questions produced the lowest accuracy for both novices and experts, possibly because strangers were able to repair earlier answers.

Combined, the results show that many features central to interpersonal interaction—expectations of truth, relational and behavioral familiarity, level of trust or suspicion, and the actual communication strategies of senders and receivers—determine honesty attributions and detection accuracy during deception. What emerges is a more complicated, dynamic picture of deception than obtained in previous noninteractive investigations. The results challenge the assumption that successful deception detection lies in identifying a stable profile of sender deception cues. Unmasking deception requires considering multiple sender and receiver factors and the dynamic interplay between the two.

The results also raise some important questions. First, why are falsifications the most difficult type of deception to detect? One possibility is that communicators are generally able to concoct plausible false answers, a speculation supported by the Burgoon, Buller, Guerrero, Afifi, and Feldman (in press) investigation, reported earlier, which found that falsifications were perceived as the most direct, clear, and complete. Another possibility is that interviewers' own truth-biases made falsifications most difficult to detect because falsifications are the most discrepant from the truthful pole of the truth-falsity continuum. Comparatively, concealments and equivocations were actually more truthful and more readily detected. To the extent that receivers lean toward judging all answers as truthful, they will de facto be more "accurate" judging answers that actually have a greater grain of truth in them, but the accuracy will depend on the answer's truth, not the interviewer's skill.

Second, why might accuracy differ depending on suspicion? It will be recalled that suspicion improved accuracy in judging equivocations but undermined it in judging concealments. One answer resides in the heightened vigilance and motivation to detect deceit that accompanied equivocation. Without the induced suspicion, such answers doubtless appeared plausible, though not very clear, and so looked much like normal, truthful discourse. But once suspicion was activated, ambiguous and evasive answers failed to withstand scrutiny, leading interviewers to judge them as less than completely truthful. By contrast, interviewers exposed to concealments reported less vigilance and motivation than other interviewers. They may therefore have been less accurate because they attended less carefully to the interviewees' answers. Alternatively, interviewees using concealment who were subjected to suspicion may have become more reticent. This would have made it more difficult for interviewers to detect that information was being withheld. Post hoc inspection of the means revealed that deceivers were in fact less honest in the suspicion than the nonsuspicion condition, a finding that would comport with this latter speculation. Thus, consistent with the premise of IDT, suspicion may have exerted its influence on the sender, the receiver, or both, by altering interviewer perceptions and

attentiveness and interviewee behavior management. This conclusion underscores the importance of studying factors such as suspicion in interactive contexts and studying their impact on both participants.

Third, why did induced suspicion exert a deleterious influence primarily on judging the honesty of acquaintances? One possibility is that the truth-bias associated with familiar others is highly resistant to doubt. When induced to be suspicious, interviewers may have been motivated to find the "truth" in their friends' answers and to overlook or rationalize away any questionable answers. With strangers, the ascription of honesty may have been less resilient and the heightened skepticism was sufficient to overcome it, producing greater accuracy. Another possibility is that other sources of suspicion more potent than the experimental suspicion manipulation were at work. As already noted, the different deception types intensified or diminished the level of suspicion. Additionally, experts entered the experiment with an already elevated level of suspicion relative to the novices. The manipulation checks revealed that experts were generally more surveillant and motivated to detect deception than the novice sample, possibly the result of generalized communication suspicion—"a predisposition toward believing that the messages produced by others are deceptive"—or state suspicion—"a belief that communication within a specific setting and at a particular time may be deceptive" (Levine & McCornack, 1991, p. 328). This may be commonplace among law enforcement, investigatory, and protective service agents. Given that experts were less accurate than novices, especially in judging the truth, their orientation toward suspiciousness may have produced their lower level of accuracy.

Overall, whether suspicion originates in a generalized personality predisposition to distrust others' communication, is induced by third parties or professional training, or arises naturally from situational cues or the act of lying itself, it transforms interaction processes and perceptions. By causing attentional shifts and interjecting information-processing biases, it can reduce detection accuracy. Of course, a lie-bias need not be a liability. There are circumstances where it is better to spot a liar than to doubt a truthteller. Experts responsible for lie detection may often find it safer to assume deceit than to trust questionable information. But if the objective is to distinguish liars from truthtellers, which is the task of jurors, employment interviewers, and police interrogators, among others, the current results and those of Levine and McCornack (1991) demonstrate that both truth– and lie–biases may seriously undermine one's ability to discern truth from falsehood. Although training and feedback may enable receivers to overcome predispositional biases, experience may not always be the best teacher if it leads to persistent suspiciousness, as occurred among our military experts. Whether different types of training and on-the-job experiences can reverse the truth-bias commonly found in interpersonal relationships without leading to a lie-bias deserves additional examination.

The influence of deception type and question type on perceptions and accuracy has implications for strategic behavior. Deceivers, to be successful, would be well-advised to engage in falsifications, which are most difficult to detect, or concealment, which allays suspicion, rather than the more common strategy of equivocation, which is likely to draw suspicion and be more readily detected. Conversely, deception detectors would be well-advised to use im-

promptu questions for which respondents have not planned answers or, if they are experts interacting with an acquaintance, to circle back and probe previously asked questions. However, this latter strategy is especially ill-advised with strangers, who may be able to repair previously unsatisfactory answers and embellish them with further, undetectable lies.[6]

Before closing, several methodological matters deserve comment. One concerns the origins of suspicion. In the current experiment, the manipulation checks revealed that suspicion arose not only through the experimental induction but also as an individual difference variable between novices and experts and as a byproduct of certain deception types. Because of this, the variables of suspicion, expertise, and deception type were confounded. This is a methodological problem that Crano (1992) has called the "madras effect." One experimental manipulation 'bleeds' into another—in this case, expertise and deception type into suspicion. Although this creates difficulties for disentangling which effects are due to suspicion per se, to one's chronic attitudes, or to the type of communication to which one is exposed, it probably reflects the reality of what is happening during deception. It is at least useful to realize that deception itself triggers varying levels of suspicion so that the role of suspicion can be factored into future predictions and explanations about deception effects.[7] Nevertheless, separating the sources of suspicion poses a challenge for future research.

Another concern is the overall strength of the suspicion manipulation. One conclusion might be that the suspicion manipulation failed. However, interviewers reported the expected differences in their expectation that the interviewee would tell the truth. Besides the problems created by experts being more suspicious at the outset of conversations than novices, the small differences in the manipulation checks also may be a measurement inaccuracy. Buller, Strzyzewski, and Comstock (1991) found that suspicion aroused by an experimental induction did not remain constant throughout their experimental interactions. Thus a manipulation check taken after the entire interaction may not reflect the strength of interviewers' suspicions at the beginning of the interactions, because the interviewees' performance has likely affected it.

On a more positive note, the measurement approach taken here makes a useful contribution by demonstrating the improvement in precision and understanding that can be gained when interval-level and multiple estimates of deception are used in place of dichotomous and single measures and when reports are obtained from both sender and receiver. Use of multiple, graded estimates takes into account the possibility that senders' messages may vary in degree of truthfulness as well as degree of directness, clarity, and relevance so that answers may be partly but not completely honest or deceptive. Including estimates from both parties permits teasing out not only the amount of discrepancy between a sender's messages and a receiver's estimate of their truthfulness but also the direction of any biases. Given the more complete information that this approach offers, we recommend its adoption in future investigations.

A third methodological note concerns the dynamics of order effects. The fact that all interviews began with honest answers may have created a cognitive bias to expect honest answers. This would naturally affect accuracy scores by reinforcing a truth-bias. The current research therefore needs replication under conditions of different orderings. For example, we are currently alternating blocks of

truthful and deceptive answers to better understand suspicion effects and to rule out the rival hypothesis that starting interactions with "known" honest answers solidifies expectations of honesty.

Finally, the acquaintance category in the relational familiarity variable was relatively broad, essentially including non-strangers. Past research by Buller and Aune (1987) suggested that acquaintances and friends did not differ from one another but did differ from strangers in their behavior during deception, leading us to utilize only two relational familiarity categories. Our results further reinforce the conclusion that strangers and acquaintances differ in their interpretation of communicative behavior, although they admittedly overlook subtle differences that may be present between very well- and less well-acquainted dyads.

This investigation is but one foray into the nature of interpersonal deception. Continued investigations that wed interpersonal communication principles with deception should contribute to our understanding of interpersonal credibility, attributions processes, and accuracy, as well as deception.

ENDNOTES

[1]It should be noted that Stiff, Kim, and Ramesh (1992) take a somewhat different view, regarding truth-bias as a cognitive heuristic that emerges as relationships develop.

[2]Because the first few participants mistook their opening introduction as a question and began deceiving on the third rather than the fourth question, we omitted question three from calculations of truthful accuracy scores. Instructions were subsequently made more explicit as to which question began the deception phase.

[3]Following are the questions used during the interview. The unexpected and repeat questions are noted.

1. Please describe your educational background.
2. What event from your childhood do you remember most fondly?
3. If you found a wallet containing $1,000 and no identification, what would you do with it? Why?
4. What types of people tend to rub you the "wrong way?"
5. How would you describe your religion and religious beliefs?
6. Unexpected question: How do you think I could improve my physical appearance?
7. If your best friend was cheating on his or her spouse, what would you do? Why?
8. Repeat question: How does your answer to the last question (Question 7) fit with your religious beliefs that you described earlier?
9. Do you thoroughly investigate the qualifications of all the candidates before you vote?
10. Describe the work you do or did.
11. Did you support sending troops to the Persian Gulf to liberate Kuwait?
12. What's the most embarrassing thing you've ever done?
13. What do you consider to be your biggest strengths and weaknesses?
14. Have you ever taken revenge on someone? If so, how?
15. If you won a million dollars in the lottery, how much would you give to charity? Why?

[4]The scale for measuring suspicion consisted of the following statements: (a) My partner used language that made it hard to understand what he/she meant, (b) I had a feeling that something was wrong with my partner's answer, and (c) I expected my partner to tell the truth. The scale for measuring vigilance consisted of the following statements: (a) I didn't pay any special attention to my partner's behavior, (b) I listened more carefully to my partner's answers than I would in normal conversation, and (c) I carefully watched my partner's behavior. The scale for measuring motivation to detect deceit consisted of the following statements: (a) I didn't care whether I understood my partner's answers, (b) I didn't try very hard to find out if my partner was honest, and (c) I was very motivated to determine if my partner was telling me the truth.

[5]A deception by suspicion interaction among experts also duplicated the interaction cited under Hypothesis 1.

[6]There was some concern that the unexpected question in this study was potentially confounded by the fact that it asked the receiver to evaluate the interviewer (i.e., on physical appearance) rather than discuss something about the receiver as other questions required. Such direct evaluation of the interviewer may have been more anxiety-producing than the other questions, making it easier to recognize. However, the repeat question, which was also to some extent unexpected because it was not on

the original list supplied to the receiver, had similar effects on accuracy, suggesting that it is easier to judge the honesty of impromptu answers.

[7]One other possibility that should also be entertained for the weak induced suspicion effects is the timing of the manipulation check—after the interaction. We have found previously that postinteraction measures of suspicion can be problematic because they can be confounded by the interaction itself and the other experimentally manipulated variables. However, such manipulation checks cannot be conducted prior to the interaction without sensitizing participants to the manipulation. The only alternative is to conduct independent manipulation checks. One conducted previously on a highly similar version of this manipulation did confirm that the suspicion induction increased expectations of deceit and motivation to detect it (see Burgoon, Buller, Dillman, & Walther, 1994). Research currently in progress is attempting to further assess the nature of suspicion and what is necessary to provoke moderate or high levels of it.

REFERENCES

Bauchner, J.E., Brandt, D.R., & Miller, G.R. (1977). The truth-deception attribution: Effects of varying levels of information availability. In B.R. Ruben (Ed.), *Communication yearbook 1* (pp. 229–243). New Brunswick, NJ: Transaction Books.

Bauchner, J.E., Kaplan, E.A., & Miller, G.R. (1980). Detecting deception: The relationship of available information to judgmental accuracy in initial encounter. *Human Communication Research, 6*, 253–264.

Bavelas, J.B., Black, A., Chovil, N., & Mullett, J. (1990). *Equivocal communication*. Newbury Park, CA: Sage.

Brandt, D.R., Miller, G.R., & Hocking, J.E. (1980). The truth-deception attribution: Effects of familiarity on the ability of observers to detect deception. *Human Communication Research, 6*, 99–110.

Brandt, D.R., Miller, G.R., & Hocking, J.E. (1982). Familiarity and lie detection: A replication and extension. *Western Journal of Speech Communication, 46*, 276–290.

Buller, D.B. (1987, November). *Detecting deception by strangers, friends and intimates*. Paper presented at the annual meeting of the Speech Communication Association, Boston.

Buller, D.B., & Aune, R.K. (1987). Nonverbal cues to deception among intimates, friends, and strangers. *Journal of Nonverbal Behavior, 11*, 269–290.

Buller, D.B., & Burgoon, J.K. (1991, August). *The language of interpersonal deception: Falsification, equivocation, and concealment*. Paper presented at the 4th Annual Conference on Language and Social Psychology, Santa Barbara, CA.

Buller, D.B., & Burgoon, J.K. (1994). Deception: Strategic and nonstrategic communication. In J.A. Daly & J.M. Wiemann (Eds.), *Strategic interpersonal communication* (pp. 191–223). Hillsdale, NJ: Erlbaum.

Buller, D.B., & Burgoon, J.K. (in press). Interpersonal deception theory. *Communication Theory*.

Buller, D.B., Comstock, J., Aune, R.K., & Strzyzewski, K.D. (1989). The effect of probing on deceivers and truthtellers. *Journal of Nonverbal Behavior, 13*, 155–170.

Buller, D.B., & Hunsaker, F.G. (in press). Interpersonal deception: XIII. Suspicion and the truth-bias of conversational participants. In J. Aitken (Ed.), *Intrapersonal communication processes reader*. Westland, MI: Hayden-McNeil.

Buller, D.B., Strzyzewski, K.D., & Comstock, J. (1991). Interpersonal deception: I. Deceivers' reactions to receivers' suspicions and probing. *Communication Monographs, 58*, 1–24.

Buller, D.B., Strzyzewski, K.D., & Hunsaker, F.G. (1991). Interpersonal deception: II. The inferiority of conversational participants as deception detectors. *Communication Monographs, 58*, 25–40.

Burgoon, J.K. (1992, November). *Applying an interpersonal perspective to deception: Effects of suspicion, deceit, and relational familiarity on perceived communication*. Paper presented at the annual meeting of the Speech Communication Association, Chicago.

Burgoon, J.K., & Buller, D.B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior, 18*, 155–184.

Burgoon, J.K., Buller, D.B., Dillman, L., & Walther, J. (1994). *Interpersonal deception: IV. Effects of suspicion on perceived communication and nonverbal behavior dynamics*. Manuscript submitted for publication.

Burgoon, J.K., Buller, D.B., Guerrero, L.K., Afifi, W., & Feldman, C.M. (in press). Interpersonal deception: XII. Information management dimensions underlying types of deceptive messages. *Communication Theory*.

Burgoon, J.K., Buller, D.B., Guerrero, L.K., & Feldman, C.M. (1993, May). *Viewing deception success from deceiver and observer perspectives: Effects of preinteractional and interactional factors*. Paper presented at the annual meeting of the International Communication Association, Washington, DC.

Comadena, M.E. (1982). Accuracy in detecting deception: Intimate and friendship relationships. In M. Burgoon (Ed.), *Communication yearbook 6* (pp. 446–472). Beverly Hills, CA: Sage.

Crano, W.B. (1992, February). *The "madras effect."* Colloquium presented to the University of Arizona, Tucson.

DePaulo, B.M., Blank, A.L., Swaim, G.W., & Hairfield, J.G. (1992). Expressiveness and expressive control. *Personality and Social Psychology Bulletin, 18*, 276–285.

DePaulo, B.M., LeMay, C.S., & Epstein, J.A. (1991). Effects of importance of success and expectations for success on effectiveness at deceiving. *Personality and Social Psychology Bulletin, 17,* 14–24.

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.

deTurck, M.A., & Goldhaber, G.M. (1988). Perjury and deceptive judgments: How the timing and modality of witness deception affects jurors' deceptive judgments. *Communication Quarterly, 36,* 276–289.

deTurck, M.A., Harszlak, J.J., Bodhorn, D.J., & Texter, L.A. (1990). The effects of training social perceivers to detect deception from behavioral cues. *Communication Quarterly, 38,* 189–199.

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., & Friesen, W.V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32,* 88–106.

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

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.R. (1991). Invited article: Face, voice, and body in detecting deceit. *Journal of Nonverbal Behavior, 15,* 125–135.

Fay, P.J., & Middleton, W. (1941). The ability to judge truthtelling or lying from the voice as transmitted over a public address system. *Journal of General Psychology, 24,* 211–215.

Gray, P. (1992, October 5). Lies, lies, lies. *Time,* pp. 32–38.

Grice, P. (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Hopper, R., & Bell, R.A. (1984). Broadening the deception construct. *Quarterly Journal of Speech, 70,* 288–302.

Kalbfleisch, P.J. (1985). Accuracy in deception detection: A quantitative review. (Doctoral dissertation, Michigan State University, 1985). *Dissertation Abstracts International, 45,* 46112B.

Kalbfleisch, P.J. (1992). Deceit, distrust and the social milieu: Application of deception research in a troubled world. *Journal of Applied Communication Research, 20,* 308–334.

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

Knapp, M.L., Cody, M.J., & Reardon, K.K. (1987). Nonverbal signals. In C.R. Berger & S.H. Chaffee (Eds.), *Handbook of communication science* (pp. 385–418). Newbury Park, CA: Sage.

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. (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.

Levine, T.R., & McCornack, S.A. (1989, November). *Distinguishing between types of suspicion and measuring a predisposition toward suspicion: Two studies validating a measure of trait suspicion.* Paper presented at the annual meeting of the Speech Communication Association, San Francisco.

Levine, T.R., & McCornack, S.A. (1991). The dark side of trust: Conceptualizing and measuring types of communicative suspicion. *Communication Quarterly, 39,* 325–339.

Littlepage, G.E., & Pineault, M.A. (1979). Detection of deceptive factual statements from the body and the face. *Personality and Social Psychology Bulletin, 5,* 325–328.

Maier, N., & Thurber, J. (1968). Accuracy of judgments of deception when an interview is watched, heard, and read. *Personnel Psychology, 21,* 23–30.

Manstead, A.S.R., Wagner, H.L., & MacDonald, C.J. (1986). Deceptive and nondeceptive communications: Sending experience, modality, and individual abilities. *Journal of Nonverbal Behavior, 10,* 147–167.

McBurney, D.L., & Comadena, M.E. (1992). *An examination of the McCornack and Parks model of deception detection.* Paper presented at the annual meeting of the Speech Communication Association, Chicago.

McCornack, S.A. (1992). Information manipulation theory. *Communication Monographs, 59,* 1–16.

McCornack, S.A., & Levine, T.R. (1990). When lovers become leery: The relationship between suspicion and accuracy in detecting deception. *Communication Monographs, 57,* 219–230.

McCornack, S.A., & Parks, M.R. (1986). Deception detection and relationship development: The other side of trust. In M.L. McLaughlin (Ed.), *Communication yearbook 9*(pp. 377–389). Beverly Hills, CA: Sage.

McCornack, S.A., & Parks, M.R. (1990). What women know that men don't: Sex differences in determining the truth behind deceptive messages. *Journal of Social and Personal Relationships, 7,* 107–118.

Metts, S. (1989). An exploratory investigation of deception in close relationships. *Journal of Social and Personal Relationships, 6,* 159–179.

Metts, S., & Chronis, H. (1986, May). *An exploratory investigation of relational deception*. Paper presented at the annual meeting of the International Communication Association, Chicago.

Miller, G.R., deTurck, M.A., & Kalbfleisch, P.J. (1983). Self-monitoring, rehearsal, and deceptive communication. *Human Communication Research, 10*, 97–117.

Miller, G.R., & Stiff, J.B. (1993). *Deceptive communication*. Newbury Park, CA: Sage.

Mongeau, P.A. (1988). *Would I lie to you honey? If I did would you be able to tell? Deception detection accuracy in ongoing relationships*. Paper presented at the annual meeting of the Western Speech Communication Association, San Diego.

O'Hair, H.D., Cody, M.J., & McLaughlin, M.L. (1981). Prepared lies, spontaneous lies, Machiavellianism and nonverbal communication. *Human Communication Research, 7*, 325–339.

O'Hair, H.D., Cody, M.J., Wang, X., & Chao, E.Y. (1990). Vocal stress and deception detection among Chinese. *Communication Quarterly, 38*, 158–169.

O'Sullivan, M., Ekman, P., & Friesen, W.V. (1988). The effect of comparisons on detecting deceit. *Journal of Nonverbal Behavior, 12*, 203–215.

Riggio, R.E., Tucker, J., & Throckmorton, D. (1987). Social skills and deception ability. *Personality and Social Psychology Bulletin, 13*, 568–577.

Silverman, I.J., Rivera, A.N., & Tedeschi, J.T. (1979). Transgression-compliance: Guilt, negative affect, or impression management? *Journal of Social Psychology, 108*, 57–62.

Stiff, J.B., Kim, H.J., & Ramesh, C.N. (1992). Truth-bias and aroused suspicion in relational deception. *Communication Research, 19*, 326–345.

Stiff, J.B., & Miller, G.R. (1986). "Come to think of it. . . ." Interrogative probes, deceptive communication and deception detection. *Human Communication Research, 12*, 339–357.

Toris, C., & DePaulo, B.M. (1985). Effects of actual deception and suspiciousness of deception on interpersonal perceptions. *Journal of Personality and Social Psychology, 47*, 1063–1073.

Turner, R.E., Edgley, C., & Olmstead, G. (1975). Information control in conversations: Honesty is not always the best policy. *Kansas Journal of Speech, 11*, 69–89.

Zebrowitz, L.A. (1990). *Social perception*. Pacific Grove, CA: Brooks/Cole.

Zuckerman, M., & Driver, R.E. (1985). Telling lies: Verbal and nonverbal correlates of deception. In A.W. Siegman & S. Feldstein (Eds.), *Multichannel integrations of nonverbal behavior* (pp. 129–148). Hillsdale, NJ: Erlbaum.

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

Zuckerman, M., Spiegel, N.H., DePaulo, B.M., & Rosenthal, R. (1982). Nonverbal strategies for decoding deception. *Journal of Nonverbal Behavior, 6*, 171–187.