Psychological Bulletin
2008, Vol. 134, No. 4, 477–492

0033-2909/08/$12.00 DOI: 10.1037/0033-2909.134.4.477

Individual Differences in Judging Deception: Accuracy and Bias

Charles F. Bond Jr.
Texas Christian University

Bella M. DePaulo
University of California at Santa Barbara

The authors report a meta-analysis of individual differences in detecting deception, confining attention to occasions when people judge strangers' veracity in real-time with no special aids. The authors have developed a statistical technique to correct nominal individual differences for differences introduced by random measurement error. Although researchers have suggested that people differ in the ability to detect lies, psychometric analyses of 247 samples reveal that these ability differences are minute. In terms of the percentage of lies detected, measurement-corrected standard deviations in judge ability are less than 1%. In accuracy, judges range no more widely than would be expected by chance, and the best judges are no more accurate than a stochastic mechanism would produce. When judging deception, people differ less in ability than in the inclination to regard others' statements as truthful. People also differ from one another as lie- and truth-tellers. They vary in the detectability of their lies. Moreover, some people are more credible than others whether lying or truth-telling. Results reveal that the outcome of a deception judgment depends more on the liar's credibility than any other individual difference.

Keywords: deception, lie detection, individual differences, social judgments

Supplemental materials: http://dx.doi.org/10.1037/0033-2909.134.4.477.supp

It has been widely believed (indeed, "virtually axiomatic"; Hubbell, Mitchell, & Gee, 2001, p. 115) that people are not very accurate at detecting deception. This is the consensus among psychologists who arrange for people to judge lies and truths and to assess the percentage of those lies and truths they correctly detect. In a large research literature, overall rates of lie/truth discrimination average less than 55%, when 50% would be expected by chance (Aamodt & Mitchell, 2006). Moreover accuracy rates vary little across studies (C. F. Bond & DePaulo, 2006).

In fact, though, what is known concerns the *average* lie detection abilities of groups of people. Although much has been said about the *mean* accuracy of these groups, there has been less evidence on individuals' abilities.

Researchers have assumed that people vary in the ability to detect lies. Buller and Burgoon (1996) posited that lie detection accuracy depends, in part, on the receiver's decoding skill; Malone and DePaulo (2001) discussed individual differences in sensitivity to deception; and O'Sullivan (2007) based some recent work on the assumption that "lie detection is an ability that can be measured" (p. 118). Presupposing that this ability exists, researchers have attempted to discover the characteristics of people who have unusual lie detection skills. No such characteristics have been uncovered. Of course, in this literature of over 200 articles, one can find a study (or two) in which a given characteristic covaries with lie detection performance. Yet across the research literature as a whole, no individual difference has been found that is consistently correlated with the detection of deceit.

In an early meta-analysis, Zuckerman, DePaulo, and Rosenthal (1981) found no reliable effects of a judge's Machiavellianism, self-monitoring, or sex on the judge's accuracy at detecting lies. In a recent quantitative review, Aamodt and Mitchell (2006) sought to relate individual differences in lie/truth discrimination accuracy to a large number of variables—including the judge's age, education, expertise, confidence, and sex. Aamodt and Mitchell found no variables that were significantly related to lie detection.

It has been supposed that people differ from one another in lie detection ability. However, this supposition may be false. Under an alternative hypothesis, differences observed in detection performances reflect nothing more than chance variation. Kraut (1980) argued that people vary little in their lie detection skills. Consistent with this view, C. F. Bond, Kahler, and Paolicelli (1985) reported negligible differences from judge-to-judge in lie/truth discrimination performances; Kraut (1978) found no relationship between a perceiver's accuracy in judging one person and that same perceiver's accuracy in judging a second person; DePaulo and Rosenthal (1979) reported that a person's success at spotting men's lies is independent of his (or her) success at spotting women's lies; and Levine, Park, and McCornack (1999) found no positive relationship between a person's accuracy in identifying lies and that same person's accuracy in identifying truths. Unfortunately, only a handful of investigators have studied individual differences in lie detection ability, and there has never been a large-scale analysis of ability differences.

Although it may seem obvious that people differ in the ability to detect lies, psychometric theory encourages us to entertain the possibility that they do not. The theory provides a framework for understanding variation in test performances. Here, each test consists of a series of truthful and deceptive statements. To each

Charles F. Bond Jr., Department of Psychology, Texas Christian University; Bella M. DePaulo, Department of Psychology, University of California at Santa Barbara.

We are grateful to David Kenny and Chris Wetzel for comments on a draft of this article.

Correspondence concerning this article should be addressed to Charles F. Bond Jr., 4128 Pebblebrook Court, Fort Worth, TX 76109. E-mail: cbond53@yahoo.com

statement, each judge gives a dichotomous response by indicating whether the statement is a lie or the truth. Each response is either correct or incorrect, and our interest is in the variation across judges in the percentage of correct responses. Under traditional psychometric theory (Lord & Novick, 1968), this variation includes two components: real variance and error variance. Real variance consists of the variance across test-takers that would be observed in a hypothetical test much longer than the test at hand—consisting of all possible test items from a universe. Here the putative universe would include all truthful and deceptive statements analogous to the ones being judged. Under psychometric theory, the *observed* variance across judges in percentage correct is larger than the *real* variance because judges see only a sample of statements. Observed variance is artifactually inflated to a degree that depends on the brevity of the test—with the briefest tests showing the highest variance.

Theoretically, it is possible for all of the variance in a set of test performances to be error variance. In such cases, we would say that the test was completely unreliable or (equivalently) that performances do not generalize over test items. For explanations of test theory, see Nunnally and Bernstein (1994); Ghiselli, Campbell, and Zedeck (1981); or Lord and Novick (1968).

Here, we offer the first large-scale analysis of individual differences in detecting deception. There has been related work. In a discussion of various measures of interpersonal sensitivity, J. A. Hall (2001) concluded that tests of accuracy for judgments of emotion show reasonable levels of internal consistency, whereas those that tap judgment accuracy across a variety of interpersonal domains do not. As part of a large-scale meta-analysis of mean lie detection accuracy, C. F. Bond and DePaulo (2006) noted that standard deviations in accuracy seem to be small. However, the meta-analysts did not test for the possibility that individual differences in accuracy are artifactual, nor did they try to estimate the magnitude of ability differences.

Individual Differences in Judging Deception

Our goal is to determine the magnitude of individual differences in deception judgments. We are primarily interested in differences across judges in percentage of correct lie/truth discrimination. With a statistical technique, we correct apparent judge-to-judge differences in accuracy for the differences that would be expected by chance. If psychometric theory is correct, chance variation should be greatest among individuals who judge only a small number of lies and truths. Perhaps our analyses will reveal that there are no real differences in the ability to detect lies. If so, this would explain why meta-analysts have found no individual- difference characteristics that are consistently related to lie detection performances. In the absence of differences in lie detection ability, it is unlikely that any nonchance correlates of lie detection performances will ever be found.

Although psychometric techniques can be used to assess abilities, they can also be used for a second purpose—to analyze test-taking biases. On personality inventories, for example, respondents differ in their inclination to agree with a statement, regardless of the statement's content (Paulhus, 1991). In a similar way, people who are judging the veracity of a series of statements might vary in their tendency to label statements as *truths*. While probing for differences among judges in accuracy at discriminating lies from truths, we also test for judge-to-judge differences in *credulity*—that is, in the general predisposition to regard others' statements as truthful.

Deception judgments have consequences (Granhag & Strömwall, 2004). Often the consequences do not depend on whether a deception judgment is correct. Some murder suspects are freed, and others are sentenced to die because of jurors' *judgments* of the suspects' truthfulness. Some international negotiations succeed, and others fail because of the negotiators' *judgments* of one another's honesty (C. F. Bond et al., 1992). Thus, it is important to understand any biases people may have toward viewing others as deceptive (or truthful).

Many have noted that the average judge labels more than 50% of the statements s/he hears as truthful, when 50% are in fact truthful (Levine et al., 1999). However, less has been said about individual differences in this tendency. People vary in chronic levels of trust (Levine & McCornack, 1991). Perhaps this influences deception judgments, such that some people are more likely than others to accept statements as truthful. If so, our psychometric analyses should reveal the magnitude of this individual difference—disentangling real variance in judge credulity from artifactual variance.

Deception involves two people—the liar as well as the judge. While gauging differences among individuals as judges of deceit, we also assess differences among them as liars. In the current article, we refer to people who lie as *senders*—recognizing that they do not invariably lie. Often, they tell the truth.

C. F. Bond et al. (1985) maintained that the outcome of a deception judgment depends more on the liar than the lie detector. Consistent with this view, the researchers found larger individual differences among people as lie- and truth-tellers than as judges. Kraut (1980) found that people who are judged to be honest by one person are judged to be honest by another. C. F. Bond and Atoum (2000) discovered that groups of people who are making judgments independently of one another reach consensus. They agree that certain individuals are lying and that others are telling the truth. People reach this consensus even when they are wrong. Judges seem to base their inferences about a person's truthfulness on the person's demeanor, in that the people who appear most honest when lying are the people who appear most honest when telling the truth (DePaulo & Rosenthal, 1979).

Although these earlier studies suggest that individuals differ as lie- and truth-tellers, no large-scale analysis of this putative difference has been reported. Here, we assess differences among senders along two dimensions: detectability and credibility. By our definition, a sender is perfectly *detectable* if that sender is always judged to be lying when s/he is telling a lie and always judged to be telling the truth when s/he is telling the truth. A sender is perfectly *credible* if that sender is invariably perceived to be truthful—whether lying or telling the truth. Perhaps our analysis of sender detectability will reveal that the veracity of some individuals is obvious and that the veracity of others is inscrutable. Maybe our analysis of sender credibility will indicate that some people are invariably believed and others invariably disbelieved, whether lying or telling the truth.

Before drawing any conclusions about differences among people who are lying and telling the truth, it will again be important to distinguish real individual differences from artifactual differ-

ences. When telling lies and truths, senders are judged for their veracity, and random error is introduced because the judgments of each sender are only a sample of all possible judgments that might have been made. Under psychometric theory, the smaller the number of times each sender is judged, the greater should be the artifactual variance in senders' indetectability and credibility. We report analyses designed to separate artifactual differences in these variables from real differences. If any real sender differences are found, their magnitude will be of interest, and we will want to compare real differences among senders with real differences among judges.

In sum, our psychometric investigation may reveal that all variation in deception judgments is illusory and reflects nothing more than chance. If there is nonrandom variation in deception judgments, our analysis should allow a partitioning of that variation between liars and lie detectors. It should reveal the extent to which a deception judgment depends on the judge's detection ability, the judge's credulity, the sender's detectability, and the sender's credibility.

Our database consists of all relevant studies of deception detection we could find. An analysis of hundreds of studies will let us draw generalizations that a handful of studies would not permit.

This meta-analytic investigation will also allow us to examine conditions that may moderate the size of individual differences in judging deception. In some of the studies in this research literature, liars have no particular motivation to succeed; in others, they are motivated. Perhaps judge-to-judge differences in lie detection ability are large when liars have motivation to succeed and small when liars lack motivation. With statistical analyses, we assess this and a number of related possibilities.

Having gauged individual differences in detection ability, credulity, detectability, and credibility in the deception judgment research literature as a whole, we assess the magnitude of those differences in various subsets of this literature—for instance, in studies in which lies are significantly discriminated from truths and in studies in which they are not.

Statistically, we measure individual differences in judging deception with standard deviations. We analyze standard deviations across judges, as well as standard deviations across senders. In supplementary analyses, we examine the range in these variables. O'Sullivan (2007) has claimed that a few people have extraordinary lie detection ability. We assess this hypothesis by comparing the top lie detection performances reported in the deception detection research literature with the top performances that would be expected by chance. More generally, we supplement our analyses of standard deviations in judge ability, judge credulity, sender detectability, and sender credibility with analyses of a second measure of the dispersion in each individual difference variable: the range.

Method

Literature Search Procedures

We used standard methods to locate relevant research. We conducted computer-based searches of Psychological Abstracts, PsycInfo, PsycLit, Communication Abstracts, Dissertation Abstracts International, WorldCat, and Yahoo through December 2006, using the phrases "deception judgment," "deception detection," and "lie detection." We searched the Social Sciences Citation Index for articles that cited key references, examined reference lists from previous reviews, and reviewed the references cited in every document that we found. Through e-mail, we requested articles from over 25 scholars who had published relevant articles.

Criteria for Inclusion and Exclusion of Studies

Our goal was to summarize all relevant English-language reports of original research on the accuracy of judgments of lies and truths available to us prior to January 2007. To be included in this review, a document had to report (or allow computation of) a measure of individual differences in accuracy at discriminating lies from truths. The measure had to be based on dichotomous judgments of statements as lies or truths by individuals who had made more than one such judgment.

Given our interest in lie/truth discrimination, we excluded studies in which individuals judged only lies and studies in which individuals judged only truths. To avoid the influence of varying degrees and types of acquaintanceship on lie detection, we excluded investigations in which judges and senders knew one another prior to participating in the study. Hoping to understand deception judgments as they are made in everyday life, we also excluded studies in which judges could draw on aids to lie detection (e.g., polygraph records or behavior codings). We excluded judgments made by people who were less than 17 years of age, leaving to child psychologists questions about the early acquisition of lie detection skills. To confine attention to statements that could properly be considered lies, we excluded reports in which senders role-played an imagined person in an imagined situation. We excluded studies in which deception judgments were made on multipoint rating scales because we could not determine from most rating-scale results a quantitative index of individual differences that would be expected by chance. We excluded studies in which each judge made only a single lie/truth judgment. In such studies, the variability among judges is completely determined by their mean judgment.

Defining Samples

Research studies in this literature exhibit two forms of interdependence: sender interdependence and judge interdependence. Senders are interdependent when the lies and truths told by a given sample of people are shown to multiple samples of judges. Judges are interdependent when researchers report multiple measures of lie/truth accuracy for a given sample of judges.

Below, we report analyses of judge differences and analyses of sender differences. For our overall analyses of judge differences, we extract one standard deviation from each independent sample of judges. For our overall analyses of sender differences, we extract one estimate from each independent sample of lie- and truth-tellers. For analyzing the moderation of individual differences, we extract multiple estimates from a sample if those esti-

mates reflect different levels of the moderator variable that we are analyzing.[1]

Coding Individual Differences

We coded individual differences in judge ability, judge credulity, sender detectability, and sender credibility. To understand these variables, readers may find it helpful to think of the results of a deception detection study as a rectangular matrix of 0 s and 1 s (see Figure 1). Each row of the matrix represents the responses on a lie/truth discrimination test given by a particular judge. Each column represents the judgments made in response to a particular sender. A judge's response to a sender is dichotomous and can be scored for either of two characteristics: accuracy or bias. A 1 in the cell (*i, j*) of the *accuracy matrix* implies that judge *i* correctly detected sender *j*'s lie or truth. A 0 in the cell (*i*, *j*) implies that judge *i* was incorrect in assessing sender *j*'s truthfulness. In this matrix, judge *i*'s accuracy in discriminating lies from truths is reflected in the marginal mean of row *i*, whereas the detectability of sender *j*'s lies and truths is reflected in the marginal mean of column *j*.

From each document that reported it, we coded the standard deviation across judges in the percentage of correct lie-or-truth judgments—that is, the standard deviation across the row marginal means of the 0–1 accuracy matrix, multiplied by 100. Whenever possible, we also coded the standard deviation across senders in the percentage of times each sender's truthfulness was judged correctly. This is the standard deviation across the column marginal means of the accuracy matrix, multiplied by 100.

Each judgment in the deception detection literature can be scored not only for accuracy but also for bias. A 1 appears in the cell (*i, j*) of the *bias matrix* if judge *i* believed that sender *j* was telling the truth; a 0 appears if judge *i* believed that sender *j* was lying. Judge *i*'s tendency to regard statements as truthful is reflected in the marginal mean for row *i* of the bias matrix, whereas sender *j*'s tendency to be believed is reflected in the marginal mean for column *j*.

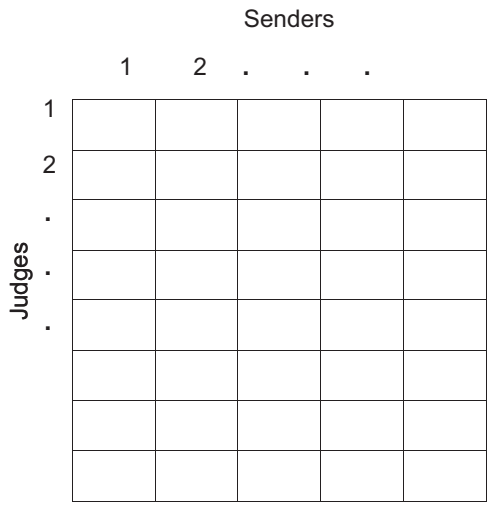

Figure 1. Results of a deception detection study.

From each document that reported it, we coded the standard deviation across judges in the percentage of messages judged to be truths—that is, the standard deviation across the row marginal means for the bias matrix, multiplied by 100. Whenever possible, we also coded the standard deviation across senders in the percentage of times each sender was judged to be telling the truth. This is the standard deviation across the column marginal means of the bias matrix, multiplied by 100.

We also sought to code the extremes on our four individual difference variables. When possible, we coded the maximum percentage of messages that any judge detected, as well as the minimum that any judge detected. We also coded the maximum and minimum for the following: percentage of messages any judge believed to be true, percentage of accuracy in judgments made of any sender, and percentage of times any sender was judged to be telling the truth.

The *number of judges* and *number of senders* were coded from each document. So was the number of lie/truth judgments made by each judge, as well as the number of lie/truth judgments made of each sender. We coded the mean percentage of correct lie/truth judgments and the mean percentage of truth judgments when available.

Other Variables

The other variables of interest to us are categorical. People perpetrate lies over various media. Here we coded *deception medium* by noting whether receivers were trying to detect lies over a video medium, an audio medium, or an audiovisual medium. We coded *liar motivation* by noting whether participants had any special motivation to dupe others. We coded *lie preparation* by noting whether participants had any time to prepare their remarks.

In some studies, senders are interacting with others as they lie and tell the truth; in other studies, they are not. For purposes of coding *liar interaction*, we regarded senders as not interacting if when lying they were alone or in the presence of a passive observer. We deemed them to be interacting if they lied in the presence of a person who was not passive. Most of the judges in this literature are college students. Others are people whose occupations are thought to provide them with special experience in detecting lies, such as mental health professionals or law enforcement agents. We noted this variable of *judge experience*. In many studies, research participants must judge whether a sender is lying or telling the truth without any baseline information about how that sender acts when being truthful; in other studies, judges get a baseline exposure of each sender's truthful behavior. We noted this variable of *baseline exposure*. We also distinguished standard deviations reported in unpublished documents from those reported in published documents. Finally, we noted whether (in each study) mean lie/truth discrimination significantly exceeded 50% by a one-tailed test with $p < .05$.

[1] In defining samples, we tried to separate individual differences from experimental effects. Thus, from an experimental study of the impact of training on lie detection (Vrij, 2000, pp. 93–97), we would extract two standard deviations: one among the judges who received training and a second among the judges who did not. For analyses of the impact on deception judgments of experimental factors, see C. F. Bond and DePaulo (2006).

For their meta-analysis of mean lie detection accuracy, C. F. Bond and DePaulo (2006) applied the current coding scheme to 123 of the 142 documents in the current compilation. They established the interrater reliability of the codes we are using here. For details, see C. F. Bond and DePaulo (2006, p. 219). As noted there, the two authors of the earlier article reached a mean of 92.2% agreement coding deception medium, liar motivation, lie preparation, liar interaction, judge experience, and baseline exposure in 24 documents randomly chosen from C. F. Bond and DePaulo's corpus. Nineteen of the documents in the current compilation did not appear in the earlier meta-analysis. C. F. Bond coded these latter documents.

Results

Our literature search uncovered 142 relevant documents. Of the documents, 89 were published, and 53 were unpublished. These documents, designated by asterisks in our References list, chronicle the efforts of 19,801 judges to assess the veracity of 2,945 senders. For a listing of study-by-study results, see Appendix A, which is available online as supplemental material. Averaging across all of the studies in this database, judges achieved a mean accuracy of 54.05% in discriminating lies from truths while rendering a mean of 55.50% truth judgments. These results are consistent with those reported in our earlier meta-analysis (C. F. Bond & DePaulo, 2006).

We analyzed individual differences in judge ability, judge credulity, sender detectability, and sender credibility. We measured the dispersion in each of these individual differences with two statistics: the standard deviation and the range.

Judge Ability

Standard deviation. We evaluated differences from judge to judge in the ability to discriminate lies from truths, measuring these differences with standard deviations. We chose to work with standard deviations, rather than variances, because standard deviations are in the familiar metric of percentage correct, whereas variances would assess ability differences as percentages squared (Howell, 2006). For comparison, it may be useful to note that the maximum possible standard deviation in a distribution of percentages is 50. This would occur if half of the percentages were equal to 0, and the other half were equal to 100. More generally, it is easy to determine the standard deviation in a distribution that consists of two equally likely percentages: a lower percentage (L) and a higher percentage (H). That standard deviation is (H − L)/2.

For analyzing standard deviations, we draw on traditional psychometric theory (Lord & Novick, 1968). We distinguish a perceiver's *observed accuracy* on a lie/truth detection test from the perceiver's *real accuracy*—defining the latter as the percentage of messages the perceiver would judge correctly on a test that included an *infinite* number of lies and truths. From standard deviations in observed accuracy, our methods will allow an inference about the standard deviation in real accuracy. Theoretically, this *real standard deviation* in lie detection abilities must be smaller than the observed standard deviation, and the difference between observed and real standard deviation should (under psychometric theory) depend on the length of the lie/truth test: the smaller the number of messages on a test, the greater should be the artifactual inflation of observed individual differences (Lord & Novick, 1968).[2]

To study individual differences in detection abilities, we began by determining the judge-to-judge standard deviation in percentage of correct lie/truth judgments from as many samples as possible. We managed to compute the value of this statistic from 247 independent samples of judges. In these samples, we find that the observed standard deviation in lie detection ability has an average value of 12.78.

According to psychometric theory, there should be an inverse relationship between the standard deviation among a set of lie detection performances, and the number of judgments each performance entails. To assess this prediction, we examined the relationship between number of judgments and observed standard deviation across our 247 samples. Results appear in Figure 2. As shown there, accuracy differences are much smaller among judges who make a large number of lie/truth judgments than judges who make only a few. This pattern of results corroborates the psychometric expectation and suggests that detection performance differences are inflated by the brevity of researchers' lie/truth tests.

Given our interest in *real* ability differences, we focused on people who make a large number of judgments. Those data appear in the right side of Figure 2. There we see a depiction of results for the longest test to date (Parker, 1978)—one that required each judge to classify 120 statements as lies or truths. It is noteworthy that the standard deviation in accuracy across judges on Parker's (1978) lie/truth test is the lowest of the 247 standard deviations reported to date—3.09%. In our view, 3.09% provides a preliminary upper bound to *real* individual differences in detection ability. However, we sought a more precise estimate. We wished to infer the magnitude of differences from judge-to-judge in percentage correct that would be obtained on a lie/truth test much longer than 120 items—a test of *infinite* length.

Before describing our method for estimating this *real* standard deviation, let us note that in each study included in Figure 2, judges classified messages into one of two categories—indicating that the message was either a lie or the truth. For purposes of analysis, let us suppose that each judge in a given sample has a certain probability of classifying a given message correctly and that this probability is the same for all messages. Then we could use the binomial distribution to model the percentage of correct judgments in the sample. Under this model, the standard deviation across perceivers in percentage correct would be inversely proportional to the square root of the number of lie/truth judgments each judge made. For details, see Lord and Novick (1968, Chapter 23). We exploited this fact to estimate the standard deviation that one would expect in a sample that made an infinite number of judgments. To do so, we fitted to the data in Figure 2 the regression equation

[2] The quantity of primary interest to us is the standard deviation among judges that would be observed if each judge responded to a lie/truth test of infinite length. In the text, we call this psychometric quantity the *real standard deviation*. Traditionally, it was called the *true standard deviation*—that is, the standard deviation among a set of true scores. As psychometricians know, these statistics are used to disattenuate correlation coefficients.

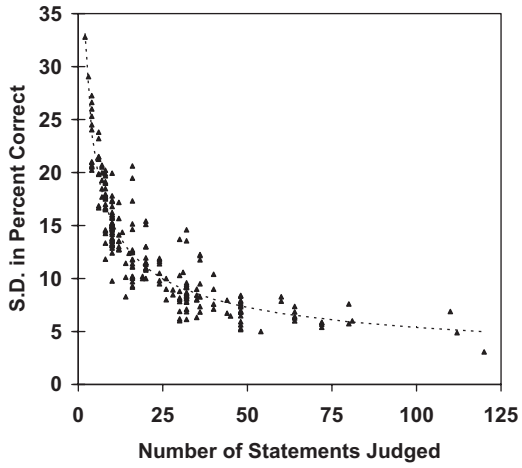

Figure 2. Standard deviation in percentage of statements judged correctly by number of statements judged.

$$s_i = a + b\left(\frac{1}{\sqrt{n_i}}\right) \quad (1)$$

where s_i is the observed standard deviation in sample i among judges in the percentage of lies and truths that they correctly detected, and n_i is the number of lie/truth judgments made by each judge in that sample.

Let us explain the logic of this equation. Note that b is the regression coefficient in the equation. A positive value for b would allow us to accommodate the pattern of results in Figure 2—the inverse relationship between the standard deviation in lie/truth accuracy across judges in a sample and the number of lie/truth judgments that each judge rendered. We use the equation to predict the standard deviation in a hypothetical sample in which each judge made an infinite number of lie/truth judgments. Given an infinitely large value of n_i, the term inside the parentheses to the right of Equation 1 becomes 0, and the predicted standard deviation across perceivers for that sample would be a—the intercept of Equation 1. This intercept provides a model-based estimate of *real standard deviation*, the standard deviation across judges in lie/truth discrimination ability that would be observed on a test of infinite length. Here, we call this estimate the *measurement-corrected standard deviation*.

Fitting Equation 1 to the data in Figure 2, we found $a = 0.80$, $b = 45.75$. This model fits the data very well. Across the 247 samples depicted in Figure 2, the predicted standard deviations correlate .92 with the actual standard deviations. We have inserted the standard deviations predicted by Equation 1 into Figure 2 as a dashed line.

A conventional hypothesis test suggests that there are real individual differences in lie detection ability—for comparison of the intercept with 0, $t(244) = 2.33$, $p < .05$. However, these ability differences are small. On a test of infinite length, the standard deviation across judges in percentage correct would be less than 1%. It would be 0.80%. The 95% confidence interval for this measurement-corrected standard deviation is 0.12%–1.47%.[3]

O'Sullivan (2007) hypothesized that differences in lie detection ability are normally distributed. In light of that conjecture, let us remind the reader that roughly 95% of the observations in a normal distribution are within two standard deviations of the mean. Let us also note that in such a distribution less than 1 observation in 2 million is more than five standard deviations above the mean.

If we make O'Sullivan's (2007) distributional assumption and further assume that the mean percentage correct lie/truth judgments is 54% (C. F. Bond & DePaulo, 2006), our current results allow a complete specification of the distribution of lie detection abilities. On a test of infinite length taken under the typical conditions of deception detection research, 95% of judges would achieve between 52.4% and 55.6% correct lie/truth discrimination (that is, the mean plus and minus two measurement-corrected standard deviations). If 2 million judges took a test of infinite length under the usual conditions, we would expect less than 1 to achieve more than 58% correct (that is, five real standard deviations above the mean).

Perhaps people differ widely in the ability to spot deception cues but not in the ability to identify truths as truths. To assess this possibility, we found 115 samples in which it was possible to code separately the standard deviation across judges in accuracy for deceptive messages and accuracy for truthful messages. Averaging across studies, there are no differences in these standard deviations, $M = 17.97$ for the standard deviation in detecting lies versus 18.29 for the standard deviation in detecting truths, $t(114) = -0.76$, *ns*.

Under the hypothesis of ability differences in lie/truth discrimination, one might suppose that individuals who were good at spotting lies would also be good at spotting truths. To assess this possibility, we determined in 154 samples the correlation between a judge's accuracy at detecting lies and that same judge's accuracy at detecting truths. In fact, the relevant correlation is negative in 97 of the 154 samples (that is, 63% of the time). For the relationship between accuracy at judging lies and accuracy at judging truths, standard meta-analytic methods yield a weighted mean r-to-Z-to-$r = -.09$ (95% confidence interval = $-.11$ to $-.07$). The most accurate judges of lies tend to be the least accurate judges of truths.

Range. We also had an interest in the range in judges' lie detection performances. In 88 independent samples of judges, we were able to determine the highest and lowest percentage accuracy achieved by any individual. We coded these statistics.

In Appendix B (which is available online as supplemental material), we derive an equation for the maximum and minimum percentage of correct lie-or-truth judgments that would be expected in a sample—if there were no ability differences among the judges in that sample (see David & Nagaraja, 2003). In applying this equation to the research literature, we assume that the proba-

[3] In the text, we estimate real standard deviations by applying Equation 1 to the pattern of observed standard deviations across studies. We also estimated a real standard deviation within each study with traditional psychometric procedures (Lord & Novick, 1968). These within-study analyses yielded results similar to those we report in the text—extremely small measurement-corrected standard deviations in judge ability, larger measurement-corrected standard deviations in judge credulity and target transparency, and the largest measurement-corrected standard deviations in target credibility. Given the sample sizes in this literature, however, within-study estimates are unstable. For more information, contact Charles F. Bond Jr.

bility of a lie/truth judgment being correct in a sample is the proportion of correct judgments observed in that sample.

We used this technique to analyze each of the 88 samples, for which we knew the range across judges in percentage correct. Results show that there is a very strong correlation between the observed range in detection accuracy in a sample and the range that would be expected if there were no ability differences (r = .88, $p < .0001$). Averaging across the 88 samples, the mean observed range in percentage correct is 44.38%; the mean range that would be expected given no ability differences from judge to judge is 44.18%. The observed range is less than 1% wider than the expected range. In fact, it is 0.20% wider. This difference is not statistically significant, $t(87) = 0.21$, *ns*. For the mean observed and expected range in judge ability across our 88 samples, see the two bars that appear in the left side of Figure 3. Each bar extends from the minimum percentage correct to the maximum percentage correct, the dark bar representing percentages observed and the light bar representing percentages expected by chance.

O'Sullivan (2007) has maintained that a few people have extraordinary skill at detecting deceit. To assess this claim, we analyzed the maximum percentage correct achieved in 93 samples of judges (the 88 in which a minimum accuracy was reported and 5 others). Averaging across all 93 samples, the observed maximum percentage correct is 75.70%; the maximum that would be expected if judges did not differ in ability is 76.44%. Observed and expected maximum performances are highly correlated (r = .81) and do not differ significantly from one another, $t(92) = -1.07$, *ns*. These data provide no evidence that the best lie detection performances in this research literature reflect any extraordinary ability. The highest detection rates are no higher than chance would produce.

Judge Credulity

Next, we analyzed differences from judge to judge in the tendency to regard others as truthful.

Standard deviation. In 162 samples, it was possible to determine the standard deviation from judge to judge in the percentage of statements classified as truthful. The mean standard deviation observed in these samples was 13.55. We hoped to partition the differences researchers observed in credulity into real differences and artifactual differences.

To do so, we fit Equation 1 to the 162 standard deviations. The equation fit the data well, for the relationship between predicted and observed standard deviations (r = .77, $p < .0001$). Again, our interest is in the real standard deviation, the standard deviation across judges in the percentage of messages they would classify as truths on a test of infinite length. Again, the intercept of Equation 1 provides our meta-analytic estimate of this quantity and indicates that perceivers do differ from one another in credulity (for this measurement-corrected standard deviation, a = 5.13%; 95% confidence interval = 3.94%–6.32%), $t(160) = 8.52$, $p < .0001$.

Although our psychometric analyses suggest that judges differ more in credulity than ability, one wonders whether similar results would be apparent in a more controlled comparison. They are. In 152 samples, we were able to determine both the standard deviation among a set of judges in percentage correct lie/truth judgments as well as the standard deviation among those same judges in percentage truth judgments. These judges vary more in credulity than in accuracy; mean observed standard deviations = 13.18 versus 12.25, respectively; for the difference, $t(151) = 3.15$, $p < .01$.

Range. It was possible in 32 samples to determine the largest and smallest percentage of statements classified as truthful by any judge. To analyze these statistics, we adapted the methods of Appendix B. These allow us to determine the range in the percentage of truth judgments judges rendered that we would observe in a sample of judges who were equal in credulity. Averaging across the 32 samples, the mean observed range in truth judgments was 50.06%; a mean of 35.86% would be expected if there were no real differences in credulity. Judges range more widely in credulity than would be expected by chance—for comparison of the means in observed and expected range, $t(31) = 6.25$, $p < .0001$. In fact, the observed range is 40% wider than what would be expected by chance. For these results, see the third and fourth bars of Figure 3.

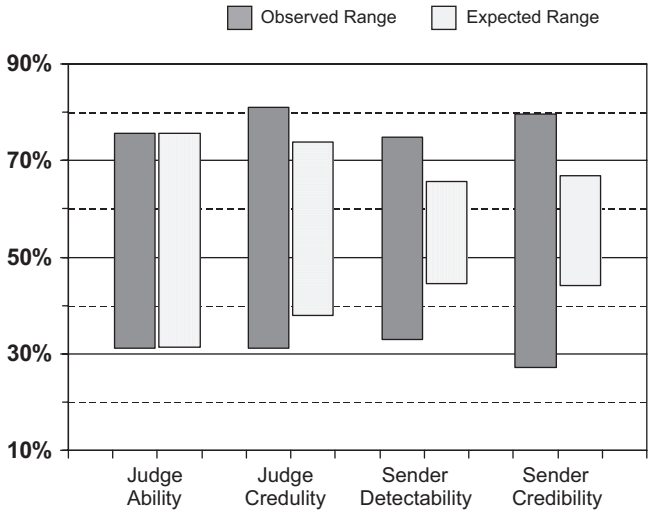

Figure 3. Observed and expected range in four individual differences.

Sender Detectability

Having found that judges show negligible differences in lie detection ability and nonnegligible differences in credulity, we turn our attention to the targets of judges' detection efforts: people who lie and tell the truth. Some researchers have claimed that people differ more as liars than as lie detectors (C. F. Bond et al., 1985).

People might differ in the transparency of their veracity. If so, it would be easy to spot some individuals' lies and truths, and it would be impossible to determine whether others were lying.

Standard deviation. We noted (in 54 samples in which data were available) the standard deviation from sender to sender in the percentage of times that judges correctly detected the sender's lies and truths. These observed standard deviations had a mean of 11.83.

We are interested in the real standard deviation in sender detectability, the standard deviation that would be observed if each sender was judged an infinite number of times. To estimate that hypothetical quantity, we used a variant of Equation 1. We predicted the standard deviation in a sample of senders from the

reciprocal of the square root of the number of judgments made of each sender in that sample. Theoretically, differences among senders should be smaller, the larger is the number of judgments made of each sender. Our analysis confirms this theoretical prediction, and the relevant regression equation fits the data well—for the correlation between observed and predicted standard deviations in 54 samples ($r = .74$, $p < .001$). From the equation, we estimate that if an infinite number of lie and truth judgments were made of each sender, some would be more detectable than others. The measurement-corrected standard deviation in detectability is 5.49% (95% confidence interval = 3.51%–7.46%), $t(52) = 5.57$, $p < .0001$.

Range. In 37 samples, we could determine the largest and smallest percentage of correct lie/truth judgments made of any sender's statements. Averaging across these samples, the mean observed range in accurate judgments received was 41.93%. The methods of Appendix B reveal that a mean range of 21.13% would be expected if there were no real differences in sender detectability. People range more widely in detectability than would be expected by chance—for comparison of mean observed with mean expected range, $t(36) = 8.19$, $p < .0001$. The observed range is 1.98 times as wide as the expected range (see the fifth and sixth bars in Figure 3).

Sender Credibility

When telling lies and truths, some people are more detectable than others—as we have discovered. In principle, people may also differ in credibility. Some may appear honest and others dishonest, regardless of their veracity.

Standard deviation. In 45 samples, we could determine a standard deviation from sender to sender in the percentage of times their statements were judged as truthful. These standard deviations show a mean value of 14.77.

To estimate the real standard deviation in sender credibility, we again used Equation 1 to predict the standard deviation in a sample of senders, substituting for n_i the number of judgments made of each sender in that sample. The equation fit moderately well—predicted standard deviations in sender credibility yielding an r of .49 with observed standard deviations ($p < .005$). From the fitted intercept of our regression equation, we infer that if people made an infinite number of judgments of senders, the senders would differ from another in credibility (measurement-corrected standard deviation = 11.58%; 95% confidence interval = 9.36%–13.80%), $t(43) = 10.52$, $p < .0001$.

From our analyses, it would appear that people vary more from one another in credibility than detectability. This difference also emerged in a controlled comparison. In 38 samples, we were able to code the standard deviation in the detectability and credibility of the same senders, on the basis of the same lie/truth judgments. The senders vary more from one another in credibility than detectability (mean observed standard deviations = 15.06% vs. 12.70%), $t(37) = 2.77$, $p < .01$.

Perhaps people vary more widely in credibility when lying than when telling the truth. To assess this possibility, we found 33 studies in which it was possible to code separately standard deviations in the percentage of truth judgments senders received when lying and in the percentage of truth judgments they received when telling the truth. Results show that, in fact, credibility differences are greater when people are lying rather than telling the truth (observed SDs = 17.98 and 15.95 for percentage of truth judgments to lies vs. truths, respectively), $t(32) = 2.59$, $p < .05$.

DePaulo and Rosenthal (1979) reported evidence of a demeanor bias—that the individuals who appear most honest when lying are the ones who appear most honest when telling the truth. In 35 samples, it was possible to determine the correlation between the percentage of truth judgments to a person's lie and that same person's truth. These data corroborate DePaulo and Rosenthal's evidence of demeanor bias—the correlation is positive in 29 of 35 samples. By standard meta-analytic methods, the relationship between percentage of truth judgments to an individual's truth and that same individual's lie yields a weighted r-to-Z-to-$r = .39$, $p < .01$ (95% confidence interval = .34–.44).

Range. It was possible in 31 independent samples to determine the largest and smallest percentage of truth judgments made about any sender's statements. Averaging across these samples, the mean observed range in truth judgments received was 52.51%; by the methods of Appendix B, a mean of 22.73% would be expected if senders did not differ in credibility. Senders range more widely in credibility than would be expected by chance—for comparison of mean observed with mean expected range, $t(30) = 12.54$, $p < .0001$. For the observed and expected range in sender credibility, see the right-most two bars in Figure 3. The observed range is 2.43 times as wide as the expected range.

Comparing Differences

In contrast to small differences in judge lie detection ability, there are substantial individual differences in the senders' apparent honesty. In terms of percentages, measurement-corrected differences in sender credibility are roughly 14 times the size of the corresponding differences in judge ability (11.58% vs. 0.80%). Differences in sender detectability and judge credulity are roughly equal to one another (5.49% and 5.13%), and each is roughly half as large as sender credibility differences (see Figure 4).

It is possible to compare the impact of these individual differences on deception judgments with the impact of situational factors—for example, the sender's veracity. Sometimes people are lying; sometimes they are telling the truth. An earlier meta-analysis revealed that on average 61.34% of truths and 52.45% of lies are judged to be truths (C. F. Bond & DePaulo, 2006). This

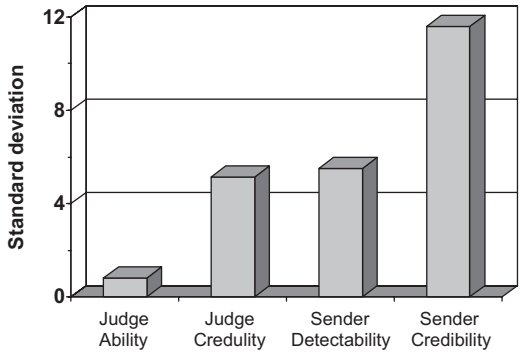

Figure 4. Measurement-corrected standard deviation in four individual differences.

implies that sender veracity introduces a standard deviation in percentage truth judgments of 4.44% (half of 61.34 – 52.45). In percentage terms, a person's credibility has more than twice the impact of the person's veracity in determining whether s/he will be judged truthful.

Although for ease of interpretability we have assessed individual differences as standard deviations, it is conventional to express them as variances. In the latter metric, the comparisons we have been making are more striking. Correcting for sampling error, the variance in sender credibility is over 200 times as large as the corresponding variance in judge ability. Relatively speaking, ability variance is trivial.

Psychologists are interested in the reliability of individual difference measures. Under traditional theory, the reliability coefficient for a measure is the ratio of the real variance in that measure to the observed variance. We estimated reliability coefficients for judge ability, judge credulity, sender detectability, and sender credibility by forming the ratio of our model-based measurement-corrected variance in each of those measures to the mean of the observed variances in each measure that have been reported in the research literature. This method yields the following reliability coefficients: judge ability = .07, judge credulity = .30, sender detectability = .58, and sender credibility = .91. In this metric, differences in detectability from sender to sender are more reliable than differences in credulity from judge to judge. Judges show no reliable differences in ability, whereas sender differences in credibility are highly reliable.

We also assessed individual differences by analyzing another statistic—the range. These analyses corroborate our conclusions. Here, a useful metric is the ratio of the mean of the observed ranges to the mean of the ranges expected by chance. Our results reveal that this ratio is 1.00 for judge ability, 1.40 for judge credulity, 1.98 for sender detectability, and 2.43 for sender credibility.

Moderator Analyses

Individual differences in judging deception could be influenced by many factors. Individual differences might, for example, be larger when deception must be inferred from audio than from video. They might be unusually large when liars are motivated not to get caught.

To assess the moderation of individual differences in judging deception, we conducted a number of analyses. Each analysis yielded a measurement-corrected standard deviation in 1 of 17 subsets of the research literature. These subsets of the literature were defined by eight variables: deception medium (audio-only, video-only, or audiovisual), liar motivation (high or low), liar interaction (some or none), lie preparation (some or none), judge experience (some or none), baseline exposure (some of none), document status (published or unpublished), and lie/truth discrimination (mean percentage correct is or is not significantly greater than 50% at one-tailed $p < .05$). With the regression-based procedure outlined in Equation 1 above, we computed measurement-corrected standard deviations within each of these subsets of the literature on our four individual-difference variables: judge ability, judge credulity, sender detectability, and sender credibility.

Let us briefly describe the pattern of results produced by our analyses: Differences in judge ability are consistently small, and differences in sender credibility are consistently large. Across 17 subsets of this research literature, measurement-corrected standard deviations in judge ability range from −0.19% (among experienced judges) to 1.47% (for video lies).[4] Measurement-corrected standard deviations in judge credulity range from 4.39% (for motivated liars) to 8.47% (for audio lies). Measurement-corrected standard deviations in sender detectability range from 2.75% (for video lies) to 16.35% (when the judged are experienced). Measurement-corrected standard deviations in sender credibility range from 6.80% (for video lies) to 18.59% (when judges are experienced). For relevant results, see Table 1.

Above, we noted that the largest of four individual differences in judging deception is sender credibility, and the smallest is judge ability. As Table 1 reveals, this pattern holds true in each of 17 subsets of the research literature. In studies in which liars are motivated, in studies of interactive deception, in studies in which judges discriminate lies from truths, and in studies that are published (as well as those that are unpublished), measurement-corrected standard deviations are highest for sender credibility, lowest for judge ability, and intermediate for judge credulity and sender detectability.

The small size of individual differences in judge ability should be apparent from an inspection of Table 1. There it is noteworthy that the *highest* measurement-corrected standard deviation in judge ability for *any* subset of this literature (1.47%) is lower than the lowest measurement-corrected standard deviation in any of the other three individual differences in any subset of the literature (2.75%). Across 17 sets of studies, these distributions do not overlap. Thus, our conclusions about the relative size of individual differences in judging deception generalize across deception media, liar motivation, liar interaction, lie preparation, judge experience, baseline exposure, publication status, and lie/truth discrimination.

Discussion

Here, we have drawn conclusions about individual differences in judging deception. Our goal was to quantify such individual differences as would be apparent if researchers could measure the differences without error. Our database consists of all relevant studies that we could find, and to this database we applied a new meta-analytic technique. In this section, we discuss our individual difference findings, consider the applicability of these findings to lie detection in the real world, and comment on our method for meta analyzing individual differences.

Individual Differences in Accuracy

Although it has become virtually axiomatic that the *mean* lie detection performances of groups of people are barely above chance, the magnitude of individual differences in detection ability was not previously known. Now, several converging lines of evidence indicate that virtually all *individuals* are barely able to

[4] Our model-based estimate for the real standard deviation among expert judges is −0.19%. Standard deviations cannot be negative; hence, our negative estimate would be anomalous if it differed significantly from zero. It does not; thus, we infer that expert judges do not differ from one another in the ability to detect lies.

Table 1
Measurement-Corrected Standard Deviations in Subsets of the Research Literature

Variable	Judge ability	Judge credulity	Sender detectability	Sender crediblity
Audio only	0.44	8.47	5.92	8.85
Video only	1.47	4.40	2.75	6.80
Audiovisual	0.71	4.86	5.15	7.97
Unmotivated liar	1.21	5.40	5.57	11.76
Motivated liar	0.32	4.39	3.85	7.41
No interaction	0.98	5.02	4.31	12.74
Interaction	0.93	4.67	5.56	7.47
Unprepared liar	0.89	6.50	4.91	12.50
Prepared liar	0.63	4.93	7.46	9.26
Inexperienced judge	0.84	4.91	5.18	11.23
Experienced judge	−0.19	6.58	16.35	18.58
No baseline exposure	0.78	5.04	5.42	11.58
Baseline exposure	1.10	5.32	8.13	—
Unpublished document	0.06	5.27	4.68	7.59
Published document	1.30	5.06	5.25	11.87
No significant mean detection	0.41	5.24	6.84	12.24
Significant mean detection	1.44	5.88	4.84	7.93

Note. A dash indicates that no standard deviations in sender credibility could be computed in this subset of the research literature.

detect lies, and that real differences in detection ability are miniscule.

In demonstrating that people differ little from one another in the ability to detect lies, we build on earlier work. It showed that accuracy in judging the veracity of one person is independent of accuracy in judging the veracity of another (Kraut, 1978), that accuracy in judging lies is not positively correlated with accuracy in judging truths (Levine et al., 1999), and that no "wizardry" need be invoked to explain why a few people get high scores on lie detection tests (C. F. Bond & Uysal, 2007). Our contribution is to demonstrate the small size of lie detection ability differences not from a single study (as had earlier researchers) but from a meta-analysis of the research literature as a whole. A second contribution is to compare the magnitude of lie detection ability differences with the size of some related individual differences.

Our meta-analytic results clarify a phenomenon that would otherwise be curious. Many psychologists have attempted to uncover the traits of individuals who are particularly gifted in divining deceit—traits like the individual's education, sex, occupation, Machiavellianism, self-monitoring, and locus of control. Although each of these individual difference variables may be related to lie detection accuracy in a few studies, once a reasonable amount of evidence accumulates over a reasonable number of laboratories, these individual-difference relationships prove to be illusory (Aamodt & Mitchell, 2006; Zuckerman et al., 1981). Across the literature as a whole, there is no replicable predictor of lie detection accuracy. This is unsurprising. As currently measured, lie detection accuracy is not a reliable individual difference. Thus, nonchance correlates of current accuracy scores are unlikely to be found.

While yielding these null individual difference results, the small size of lie detection ability differences has made a positive contribution to experimental research. There, investigators test for the impact of situational factors on human lie detection, assessing these situational effects against an error term that consists of differences among judges' lie detection performances. Insofar as judges' performances show little variability, deception researchers have had high statistical power to uncover the effects that interest them. Small ability differences have also allowed many investigators to find that rates of lie/truth discrimination are statistically significant, even when the rates are only slightly above 50%. To accrue the benefits of high statistical power, those who conduct deception experiments should plan to give their research participants long lie detection tests. If the researcher's error term is to consist of the variability among judges' lie detection performances, this error term will be smaller the larger the number of lies and truths judged.

Although people hardly vary in the ability to detect deception, they differ in their detectability as liars. When lying, some people get caught, and others elude capture. Thus, in an individual-difference sense, the accuracy of a deception judgment depends more on the liar than the judge.

Having noted that on the whole lies can barely be distinguished from truths (C. F. Bond & DePaulo, 2006), let us mention a complicating factor. Some individuals tell lies and truths that are readily distinguishable, and their transparency merits discussion. Although we do not know why the veracity of certain individuals is obvious, perhaps these people have ethical compunctions against lying; perhaps they cannot regulate deception-related emotions; perhaps they are poor at masking those emotions; perhaps for them

the truth is so cognitively prepotent that they have trouble concocting plausible tales. Future research will be needed to explain why (as lie- and truth-tellers) some people are highly transparent and others opaque. In the meantime, it is worth reiterating that there *is* an individual difference in detecting deception. Although differences from judge to judge in lie detection are small, differences from sender to sender in lie detectability are much larger.

Individual Differences in Bias

Deception judgments can have important consequences, and sometimes the consequences do not depend on whether the judgment is correct (C. F. Bond et al, 1992). Thus, we sought to understand individual differences that might bias one person to judge another as truthful (on the one hand) or deceptive (on the other).

On average, people show a bias toward judging others as truthful (Levine et al., 1999). However, individual differences in this judgmental bias have scarcely been discussed. Here, we found that people do, in fact, vary in the tendency to regard others as truthful. They vary from one another in ways that cannot be attributed to measurement error. This individual difference is related to a broader suspiciousness: People who are most leery of communications in general are the ones most likely to regard others as lying (Levine & McCornack, 1991). Perhaps these individuals have learned to be suspicious because they have often been the victims of deceit, or maybe those who are chronically suspicious of deception are themselves habitual liars. However they are explained, individual judge differences in the bias to perceive others as truthful are roughly equal in magnitude to individual sender differences in lie–truth detectability.

The largest determinant of a deception judgment is not, however, the judge's degree of truth bias or the sender's detectability. Instead, it is the credibility of the person being judged—some individuals appear substantially more truthful than others. In fact, a person's credibility has a bigger impact than the person's honesty on whether s/he will be seen to be telling the truth. High credibility liars are more likely to be believed than low credibility truth-tellers.

Let us try to explain why some people appear more truthful than others. In our view, differences in apparent honesty emerge soon after birth by virtue of facial anatomy. Some infants are anatomically gifted with an honest-looking face; others are facially disadvantaged. The gifted have baby faces, and the disadvantaged look mature (Masip, Garrido, & Herrero, 2004). Individual differences in facial honesty carry forward over the lifespan (Zebrowitz, Voinescu, & Collins, 1996) and help explain why some people are more likely than others to be seen as telling the truth.

Facial anatomy, however, is not sufficient to explain the accumulated research findings. As our meta-analytic results show, people differ more in credibility when lying than when telling the truth. Differential practice may account for this effect. As youngsters, all people have occasion to lie. Some children, the facially honest, will discover that they can avoid punishment by lying. This social reinforcement motivates them to continue lying, and with practice, they hone their deception skills. Others, children who have a dishonest face, learn that lying does them no good. Indeed, deceptive-looking youngsters may be punished for their failed attempts at deception. Thus, they learn to refrain from lying and never develop whatever behavioral potential for deception they might have had. So the deceptively rich get richer, and the poor stay the same. In this view (C. F. Bond & Robinson, 1988), social reinforcement contingencies augment anatomical differences to explain why some people appear more honest than others.

Real World Applicability

Having drawn conclusions about individual differences in judging deception from experimental studies, let us comment on the applicability of our findings to lie detection in the real world.

Critics characterize deception research as artificial. They argue that experimental deceptions are trivial and that research participants tell lies in an asocial context. They note that experimenters study lies between strangers and deprive would-be detectors of information about how liars appear when telling the truth. Perhaps in experimental research, liars act similarly to truth-tellers, and if judges do not differ from one another in detecting deceit, it is because they have no cues to detect.

Although these criticisms can be made of many studies of deception detection, they cannot be made of others. Researchers have studied high-stakes lies—lies told by murderers, for example (Vrij & Mann, 2001). Researchers have studied naturalistic deceptive interactions (Kassin & Fong, 1999) and have given judges baseline information about senders' truthful behaviors (Feeley, deTurck, & Young, 1995). Researchers have studied lies that *can* be discriminated from truths—with 72% accuracy, in one case (Vrij, Mann, Robbins, & Robinson, 2006).

Results show that our individual difference findings are not restricted to the most artificial of research studies; rather they are consistent across the deception detection research literature as a whole. Yes, people barely differ from one another in the ability to detect low-stakes lies told in noninteractive contexts when there is no evidence that the lies can be detected. Yet, there are also small differences in lie detection ability among individuals who judge motivated lies, among people who judge interactive lies, and among those who judge lies that have (in fact) been detected. Differences in judge ability are small whether judges receive baseline information about the way liars act when telling the truth.

As judge differences in detection ability are small in all parts of this research literature, sender differences in detectability are larger. In studies of low-stakes lies told in noninteractive contexts in which there is no evidence that the lies were discriminated from truths, people differ from one another in their detectability as lie- and truth-tellers. If artificial methodology masks individual differences in the ability to detect lies, it is peculiar that it leaves unmasked the corresponding differences in sender detectability.

Having noted that our meta-analytic findings are not specific to trivial lies, asocial lies, or undetectable lies, we do not mean to imply that this pattern of individual differences would emerge in every instance of lie detection in the real world. In our view, experimenters have not yet captured several features of real world lie detection. In the real world (but not the laboratory), lie detection requires unprompted suspicion, involves nonbehavioral evidence, and entails nonrandom liar–judge pairings

Experimenters forewarn the judges in their studies that they will be seeing deception and instruct them to consider whether each message they encounter is a lie. Under these conditions, people differ in the tendency to regard others as deceptive. Outside the

laboratory, different conditions obtain. To detect a lie, one must first come to suspect deceit, and there may be individual differences in prerequisite suspicion. The possibility of deception may rarely occur to some people and be chronically salient to others. If so, individual differences in the tendency to regard others as deceptive may be larger outside the laboratory than the differences we have found here. It is also possible that people differ in responsiveness to real-world indications that suspicion is warranted. If this alertness is as much an acuity as a generalized distrust, there could be real-world individual differences in accuracy at detecting deception—differences that are obscured when experimenters prompt all judges to suspect deceit. These possibilities should be explored.

In experiments, judges must detect deception solely from the behavior and speech people display when lying. Outside the laboratory, people infer deception from other forms of evidence. They rely on motivational information, physical evidence, and information from third parties. Research indicates that in the real world people rarely detect deception at the time a lie is told. Rather, they infer deceit days, weeks, or months later (H. S. Park, Levine, McCornack, Morrison, & Ferrara, 2002). Perhaps there are individual differences in people's ability to use nonbehavioral cues to deceit, and people differ in sensitivity to evidence of lies they were told earlier. It is also possible that some people are better than others in spinning fabrications that will be immune to nonbehavioral evidence and resistant to delayed exposure. New research paradigms would be needed to uncover these skills.

In the natural ecology of deceit, individuals who have different traits may gravitate toward different interaction partners. It is conceivable, for instance, that certain kinds of individuals tend to interact with highly detectable liars, and others with people whose lies are opaque. If so, the former would achieve higher levels of real world lie detection than the latter. New naturalistic research would be required to explore this possibility.

Meta-Analytic Method

Finally, let us comment on our statistical method for assessing individual differences in judging deception. Our goal was to estimate a psychometric quantity: the standard deviation among a group of individuals in the percentage of lies and truths they would correctly detect, if each individual took a lie detection test of infinite length. We constructed a regression equation to estimate this quantity. In particular, our equation predicted the standard deviation in lie detection across the judges in a sample from the reciprocal of the square root of the number of lies and truths that each individual judged. Applying this equation to data from a large research literature, we found that it could accurately predict the standard deviations that investigators observed. We could then use the fitted equation to make predictions. For our hypothetical sample of individuals who made an infinite number of lie/truth judgments, the predictor variable in this equation would be 0 (i.e., one divided by infinity); hence, we would predict a standard deviation equal to the *y*-intercept of the equation. This was the psychometric quantity of interest—the real standard deviation, a standard deviation corrected for random error in the sampling of lies and truths.

Our regression-based method invites comparison with other methods. For estimating a real standard deviation in the ability of a single sample of judges, there are traditional procedures (Lord & Novick, 1968). Given access to raw data from a number of samples, one could estimate a real standard deviation within each sample and cumulate the estimates across samples. A cumulated within-study estimate of the real standard deviation might require fewer assumptions than the estimate we report here. However, within-study psychometric techniques are harder to understand than our regression-based approach. Also, the best within-study procedures require access to primary data that meta-analysts lack.

In its goal of correcting differences for statistical error, our regression-based procedure is reminiscent of random-effects meta-analysis (Hedges & Vevea, 1998). However, the similarity may be more apparent than real. Random-effects meta-analysis concerns the differences across *studies* in a summary statistic. It focuses on participant sampling error—random variability across the research findings in a literature that results from the fact that investigators study samples of research participants rather than populations. The smaller the number of research participants in a study, the greater is the participant sampling error introduced into the outcome of that study. Hunter and Schmidt (1990) developed procedures to estimate the participant sampling error that one would expect across the correlation coefficients in a research literature from the number of participants on which each correlation in the literature was based. These authors have advised meta-analysts to subtract this sampling error variance from the variance across the literature in the correlation coefficients observed and to regard the difference as a variance among population correlation coefficients. Applying this logic to research on deception detection, C. F. Bond and DePaulo (2006) used a random-effects meta-analysis to estimate that the standard deviation across studies in the percentage of lies and truths detected in this research literature would be 4.52% if an infinite number of individuals made judgments in each study.

The current regression-based procedure is, by contrast, not intended to assess differences across *studies* but rather differences across *individuals*. It estimates how much individuals would differ from one another as lie detectors on a test that was infinitely long and how much individuals would differ from one another as liars if each was subject to an infinite number of judgments. Thus, our analyses of judge differences correct for error in the sampling of lies and truths, whereas our analyses of sender differences correct for error in the sampling of judgments.

Like the methods outlined by C. F. Bond, Wiitala, and Richard (2003), our regression-based approach yields results that are in the raw metric used by researchers—percentage of correct lie/truth judgments, for instance. By contrast, most random-effects meta-analyses produce findings in a metric that has been subjected to statistical standardization, then squared—the variance among population standardized mean differences, for example (Hedges & Vevea, 1998). For advantages of retaining the raw metric, see C. F. Bond, Wiitala, and Richard (2003).

We invite colleagues to study other individual differences with our method. They might, for instance, use this technique to quantify individual differences among senders and receivers in the communication of emotion (J. A. Hall, 1984), empathic accuracy (Davis & Krauss, 1997), or meta-perception (Kenny, 1994). Here, we find that judges vary little in the ability to detect deception, whereas senders vary substantially in their tendency to appear deceptive. Perhaps this reflects a more general phenomenon—for social perception to depend more on the sender than on the

receiver (Kenny & LaVoie, 1984). Further meta-analytic work will be needed to evaluate this possibility.

In the meantime, we have analyzed individual differences in a large research literature. Here, deception judgments depend more on the liar than the judge.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Aamodt, M. G., & Mitchell, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examiner, 15,* 6–11.

*Al-Simadi, F. A. (2000). Detection of deceptive behavior: A cross-cultural test. *Social Behavior & Personality, 28,* 455–461.

*Anderson, D. E., DePaulo, B. M., & Ansfield, M. E. (2002). The development of deception detection skill: A longitudinal study of same-sex friends. *Personality and Social Psychology Bulletin, 28,* 536–543.

*Ask, K., & Granhag, P. A. (2003). Individual determinants of deception detection performance: Need for closure, attributional complexity, and absorption. *Goteborg Psychological Reports, 33,* 1–13.

*Atmiyanandana, V. (1976). *An experimental study of the detection of deception in cross-cultural communication.* Unpublished doctoral dissertation, Florida State University, Tallahassee.

*Bailey, J. T. (2002). *Detecting deception when motivated: The effects of accountability and training on veracity judgments.* Unpublished mater's thesis, Ohio University, Athens.

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

*Berger, R. E. (1977). Machiavellianism and detecting deception in facial nonverbal communication. *Towson State University Journal of Psychology, 1,* 25–31.

*Billings, F. J. (2004). *Psychopathy and the ability to deceive.* Unpublished doctoral dissertation, University of Texas at El Paso.

*Blair, J. P. (2006). *The impact of external information on deception detection.* Paper presented at the annual meeting of the National Communication Association, San Antonio, TX.

*Blair, J. P., & McCamey, W. P. (2002). Detection of deception: An analysis of the behavioral analysis interview technique. *Illinois Law Enforcement Executive Forum, 2,* 165–169.

*Bond, C. F., Jr. (2006). *Accuracy in judging self-initiated lies.* Unpublished data, Texas Christian University, Fort Worth.

*Bond, C. F., Jr., & Atoum, A. O. (2000). International deception. *Personality and Social Psychology Bulletin, 26,* 385–395.

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214–234.

*Bond, C. F., Jr., & Fahey, W. E. (1987). False suspicion and the misperception of deceit. *British Journal of Social Psychology, 26,* 41–46.

*Bond, C. F., Jr., & Gresky, D. P. (2006). *Investigations of deceit.* Paper presented at the annual meeting of the National Communication Association, San Antonio, TX.

*Bond, C. F., Jr., Howard, A. H., & Hutchison, J. (2006). *Overlooking the obvious: Incentives to lie.* Unpublished manuscript, Texas Christian University, Fort Worth.

*Bond, C. F., Jr., Kahler, K. N., & Paolicelli, L. M. (1985). The miscommunication of deception: An adaptive perspective. *Journal of Experimental Social Psychology, 21,* 331–345.

*Bond, C. F., Jr., Lashley, B. R., & Kirk, C. T. (2002). *Responding to deception.* Unpublished manuscript, Texas Christian University, Fort Worth.

*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.

*Bond, C. F., Jr., Paulson, R. M., & Thomas, B. J. (2003). *Multiple-audience messages.* Unpublished data, Texas Christian University, Fort Worth.

Bond, C. F., Jr., & Robinson, M. (1988). The evolution of deception. *Journal of Nonverbal Behavior, 12,* 295–307.

*Bond, C. F., Jr., Thomas, B. J., & Paulsen, R. M. (2004). Maintaining lies: The multiple-audience problem. *Journal of Experimental Social Psychology, 40,* 29–40.

Bond, C. F., Jr., & Uysal, A. (2007). On lie detection "wizards." *Law and Human Behavior, 31,* 109–115.

Bond, C. F., Jr., Wiitala, W. L., & Richard, F. D. (2003). Meta-analysis of raw mean differences. *Psychological Methods, 8,* 406–418.

*Bond, G. D. (2006). *Deception detection expertise.* Unpublished manuscript, Winston-Salem State University, NC.

*Bond, G. D., Malloy, D. M., Arias, E. A., Nunn, S. N., & Thompson, L. A. (2005). Lie-biased decision making in prison. *Communication Reports, 18,* 1–11.

*Bond, G. D., Malloy, D. M., Thompson, L. A., Arias, E. A., & Nunn, S. N. (2004). Post-probe decision making in a prison context. *Communication Monographs, 71,* 269–283.

*Boone, R. T., Blumenthal, J. A., Simon, T., Cunningham, J. G., & Tucker, J. S. (1998). *Discriminating the truth: The effect of presentation, order, and gender on the detection of deception in an adversarial dispute.* Unpublished manuscript, Brandeis University, Waltham, MA.

*Brand, R. J., Hodges, S. D., & Williams, J. L. (2003). *Sex differences in encoding sexual attraction.* Paper presented at the annual meeting of the American Psychological Society, Atlanta, GA.

*Brandt, D. R., Miller, G. R., & Hocking, J. E. (1980b). Effects of self-monitoring and familiarity on deception detection. *Communication Quarterly, 28,* 3–10.

*Buchanan, J. N. (1998). *UJOS as lie detectors: Exploring university judicial officers' judgments of deception.* Unpublished doctoral dissertation, Florida State University, Tallahassee.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6,* 203–242.

*Cardena, E. (1983). *The face of deception: Perception of lying and facial expression.* Unpublished master's thesis, York University, Toronto, Ontario, Canada.

*Chahal, K., & Cassidy, T. (1995). Deception and its detection in children: A study of adult accuracy. *Psychology, Crime, & Law, 1,* 237–245.

*Cheng, H. C. (1996). *A study of baseline familiarity on deception detection: Effects of rehearsal and consequences for lying.* Unpublished master's thesis, State University of New York, Buffalo.

*Childers, C. D. (1980). *An exploration of multiple channel evaluations in attributions of deception.* Unpublished master's thesis, University of Mississippi, Oxford.

*Christensen, D. (1980). *Decoding of intended versus unintended nonverbal messages as a function of social skill and anxiety.* Unpublished doctoral dissertation, University of Connecticut, Storrs.

*Clark, L. M. (1983). *Training humans to become better decoders of deception.* Unpublished master's thesis, University of Georgia, Athens.

David, H. A., & Nagaraja, H. N. (2003). *Order statistics.* Hoboken, NJ: Wiley.

Davis, M. H., & Krauss, L. A. (1997). Personality and empathic accuracy. In W. Ickes (Ed.), *Empathic accuracy* (pp. 144–168). New York: Guilford Press.

*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. (1979). Telling lies. *Journal of Personality and Social Psychology, 37,* 1713–1722.

*deTurck, M. A. (1991). Training observers to detect spontaneous deception: Effects of gender. *Communication Reports, 4,* 81–89.

*deTurck, M. A., Feeley, T. H., & Anastasiou, L. (1997). *Effects of motivation to deceive and rehearsal on deception detection.* Unpublished manuscript.

*deTurck, M. A., Feeley, T. H., & Roman, L. A. (1997). Vocal and visual cue training in behavioral lie detection. *Communication Research Reports, 14,* 249–259.

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

*Edelstein, R. S., Luten, T. L., Ekman, P., & Goodman, G. S. (2006). Detecting lies in children and adults. *Law and Human Behavior, 30,* 1–10.

*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., & Frank, M. G. (1999). A few can catch a liar. *Psychological Science, 10,* 263–266.

*Elaad, E. (2003). Effects of feedback on the overestimated capacity to detect lies and the underestimated ability to tell lies. *Applied Cognitive Psychology, 17,* 249–263.

*Etcoff, N. L., Ekman, P., Magee, J. J., & Frank, M. G. (2000, May 11). Lie detection and language comprehension. *Nature, 405,* 139.

*Fan, R. M., Wagner, H. L., & Manstead, A. S. R. (1995). Anchoring, familiarity, and confidence in the detection of deception. *Basic and Applied Social Psychology, 17,* 83–96.

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

*Feeley, T. H., & deTurck, M. A. (1997). Case-relevant vs. case-irrelevant questioning in experimental lie detection. *Communication Reports, 10,* 35–45.

*Feeley, T. H., deTurck, M. A., & Young, M. J. (1995). Baseline familiarity in lie detection. *Communication Research Reports, 12,* 160–169.

*Fiedler, K., & Walka, I. (1993). Training lie detectors to use nonverbal cues instead of global heuristics. *Human Communication Research, 20,* 199–223.

*Fontenot, K. A. (1993). *The relationship of conversational sensitivity and employment interview experience to deception detection in employment interviews.* Unpublished doctoral dissertation, Louisiana State University, Baton Rouge.

*Forrest, J. A. (2001). *A social cognitive model of detecting deception.* Unpublished doctoral dissertation, University of Massachusetts, Boston.

*Fraidin, S. N., Hollingshead, A. B., & Kruger, J. (2002). *Effects of suspicion on perceptions, judgments, and truth bias in lie detection.* Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

*Frank, M. G. (1989). *Human lie detection ability as a function of the liar's motivation.* Unpublished doctoral dissertation, Cornell University, Ithaca, NY.

*Frank, M. G., & Ekman, P. (1997). The ability to detect deceit generalizes across different types of high-stakes lies. *Journal of Personality and Social Psychology, 72,* 1429–1439.

*Frank, M. G., & Ekman, P. (2005). Appearing truthful generalizes across different deception situations. *Journal of Personality and Social Psychology, 86,* 486–495.

*Frank, M. G., Paolantonio, N., Feeley, T. H., & Servoss, T. J. (2004). Individual and small group accuracy in judging truthful and deceptive communication. *Group Decision and Negotiation, 13,* 45–59.

*Galarza, L. (1996). *On the accuracy of detecting deception in selection interviews: The effects of applicant rehearsal, applicant job interest, and self-monitoring.* Unpublished master's thesis, Rice University, Houston, TX.

*George, J. F., Marett, K., Burgoon, J. K., Crews, J., Cao, J., Lin, M., & Biros, D. P. (2004). *Training to detect deception: An experimental investigation.* Proceedings of the 37th Hawaii International Conference on System Sciences, Big Island, HI.

Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences.* San Francisco: Freedman.

*Grandpre, J. R. (1993). *To catch a lie: An examination of the behavioral cues associated with deception detection.* Unpublished master's thesis, University of Wyoming, Laramie.

Granhag, P. A., & Strömwall, L. A. (2004). *The detection of deception in forensic contexts.* New York: Cambridge University Press.

Hall, J. A. (1984). *Nonverbal sex differences: Communication accuracy and expressive style.* Baltimore: Johns Hopkins.

Hall, J. A. (2001). The PONS test and the psychometric approach to measuring interpersonal sensitivity. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 143–182). Mahwah, NJ: Erlbaum.

*Hall, S. (1989). *The generalizability of learning to detect deception in effective and ineffective deceivers.* Unpublished doctoral thesis, Auburn University, AL.

*Harrison, A. A., Hwalek, M., Raney, D., & Fritz, J. G. (1978). Cues to deception in an interview situation. *Social Psychology, 41,* 156–161.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3,* 486–504.

*Hemsley, G. D. (1977). *Experimental studies in the behavioral indicants of deception.* Unpublished doctoral dissertation, University of Toronto, Ontario, Canada.

*Hendershot, J. (1981). *Detection of deception in low and high socialization subjects with trained and untrained judges.* Unpublished master's thesis, Auburn University, AL.

*Hocking, J. E., Bauchner, J., Kaminski, E. P., & Miller, G. R. (1979). Detecting deceptive communication from verbal, visual, and paralinguistic cues. *Human Communication Research, 6,* 33–46.

*Horn, D. B. (2001). *Seeing is believing: Video quality and lie detection.* Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Howell, D. C. (2006). *Statistical methods for psychology* (6th ed.). Belmont, CA: Wadsworth.

Hubbell, A. P., Mitchell, M. M., & Gee, J. C. (2001). The relative effects of timing of suspicion and outcome involvement on biased message processing. *Communication Monographs, 68,* 115–132.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Beverly Hills, CA: Sage.

*Hurd, K., & Noller, P. (1988). Decoding deception: A look at the process. *Journal of Nonverbal Behavior, 12,* 217–233.

*Johnson, A. K., Barnacz, A., Yokkaichi, T., Rubio, J., Racioppi, C., Shackelford, T. K., et al. (2005). Me, myself, and lie: The role of self-awareness in deception. *Personality and Individual Differences, 38,* 1847–1853.

*Kassin, S. M., & Fong, C. T. (1999). "I'm innocent!": Effects of training on judgments of truth and deception in the interrogation room. *Law & Human Behavior, 23,* 499–516.

*Kassin, S. M., Meissner, C. A., & Norwick, R. J. (2005). "I'd know a false confession if I saw one": A comparative study of college students and police investigators. *Law & Human Behavior, 29,* 211–227.

*Keens, C. H. W., & Broadhurst, R. (2005). The detection of deception: The effects of first and second language on lie detection ability. *Psychiatry, Psychology, and Law, 12,* 107–118.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis.* New York: Guilford Press.

Kenny, D. A., & LaVoie, L. (1984). The social relations model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 141–182). New York: Academic Press.

*Kirk, C. T. (1993). *A nonspecific factor in deception judgments.* Unpublished master's thesis, Texas Christian University, Fort Worth.

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

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

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

*Landry, K. L., & Brigham, J. C. (1992). The effect of training in criteria-based content analysis on the ability to detect deception in adults. *Law & Human Behavior, 16,* 663–676.

*Lane, J. D., & DePaulo, B. M. (1999). Completing Coyne's cycle: Dysphorics' ability to detect deception. *Journal of Research in Personality, 33,* 311–329.

*Lashley, B. R. (1993). *Anxiety cues versus verbal–nonverbal inconsistency: Two factors in deception judgments.* Unpublished master's thesis, Texas Christian University, Fort Worth.

*Lavrakas, P. J., & Maier, R. A. (1979). Differences in human ability to judge veracity from the audio medium. *Journal of Research in Personality, 13,* 139–153.

*Leach, A., Talwar, V., Lee, K., Balan, N., & Lindsay, R. C. L. (2004). "Intuitive" lie detection of children's deception by law enforcement officials and university students. *Law & Human Behavior, 28,* 661–685.

*Levine, T. R., Feeley, T. H., McCornack, S. A., Hughes, M., & Harms, C. M. (2005). Testing the effects of nonverbal behavior training on accuracy in deception detection with the inclusion of bogus training control group. *Western Journal of Communication, 69,* 203–218.

*Levine, T. R., Kim, R. K., Park, H. S., & Hughes, M. (2006). Deception detection accuracy is a predictable linear function message veracity base-rate: A formal test of Park and Levine's probability model. *Communication Monographs, 73,* 243–260.

Levine, T. R., & McCornack, S. A. (1991). The dark side of trust: Conceptualizing and measuring types of communication suspicion. *Communication Quarterly, 39,* 325–340.

*Levine, T. R., & McCornack, S. A. (2001). Behavioral adaptation, confidence, and heuristic-based explanations of the probing effect. *Human Communication Research, 27,* 471–502.

*Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect." *Communication Monographs, 66,* 125–144.

*Lin, Y. C. (1999). *A study of training on deception detection: The effects of the specific six cues versus heuristics on deception detection accuracy.* Unpublished master's thesis, State University of New York, Buffalo.

*Littlepage, G. E., Maddox, J., & Pineault, M. A. (1985). Recognition of discrepant nonverbal messages and detection of deception. *Perceptual and Motor Skills, 60,* 119–124.

*Littlepage, G. E., McKinnie, R., & Pineault, M. A. (1983). Relationship between nonverbal sensitivities and detection of deception. *Perceptual and Motor Skills, 57,* 651–657.

*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.

*Littlepage, G. E., & Pineault, M. A. (1981). Detection of truthful and deceptive interpersonal communications across information transmission modes. *Journal of Social Psychology, 114,* 57–68.

*Littlepage, G. E., & Pineault, M. A. (1985). Detection of deception of planned versus spontaneous communications. *Journal of Social Psychology, 125,* 195–201.

*Littlepage, G. E., Tang, D. W., & Pineault, M. A. (1986). Nonverbal and content factors in the detection of deception in planned and spontaneous communications. *Journal of Social Behavior and Personality, 1,* 439–450.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

*Lusby, D. J. (1999). *Effects of motivation to lie and sanctioned deception of the accuracy of observers' veracity judgments.* Unpublished master's thesis, State University of New York, Buffalo.

*Malcolm, S. R., & Keenan, J. P. (2005). Hemispheric asymmetry and deception detection. *Laterality, 10,* 103–110.

Malone, B., & DePaulo, B. M. (2001). Measuring sensitivity to deception. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 103–124). Mahwah, NJ: Erlbaum.

*Mann, S., & Vrij, A. (2006). Police officers' judgments of veracity, tenseness, cognitive load, and attempted behavioural control in real-life police interviews. *Psychology, Crime, & Law, 12,* 307–319.

*Mann, S., Vrij, A., & Bull, R. (2004). Detecting true lies: Police officers' ability to detect suspects' lies. *Journal of Applied Psychology, 89,* 137–149.

*Marchioni, P. M. (1980). *The detection of deception in the body and the face.* Unpublished master's thesis, University of Texas at El Paso.

*Masip, J. (2002). *Credibility assessment of the testimony on the basis of behavioral indicators in criminal justice settings.* Unpublished doctoral dissertation, University of Salamanca, Madrid, Spain.

*Masip, J., Garrido, E., & Herrero, C. (2002). *When did you conclude she was lying? The impact of the moment the deception about the sender's veracity is made and the sender's facial appearance on police officers' credibility judgments.* Manuscript submitted for publication.

Masip, J., Garrido, E., & Herrero, C. (2004). Facial appearance and impressions of credibility: The effects of facial babyishness and age on person perception. *International Journal of Psychology, 39,* 276–289.

*Masip, J., Garrido, E., & Herrero, C. (2006a). *Heuristic versus systematic processing of information in detecting deception.* Unpublished manuscript, University of Salamanca, Madrid, Spain.

*Masip, J., Garrido, E., & Herrero, C. (2006b). Observers' decision moment in deception detection experiments: Its impact on judgment, accuracy, and confidence. *International Journal of Psychology, 41,* 304–319.

*Meissner, C. A., & Kassin, S. M. (2002). "He's guilty": Investigator bias and judgments of truth and deception. *Law & Human Behavior, 26,* 469–480.

*Millar, M. G., & Millar, K. U. (1997a). The effects of cognitive capacity and suspicion on truth bias. *Communication Research, 24,* 556–570.

*Millar, M. G., & Millar, K. (1997b). Effects of situational variables on judgments about deception and detection accuracy. *Basic and Applied Social Psychology, 19,* 401–410.

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

*Morris, K. A. (2003). Teaching students about classic findings on the detection of deception. *Teaching of Psychology, 30,* 111–113.

*Murray, J. (1983). *The detection of interviewees' verbal deceptions from their accompanying overt nonverbal behavior.* Unpublished doctoral dissertation, Florida State University, Tallahassee.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

*Oberleitner, D., McLarney-Vesotski, A., Bernieri, F., & Okdie, B. (2004). *Musical mood induction and its effects on lie detection accuracy.* Paper presented at the annual meeting of the Society for Personality and Social Psychology, Austin, TX.

*Oldfield, S. E. (1999). *The influence of self-monitoring on the detection of deception.* Unpublished master's thesis, Radford University, VA.

*O'Sullivan, M. (2003). The fundamental attribution error in detecting deception: The boy-who-cried-wolf effect. *Personality and Social Psychology Bulletin, 29,* 1316–1327.

O'Sullivan, M. (2007). Unicorns or Tiger Woods: Are lie detection wizards myths or rarities? A response to On Lie Detection "Wizards" by Bond and Uysal. *Law & Human Behavior, 31,* 117–123.

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

*Park, E. S., Levine, T. R., Harms, C. M., & Ferrara, M. H. (2002). Group and individual accuracy in deception detection. *Communication Research Reports, 19,* 99–106.

Park, H. S., Levine, T. R., McCornack, S. A., Morrison, K., & Ferrara, S. (2002). How people really detect lies. *Communication Monographs, 69,* 144–157.

*Parker, R. J. (1978). *Age, sex, and the detection of deception through nonverbal cues.* Unpublished doctoral dissertation, California School of Professional Psychology.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp.17–59). New York: Academic Press.

*Pietras, K. R., & Bond, C. F., Jr. (1985). *The reliability of lie detection judgments.* Paper presented at the annual meeting of the Eastern Psychological Association, Boston.

*Porter, S., Campbell, M. A., Stapleton, J., & Birt, A. R. (2002). The influence of judge, target, and stimulus characteristics on the accuracy of detecting deceit. *Canadian Journal of Behavioural Science, 34,* 172–185.

*Porter, S., McCabe, S., Woodworth, M., & Peace, K. A. (2006). "Genius is 1% inspiration and 99% perspiration," or is it? An investigation of the impact of motivation and feedback on deception detection. *Legal and Criminological Psychology, 12,* 297–309.

*Porter, S., Woodworth, M., & Birt, A. R. (2000). Truth, lies, and videotape: An investigation of the ability of federal parole officers to detect deception. *Law & Human Behavior, 24,* 643–658.

*Raichle, W. (1990). *The roles of verbal and nonverbal behavior in impression formation during an employment interview.* Unpublished doctoral dissertation, New York University.

*Rockwell, P. (1996). Hemispheric differences in detection of deception with content-filtered speech. *Perceptual and Motor Skills, 82,* 1241–1242.

*Rudin, J. P. (1999). *Effect of reward size in performance of a deception detection task.* Proceedings of the annual meeting of the Southwest Academy of Management, Houston, TX.

*Sahlman, J. M., & Koper, R. J. (1992, May). *Do you hear what I hear? Deception detection by the blind.* Paper presented at the annual meeting of the International Communication Association, Miami, FL.

*Sakai, D. J. (1981). *Nonverbal communication in the detection of deception among women and men.* Unpublished doctoral dissertation, University of California, Davis.

*Schoephoerster, B. T. (1996). *Deception detection accuracy: The effects of suspicion and antisocial personality traits.* Unpublished master's thesis, University of Nevada at Las Vegas.

*Seager, P. B. (2001). *Improving the ability of people to detect lies.* Unpublished doctoral dissertation, University of Hertfordshire, United Kingdom.

*Steinberg, J. R., & Bernieri, F. (2001). *How locus of control affects accuracy of detecting deception.* Paper presented at the annual meeting of the American Psychological Society, Toronto, Ontario, Canada.

*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.

*Swann, W. B., Jr., Silvera, D. H., & Proske, C. U. (1995). On "knowing your partner": Dangerous illusions in the age of AIDS? *Personal Relationships, 2,* 173–186.

*Swinkels, A. H. (1989). *The effects of cognitive busyness on human lie detection ability.* Unpublished doctoral dissertation, University of Texas at Austin.

*Taylor, R. (1999). *Relationships between accuracy of detecting deception and beliefs about the cues involved.* Paper presented at the Division of Forensic Psychology annual conference, Cambridge, England.

*Tetterton, V. S., & Warren, A. R. (2005). Using witness confidence can impair the ability to detect deception. *Criminal Justice and Behavior, 32,* 433–451.

*Tornqvist, J. S. (2002). *The detectability and believability of lies and truths in the e-mail medium.* Unpublished doctoral dissertation, University of Virginia, Charlottesville.

*Vrij, A. (1993). Credibility judgments of detectives: The impact of nonverbal behavior, social skills, and physical characteristics on impression formation. *Journal of Social Psychology, 133,* 601–610.

*Vrij, A. (1994). The impact of information and setting on detection of deception by police detectives. *Journal of Nonverbal Behavior, 18,* 117–136.

Vrij, A. (2000). *Detecting lies and deceit: The psychology of lying and the implications for professional practice.* New York: Wiley.

*Vrij, A., & Baxter, M. (1999). Accuracy and confidence in detecting truths and lies in elaborations and denials: Truth bias, lie bias, and individual differences. *Expert Evidence, 7,* 25–36.

*Vrij, A., & Graham, S. (1997). Individual differences between liars and the ability to detect lies. *Expert Evidence, 5,* 144–148.

*Vrij, A., Harden, F., Terry, J., Edward, K., & Bull, R. (2001). The influence of personal characteristics, stakes, and lie complexity on the accuracy and confidence to detect deceit. In R. Roesch, R. R. Corrado, & R. Dempster (Eds.), *Psychology in the courts* (pp. 289–302). London: Routledge.

*Vrij, A., & Mann, S. (2001). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. *Applied Cognitive Psychology, 15,* 187–203.

*Vrij, A., Mann, S., Kristen, S., & Fisher, R. P. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law & Human Behavior, 31,* 499–518.

*Vrij, A., Mann, S., Robbins, E., & Robinson, M. (2006). Police officers ability to detect deception in high stakes situations and in repeated lie detection tests. *Applied Cognitive Psychology, 20,* 741–755.

*Wilson, S. J. (1975). *Channel differences in the detection of deception.* Unpublished doctoral dissertation, Florida State University, Tallahassee.

*Yang, C. C. (1996). *The effects of training, rehearsal, and consequences for lying on deception detection accuracy.* Unpublished master's thesis, State University of New York, Buffalo.

Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). "Wide-eyed" and "crooked-face": Determinants of perceived and real honesty over the life span. *Personality and Social Psychology Bulletin, 22,* 1258–1269.

*Zuckerman, M., DeFrank, R. S., Hall, J. A., Larrance, D. T., & Rosenthal, R. (1979). Facial and vocal cues of deception and honesty. *Journal of Experimental Social Psychology, 15,* 378–396.

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. 1–60). New York: Academic Press.

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

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

*Zuckerman, M., Koestner, R., Colella, M. J., & Alton, A. O. (1984). Anchoring in the detection of deception and leakage. *Journal of Personality and Social Psychology, 47,* 301–311.

Received March 12, 2007
Revision received November 15, 2007
Accepted November 16, 2007 ■