Chapter 10
Is *Le Mot Juste*? The Contexualization of Words by Expert Lie Detectors

Maureen O'Sullivan

Some of the chapters in this book discuss the ways in which language samples can be analyzed to determine credibility (e.g., Colwell, Hiscock-Anisman, & Fede, this volume; Griesel, Ternes, Schraml, Cooper, & Yuille, this volume). This chapter illustrates how expert lie detectors use information from a single word in discerning the truthfulness of others. These illustrations were obtained from in-depth interviews with highly accurate lie detectors (O'Sullivan & Ekman, 2004) who received scores of 80% or more on at least two of three different lie detection tests. The three tests were not easy, since average scores on the measures are close to 50%. Although the base rate occurrence of such expert lie detectors varies from group to group, the expert lie detectors in this analysis are at least two standard deviations above the mean in their lie detection abilities. For example, using the criterion described, no expert lie detector has been found among college students, although thousands have been examined. Although there are now a sufficient number of experts ($n = 50$) to aggregate their responses and compare them with their matched controls, another value of the project[1] is the opportunity to compare the description of the lie detection enterprise that results from the efforts of a single expert with the contributions to knowledge made by scores of scientists using a wide variety of methodologies. A brief review of these methodologies is offered in order to situate the kind of information obtained from individual interviews in the broader research endeavor. This review is, of necessity, cursory. Many subtle distinctions are disregarded in the effort to describe briefly each approach.

[1] Editorial note: "The project" refers to the "Truth Wizard Project"- O'Sullivan and Ekman's research project that sought to identify expert lie detectors, who obtained highly accurate scores on at least two of three videotaped lie detection tasks.

M. O'Sullivan (✉)
University of San Francisco, San Francisco, CA, USA

B.S. Cooper et al. (eds.), *Applied Issues in Investigative Interviewing, Eyewitness Memory, and Credibility Assessment*, DOI 10.1007/978-1-4614-5547-9_10,

Five Ways in Which the Contribution of Language to Lie Detection has Been Studied

Some research has focused on the content of speech to determine whether language alone can provide clues to deception, or, alternately, to assess the relative importance of language in the detection of deception. At least five different methods of addressing these questions can be discriminated: (1) language analysis of honest and deceptive verbal content alone (e.g., based on transcripts or other written documents); (2) behavioral measurement of verbal and nonverbal behaviors in honest and deceptive videotaped or audiotaped materials; (3) comparison of communication channels (e.g., verbal or nonverbal) to determine which is more effective in accurate lie detection; (4) statistical models of how individuals use different clues in making summary judgments of honest and deceptive individuals; and (5) soliciting and analyzing the reasons people give for deciding that someone is lying or telling the truth. A variety of approaches may be further distinguished within each of these five research paradigms but, except for the last paradigm (i.e., soliciting reasons for the truth vs. lie decision), those distinctions are ignored in the present chapter.

(1) Verbal Content Analysis

Within the research tradition of language analysis based on written materials, several different approaches to the analysis of an entire statement have been used, for example, Criteria-Based Content Analysis (Porter & Yuille, 1995; Steller & Koehnken, 1989; Undeutsch, 1982) and Reality Monitoring (Masip, Sporer, Garrido, & Herrero, 2005). Although each of these methods varies in both the speech qualities it deems to be most important and the objectivity with which the analysis of the materials can be completed, they share the belief that a substantial amount of language is necessary to assess credibility based on factors such as the immediacy of the language used, the quantity and quality of details provided, the consistency or coherence of the account, and its spontaneity (see Colwell et al., this volume; Griesel et al., this volume). These judgments are usually global, based on a complete story or account.

A somewhat different language-only approach is provided by Pennebaker and his colleagues (Newman, Pennebaker, Berry, & Richards, 2003). Their approach proceeds from the premise that word counts alone can provide a means of differentiating honest and lying communications (also see Hancock & Woodworth, this volume). Unlike the other theories mentioned above, in which the meaning of the statement and the overall coherence or consistency of the story is central to the final determination of truthfulness, the Pennebaker approach proposes that a sufficiently sophisticated analysis of parts of speech and combinations of words can provide a competing method of language-only lie detection. Computer-generated word counts, independent of the overall content of the story or other written communication, and various statistical models such as logistic regression and Latent Semantic Analysis (Campbell

& Pennebaker, 2003) are used to differentiate lying and truthful and more or less traumatic communications (Cohn, Mehl, & Pennebaker, 2004). The appeal of the method is its efficiency (e.g., computer counts rather than people assessment) and the counter-intuitive use of pronouns and other parts of speech to distinguish communications varying in truthfulness and/or emotionality. For example, although different kinds of lies resulted in somewhat different language styles, across five studies, "… deceptive communications were characterized by fewer first-person singular pronouns, fewer third person pronouns, more negative emotion words, fewer exclusive words, and more motion verbs" (Newman et al., 2003, p. 670).

Obviously, the use of fewer first person singular pronouns and fewer motion words is consistent with the lessened immediacy of a statement suggested by the Undeutsch Hypothesis[2] (Undeutsch, 1982) and Criteria-Based Content Analysis (Yuille, 1989). What differs is the method of obtaining this information and the loss of the "gist" of the communication.

Although most language-only techniques suggest that both the individual words used and the context within which they occur contain information that can be used to distinguish truth and deception, they differ in terms of the source of the materials they analyze. While many language-only assessment methods attempt to determine the credibility of actual suspect or witness statements made for forensic purposes, the Pennebaker group has tended to use truthful and deceptive materials produced in the laboratory or obtained from non-forensic sources. There is some evidence that different clues are available in different kinds of lies: high stakes vs. low stakes (Ekman, 2001), emotional vs. nonemotional (Warren, 2007), lies about facts vs. lies about feelings (O'Sullivan, 2008), sanctioned vs. unsanctioned lies (Feeley & deTurck, 1998), more vs. less complex lies (Vrij & Heaven, 1999), as well as the relationship between the liar and lie catcher (Burgoon, Buller, White, Afifi, & Buslig, 1999; Ekman, 2001). Differences in these variables as well as interest in and experience with those kinds of lies may also affect the accuracy of those seeking to uncover them (see ten Brinke & Porter, this volume).

(2) Behavioral Measurement of Verbal and Nonverbal Behaviors

Although no researcher interested in the accurate detection of deception would disregard the importance of verbal clues, some researchers have thought a more complete picture of honest vs. deceptive communication results from the simultaneous analysis of both verbal and nonverbal behaviors. So, a second approach to the question of the relative importance of language in deception detection can be seen in the scores of articles in which researchers have actually measured both nonverbal behaviors such as facial expressions, hand gestures, and body postures (Ekman &

[2] Editorial note: The Undeutsch hypothesis states that statements based on experienced events differ in quantity and quality from fictitious accounts.

Friesen, 1969; Ekman, Friesen, & O'Sullivan, 1988; Granhag & Strömwall, 2002; Porter, Doucette, Woodworth, Earle, & MacNeil, 2008) as well as vocal quality (e.g., pitch; Ekman, O'Sullivan, Friesen, & Scherer, 1991; Rockwell, Buller, & Burgoon, 1997) and verbal characteristics (e.g., number of words, content, detail; Kraut, 1978; Vrij, Akehurst, Soukara, & Bull, 2006). This literature has produced many widely varying conclusions. The reasons for these inconsistencies are many. The kinds of lies sampled reflect the entire continuum of ecological validity. Some of the lies are high stakes, causing significant behavioral changes in the liars and truth tellers. Others are trivial lies of politeness or courtesy which may have insufficient emotional or cognitive arousal to result in behavior. The liars and truth tellers studied include paroled felons (Bond, 2008), convicted murderers, or crime suspects (Vrij & Mann, 2001), while others use college students or children (Feldman & Jenkins, 1979; Vrij et al., 2006). The degree of sophistication and/or objectivity of the behavioral measures used are also highly variable. Some researchers merely ask observers to make Likert ratings as to whether a particular behavior, such as a smile, has occurred. Other researchers count the frequency of occurrence of any smile-like behavior while, still others, use muscle movement coding systems (Ekman & Friesen, 1978) that distinguish whether or not a very subtle movement has occurred around the eye (i.e., as a result of the *orbicularis oculi* muscle firing) simultaneously with the movement of the smiling typical lip corner raise occasioned by the action of the *zygomatic major*.

Since 2000, several meta-analyses of these studies have been provided (Aamodt & Custer, 2006; DePaulo et al., 2003; Sporer & Schwandt, 2006; Sporer & Schwandt, 2007). These summaries are very useful compendia, but they provide little but a rough guide to the relevant verbal and nonverbal behaviors that distinguish honest and deceptive behavior. By summing over such disparate methodologies, subtle differences that are very useful in real-world interviewing and in real-world assessments of the honesty of a particular individual may be lost. Nonetheless, despite the confusing variation provided by a plethora of lie types, subjects, and measurement methods, DePaulo et al. reported many variables which had significant *d*'s (i.e., a behavior discriminated honest and deceptive samples significantly, either across several studies or so strongly in a single study that its effect was not eradicated in the meta-analysis). Even examining only those effect sizes (i.e., *d*'s) of 0.50 or above (i.e., consistent with a moderate effect), objectively measured verbal and vocal behaviors were identified that consistently differentiated honest and deceptive samples. The value of such meta-analyses is that they demonstrate the replicability of clues, both verbal and nonverbal, across many kinds of lie detection materials. They also demonstrate that at least some lie detection materials have significant clues to deception in them.

The above is important because a continuing bias in the field of lie detection accuracy research is the lack of lie detection accuracy of most of the subjects studied. Bond and DePaulo (2006) reported a mean accuracy of 54.3% over 20,000 subjects and, although they summarized that result as "mean lie-truth discrimination abilities are nontrivial, with … a *d* of roughly 0.40 … an effect that is at roughly the 60th percentile in size, relative to others that have been meta-analyzed by social

psychologists" (p. 214), they then went on to claim that, although there is evidence of reliable observer truth bias and target demeanor credibility, there is no evidence for any lie detection accuracy (Bond & DePaulo, 2008).[3] O'Sullivan (2008), however, questioned their conclusions on theoretical, methodological, and empirical grounds.

Ekman (2001) has long argued that many lie detection scenarios used in determining accuracy are too low stakes to provide the cognitive and/or emotional clues necessary to produce discernible clues to deception. O'Sullivan (2008) reiterated this observation and argued that, in addition, most of the 20,000 subjects surveyed in the aforementioned meta-analysis were college students with little life experience, feedback, or motivation to support accurate lie detection in the low stakes lies provided in most studies.

(3) Modality Dissection

In the studies just reviewed, the relative importance of different kinds of verbal and nonverbal behaviors was examined through a direct measurement of the liars' and truth tellers' behaviors and a frequency count of whether a particular kind of behavior occurred or not was provided. Of course, only those behaviors which researchers chose to measure were included in the analyses. Behavioral measurement, of necessity, reflects the interests and expertise of the people doing the often costly and always tedious behavioral analyses. But whether real-life observers actually attend to the clues that researchers so laboriously assess is another question. The naive observer and the sophisticated scientist may have non-overlapping sets of clues to which they attend. A cognitive scientist, steeped in the knowledge of the fallibility of human memory, will judge admitted lack of memory as more believable than claimed total recall; the untutored observer may come to the opposite conclusion. Similarly, a facial expression expert might use rapidly occurring signs of emotion, i.e., microexpressions (Ekman, 2003), as clues to suppressed or repressed emotion that might be related to lying, but most observers may neither perceive such clues nor be able to interpret them accurately.

This difficulty was addressed in a third type of analysis which sought to determine the relative importance of one kind of communication compared with another by limiting the information that observers are given and examining the accuracy of the judgments they make under each of the different viewing or listening conditions. DePaulo, Rosenthal, Eisenstat, Rogers, and Finkestein (1978), for example, showed observers lie detection scenarios in three formats: video alone, audio alone, or combined audio and video information (i.e., the usual audiovisual format). They found

[3] Editorial note: Bond and DePaulo (2008) suggest that there is very little variation in individuals' ability to detect deception, that detection accuracy ranges no more widely than would be expected by chance, and that the most accurate judges are no more accurate than a stochastic mechanism would produce.

a video primacy effect, with those judgments based on the video only being more accurate than those made from the audio material. DePaulo, Lanier, and Davis (1983), however, showed observers verbal only (i.e., transcript), audio only (i.e., verbal and audio), visual only, and the complete audiovisual recording. Lies told by more motivated senders were more readily detected with more information, that is, in the audio and the audiovisual conditions. Visual only and verbal only conditions were not associated with deception detectability.

A variant of the above paradigm is the attempt to study experimentally the discrepancy between different channels of communication. The overall lack of coherence or consistency of a statement is a frequently reported characteristic of statements judged to be less credible in the forensic context (see Connolly & Price, this volume; Fisher, Vrij, & Leins, this volume). In the nonverbal area, many theorists (Bugental, Kaswan, & Love, 1970) have suggested that the perception of verbal/nonverbal discrepancies is one of the more important clues in accurate lie detection. As is discussed at the end of this chapter, that is certainly an important characteristic of the lie detection strategies of the "truth wizards," but they are sensitive not only to verbal/nonverbal discrepancies, but also to discrepancies concerning demographic characteristics of the liar or truth teller—age, race, social class, gender, personality type, interpersonal style, and many other variables that differ among individuals.

(4) Processes Involved in Judging Others as Deceptive

O'Sullivan (2005) has argued that discerning the truthfulness of others is a particular example of the more general ability referred to as empathic accuracy (Ickes, 1993), understanding others (Funder, 1999), social-emotional intelligence (Mayer, Salovey, & Caruso, 2002), interpersonal and intrapersonal intelligence (Gardner, 1993) and behavioral cognition (O'Sullivan & Guilford, 1975). In addition to the differences in variance provided by the different kinds of lies examined (discussed above), the processes involved in accurate lie detection probably involve more than merely the perception and interpretation of lie-related clues. The finding (Warren, 2007) that individuals with greater sensitivity to subtle facial expressions of emotion are more accurate in detecting emotional lies but not non-emotional ones is consistent with this view. Similarly, O'Sullivan reports (2008) that, among expert lie detectors, police professionals are significantly better than therapists in detecting lies about a theft, but significantly less accurate in detecting lies about feelings (see Table 10.1).

Years of research on social cognition (Fiske, 1992) suggests that judgments of honesty or deceptiveness, like all judgments made under uncertainty, will be characterized by the cognitive biases and heuristics that mark other kinds of social assessments. A well-known bias in lie detection studies is the truth bias (Zuckerman, DeFrank, Hall, Larrance, & Rosenthal, 1979) in which observers have reliable tendencies (Bond & DePaulo, 2008) to call people honest, regardless of the base rate

Table 10.1 Lie detection accuracy in percentages for "truth wizard" police professionals and therapists

Expert	N	Lie scenario					
Group		Opinion		Crime		Emotion	
		Mean	S.D.	Mean	S.D.	Mean	S.D.
Police	18	88	3	88	8	69	15
Therapists/Psychologists	10	89	.3	61	15	82	12

or actual honesty of the people they are judging (i.e., even if told that about half the people they will see are lying, many judges will rate 70% or more of the targets as honest). Ekman (2001) and Meissner and Kassin (2002) have reported a deceptive bias among police professionals, in which their tendency is to presume that those they are interviewing are deceptive, even when they have been instructed that is not the case (see Yarbrough, Hervé, & Harms, this volume).

The mere act of labeling someone as a liar also seems to be problematic, although this "accusatory reluctance" (O'Sullivan, 2003) may be circumvented by changes in how the judgment of honesty is obtained. DePaulo (1998) found that observers, who were only at chance in labeling targets as deceptive, obtained significantly higher accuracy rates if they were asked to characterize the targets as comfortable or uncomfortable. Similarly, Mann and Vrij (2006) found that observers were more accurate if they were asked to categorize liars and truth tellers as "thinking hard" rather than "lying vs. truthful."

O'Sullivan (2003) demonstrated that observers' fundamental attribution (Ross & Nisbett, 1991) about someone based on a 1-sec still photograph of him was significantly correlated with their later judgment of that man's truthfulness in a 1-min interview. She called this the "boy-who-cried-wolf" effect because the effect was much stronger for those initially judged as untrustworthy. Although observers would sometimes judge a person rated as trustworthy on the basis of a photograph as lying in an interview, they rarely rated someone they thought to be an untrustworthy person as telling the truth.

These above noted studies suggest that the processes involved in judging whether someone is lying or telling the truth are complex ones and only a little work has been done to untangle the processes involved. Two different approaches have been reported.

Ekman, Friesen, O'Sullivan, and Scherer (1980) obtained personality/trait judgments of young women videotaped describing their feelings truthfully or lying about them. In earlier research, it had been found that observers were only at chance in their judgments of the women's veracity. Various groups of observers were shown the same 15 women in different formats: speech only, face only, body only and total audio visual recording. About half of the women were lying, although the observers were not told that deception was involved. Each of the women was rated on 14 seven-point bipolar scales such as outgoing-inhibited and calm-agitated. The ratings of the observer group which saw and heard the entire audiovisual record was used as the criterion and the ratings of the other three groups were regressed against

them. These multiple regression analyses were conducted separately for the truthful nurses and the deceptive ones. Two different judgment strategies were found. If the target nurses were telling the truth, different channels were more highly correlated with the total criterion judgment depending on what the trait was. For example, in the honest condition, ratings from the face-only and speech-only conditions were both correlated with total judgments of outgoingness, sociability and expressivity. Body ratings were more often correlated with total judgments of calmness, stability and relaxation. However, these same judges, when judging the deceptive nurses, tended to use only the verbal channel. That is, the ratings made on the basis of the voice only were most highly correlated with the total audio visual criterion (note: the same judges made the ratings of both the honest and deceptive nurses).

The above study presented the entire audio channel to the observers, so it was not clear whether judges were attending to the content of speech or vocal quality, or both. In a follow-up study, O'Sullivan, Ekman, Friesen, and Scherer (1985) had one group of judges read the transcripts of the interviews; a second group heard an audio tape which had been content filtered. This retained the rhythm and cadence of speech, but removed the meaning of the words spoken. The ratings of the full speech group from the earlier study were used as the criterion, and the ratings based on the transcript and the content-filtered speech was used as the predictors. When the women were lying, ratings based on the transcript alone were significantly more highly correlated with the ratings based on the complete speech recording than those made based on content filtered speech. In the honest interviews (i.e., with the same women), observers showed the opposite pattern, attending significantly more frequently to vocal quality.

These two noted studies are important because they suggest that, although observers are loath to label someone as a liar, they process the behavioral and verbal information produced by liars and truth tellers differently. And despite the many studies and folk wisdom about the importance of attending to discrepancies in information, the results of these studies are consistent with the social cognition finding that most people are cognitively lazy (Fiske, 1992) and may attempt to solve discrepancies by attending to the channel for which the target is most responsible—her words.

Heinrich and Borkenau (1998) proceeded from a different premise in attempting to understand the cognitive strategies used by lie catchers in understanding others. They hypothesized that the human default option in social judgments is to assess the overall character of an individual. They demonstrated that ratings of the Agreeableness of the individual (i.e., measured by a four item scale based on the Big Five model of personality, and containing a rating on scrupulousness vs. unscrupulousness) were significantly correlated with ratings of deceptiveness, but not with other personality factors such as Extraversion, Neuroticism and Openness. They argued that overall judgments of character are more predictive of more behaviors than particularized truthfulness assessments (also see ten Brinke & Porter, this volume). Their results are consistent with those reported by O'Sullivan (2003) but, while they emphasized the observer's assessment of personality characteristics, O'Sullivan conceptualized this relationship between trait judgments of trustworthiness and state judgments of honesty as an example of the fundamental attribution error.

(5) Beliefs About Clues to Deception

A fifth type of study has tried to evaluate the relative importance of verbal and nonverbal behaviors by asking people to report what they believe the clues to lying to be or to describe how they arrived at their judgment of someone as truthful or deceptive. The kinds of reasons solicited, and the kinds of studies used to examine them can be categorized into four groups: The General Belief Group, The Expert Belief Group, The Free Response to Item Group and the Personal Remembered Lie Group.

The General Belief Group includes early work by Zuckerman and Driver (1985) and more recent work by Bond and Rao (2004) who asked individuals what they believe to be important clues to deception. This literature suggests that eye gaze aversion is widely believed to be a useful clue to deception, although laboratory evidence suggests that some liars actually increase their eye gaze when lying (Zuckerman & Driver). DePaulo et al. (2003) concluded that increased eye gaze aversion may occur when lying about something one is ashamed of, but not in other circumstances. Park, Levine, Harms, and, Ferrara (2002) had students rate the importance of 11 behaviors on a seven-point scale. The behaviors included eye contact; plausibility or consistency of verbal statements; body movements such as fidgeting, speech fluency or disfluency, "intuition" or "gut feeling"; consistency of verbal statements; nervous nonverbal behaviors (i.e., other than body movements); random guessing; facial expressions; consistency of nonverbal behavior; and content of verbal statements. Unlike the earlier Zuckerman and DePaulo studies, the Park et al. students reported speech fluency and the plausibility and content of verbal statements to be the more important behaviors in detecting deception.

A second group of studies (i.e., The Expert Belief Group) compared the beliefs about clues to deception of various lie detector professionals. Vrij and Semin (1996) contrasted the ratings made by prisoners, students and police professionals of the importance of 16 behaviors studied previously or reported in the experimental literature. These included: gaze behavior, smiling, head movements, trunk movements, shifting positions, foot/leg movements, gestures, self-touches, hand and finger movements, shoulder shrugs, response length, speech rate, latency period, ah-filled pauses, non-ah speech disturbances and pitch of voice. Although differences were found between students and police professionals in gestures and shoulder shrugs, with police officers finding them more important than the college students did, the more significant differences were between the prisoners and the other two groups with the prisoners reporting many more behaviors to be unimportant in lie detection.

Strömwell and Granhag (2003) asked judges, prosecutors and police about the importance to lie detection of two verbal clues to deception (i.e., number of details and consistency) and three nonverbal behaviors (i.e., gaze aversion, pitch of voice and body movements) using the same four-point scale as Vrij and Semin (1996). They found that judges believed that verbal content clues were more reliable than nonverbal clues significantly more frequently than police officers did. Police were more likely

to view nonverbal behavior as more reliable. Hartwig, Granhag, Strömwall, and Vrij (2004) contrasted the ratings given by police officers following an interrogation they performed vs. one they watched on videotape. Seven kinds of verbal clues were rated: completeness of the statement, confidence in the statement, consistency, details, plausibility, whether the story seemed rehearsed, and general clarity of the statement. The nonverbal clues assessed were: body movements, general behavioral trustworthiness, gaze aversion and nervousness. When rating the use of verbal vs. nonverbal cues, police rated verbal cue usage as significantly greater. When the percentage of time that individual police interviewers or observers actually used any of the seven different cues was used as a measure, a somewhat different pattern was found. General behavioral credibility was cited 24% of the time by interrogators and 16% of the time by observers. Observers cited body movements 20% of the time and general nervousness 15% of the time. Statement plausibility and general statement credibility were each cited 12% of the time. The other cues were all mentioned fewer than 9% of the time. A limitation of all these studies, whether using students or lie detection professionals as respondents, is that the respondents were presented with a predetermined and limited number of clues selected by the experimenters.

A third group of studies (i.e., The Free Response to Item Group) is exemplified by Ekman and O'Sullivan (1991) who contrasted the kinds of reasons accurate vs. inaccurate observers produced after their decisions on two different lie detection items (note: the observers were classified on the basis of whether they were accurate on a particular item, not on their overall accuracy). After deciding whether a particular individual was lying or telling the truth, the observers wrote down their reasons after each item. Every different reason provided was classified as verbal or nonverbal. Ekman and O'Sullivan found that, when observers got an item correct, they were more likely to report using nonverbal clues or a combination of verbal and nonverbal behaviors to make their decision. Those observers who were inaccurate on the item reported using only verbal clues. The kind of lie used in this study, however, was one in which the targets were lying about the emotions they were feeling as they watched an extremely upsetting film. A recent study by Warren (2007) suggests that sensitivity to nonverbal behavior is relevant to the detection of emotional lies but is uncorrelated with non-emotional lies. This report of different clues being used in making different kinds of inferences is consistent with research on empathic accuracy by Hall and Mast (2007). They found that verbal clues contributed the most to accuracy when thoughts were being inferred, but that visual nonverbal clues contributed more to inferring feelings.

Using a videotaped deception scenario in which people lied or told the truth about strongly held opinions, so that both feelings and thoughts were involved, Soohoo and O'Sullivan (2001) hypothesized that accurate and inaccurate judges would be more likely to use different strategies with different items. Using the written reasons given by 65 college students for correct or incorrect answers to two items, they demonstrated that accurate judges used different clues in the two items they got correct, whereas the inaccurate judges tended to use the same kind of clue (i.e., verbal vs. nonverbal) regardless of the relevance of those clues in the items they were judging.

A fourth and quite different approach to the verbal reporting of reasons for detecting deception is the Personal Remembered Lie approach examined by Park, Levine, McCornack, Morrison, and Ferrara (2002) in which they asked subjects to recall a lie that they had discovered in their own life; they then asked them how they discovered that lie. The authors argue that, with this paradigm, the results obtained are strikingly different from those obtained by research in the General Belief, Expert Belief or Free Response to Item approaches. When people reported how they detected a lie in their own life, the more important variables were physical evidence, confirmation by a third party, or the liar confessing the truth. Verbal and nonverbal behaviors of the sort studied by social scientists were rarely reported. Obviously, a personal recollected lie is more likely to be a serious one, which might have required a higher degree of certainty before being acknowledged.

In all of the studies reviewed, with the exception of Ekman and O'Sullivan (1991) and Soohoo and O'Sullivan (2001), the reasons were obtained from college students or police officers of no special lie detection ability. The last methodology to be considered is a variant of the Free Response to Item approach in which highly expert lie detectors are interviewed using a think aloud procedure to determine the emotional and cognitive processes they use in arriving at their correct decisions. Sample responses were chosen to illustrate how some of these experts use single words contextualized by their understanding of the individual and/or of their knowledge of what usual methods of discourse sound like.

The think aloud protocol was developed by Chase and Simon (1988) for their analysis of expert chess players. Since then, it has been widely used to study expertise in many different areas (Ericcson, 1996; Ericcson & Simon, 1998). Recently, Bond (2008) and O'Sullivan (2007; 2008) have used different versions of this technique to examine lie detection expertise. O'Sullivan has been doing research for the last 5 years on a group of extremely rare expert lie detectors (i.e., "truth wizards"), now totaling 50, who obtained highly accurate scores on at least two of three videotaped lie detection tasks. Although the study is ongoing, preliminary evidence suggests that any approach that groups reasons across items may not be the most useful, at least to understand how a truth wizard uses verbal material to assess truthfulness of a given individual. As Kraut (1978) noted: "A danger in research on impression management and on the detection of lying is to treat verbal and nonverbal cues associated with deception as if they were analogous to cues associated with emotion, and, thereby, underestimate the importance of the context in providing them with meanings" (p. 389).

"Truth wizards" are very sensitive to incongruities in communication, but this need not be between communication channels only (e.g., face vs. words). Many kinds of inconsistencies are attended to, including changes in the use of particular words within a statement and inconsistencies between what the person's overall appearance or manner suggests and what she is saying at a particular moment. What expert lie detectors seem to do when presented with videotaped material and asked to judge whether the person shown is lying or telling the truth is to contextualize the individual, to make sense of what it is that particular person is doing and saying in that particular situation. This takes into account the quality of the interview, the relationship with the

interviewer (Burgoon et al., 1999), the interactive style of the liar or truth teller and a host of other factors (see Griesel et al., this volume). Even individual words are given great weight.

Following are verbatim examples of this assessment of the "rightness" of words (*le mot juste*) from interviews done with some of the expert lie detectors (note: each "truth wizard" was also matched with a control who participated in the same kind of interviews). The "truth wizards" and their controls had already seen the following three lie detection videos since their responses to them (i.e., whether each individual was lying or telling the truth) were the bases for their identification as "truth wizards" or controls. The three lies included one about a strongly held opinion and another about stealing money. Both of these lies involved a significant cash reward and a threatened punishment (for more details, see Frank & Ekman, 1997). The third lie showed women lying or telling the truth about their feelings as they watched either a distressing surgical film or pleasant nature films. The subjects were nursing students motivated by a letter from the Dean of their school and the belief that their ability to control their emotional display was important to their professional success (see Ekman et al., 1988 for more details).

In the "debriefing" think aloud part of the study, each expert watched the video again and was encouraged to say whatever came into his or her mind (e.g., what she thought about the person she was looking at; what ideas went through her mind; what feelings she experienced, etc.). The interviewer adapted her style to maximize the quantity and quality of the verbal output provided by each expert. If an expert preferred to watch the entire interview before commenting, or wanted to go back and forth in the tape, that is what was done. If the expert made a movement or a sound or a facial expression, the interviewer would stop the tape and ask about it (e.g., "you just had a little smile on your face. What was that about?"; "When she said 'I am enjoying it', you cocked your head to the side. What were you thinking?"). At the end of each item, if the expert had not commented on a striking behavior that most observers mentioned, the interviewer would inquire about it (e.g., "Many people comment on the way he moves his eyebrows. What did you think about that?"). All 30 items were reviewed in this manner. The interviews took from 2 to 4 h to complete and were transcribed by one individual and checked by a second. What follows are a selection of comments made by expert lie detectors that illustrate the way in which they interpret single words.

The first expert to discuss is also the youngest involved in The Truth Wizard Project. Abigail was a 26 year old third year law student when she was identified as an expert lie detector. During the debriefing of one of the crime video items, she watched an interview with a young Asian man whom most observers think is telling the truth about whether he stole $50 (i.e., a great deal of money in 1995 when the experiment was conducted). Most people were positively impressed by his consistent and unwavering eye contact with the interviewer as well as by the lack of hesitation with which he answered the questions he was asked about whether he took the money. But Abigail was struck by the sound of a single word. She said:

"Did you hear how he said 'money?' … it was soft, and special, not at all like how he said the other words in the interview. Money is important to him. So when he says he doesn't need the money, I don't believe him." She went on to elaborate

that needing money was not the issue for him; he just really liked money. It gave him pleasure.

Pennebaker and others (e.g., Newman et al., 2003) have written about the importance of pronoun use. The research methods used in those studies, however, depend on an objective counting of the frequency of occurrence of such parts of speech. Many of the expert lie detectors, particularly those in legal professions, such as law enforcement personnel or arbitrators, pay attention not only to pronouns but to the context of the statement in which the pronouns occur. Although most of the "truth wizards" attend to both verbal and nonverbal behaviors, one of them depends almost exclusively on language.

Daniel is a well-known arbitrator. His language analysis, however, is contextualized by his astute assessment of the kind of person speaking: his intelligence, education, personality, social class, etc. In response to a young man saying that he did not support the death penalty, when he actually did, Daniel said:

> I don't (remember) what the question was, but (the man in the interview) answered it, 'I don't think *they* should be executed', which, whatever the question … is an odd way to put it. 'I don't think *people* should be executed,' 'I don't think *criminals* should be executed,' but the use of the term, *they* is interesting. Not conclusive, but interesting … it suggests … a depersonalization, an alienation (that I want to pay more attention to) …

Here is what Daniel had to say about another young man who was also lying about his belief in the death penalty:

> There were places in the conversation, (where) … he was kind of tentative … where I would not expect him to (be) … He talks quickly and without interruption at times, when I would expect him to be hesitant. He's looking directly at the interviewer, holding his gaze for a significant amount of time. But he's also looking down … from time to time. His facial expression is not a natural facial expression … the sequence is not right between what he's saying, how fast he's saying it, how his eyes are looking, how he's carrying his facial expression. So, I don't believe this guy.

All human beings have expectations of social behavior that they use to evaluate the information that occurs in their relationships with others. Daniel explicitly refers to it in his explanation of his thought processes. Other expert lie detectors do the same thing.

Linda is a retired FBI agent whose interest in language is suggested by her license plate, "lemotjust." She noticed that a young man who was telling the truth when he said he supported the death penalty for murderers focused on the victims, rather than the murderer. Linda said:

> See, and then he goes into 'before they kill someone' (else) … it's personal … personal responsibility to him.

About another young man who was lying about his opinion, Linda said:

> He probably would have rather had (the answer) be a yes or no, (but he is asked) 'What is your position?', so he has to say more than he thought he was going to have to say, so … he has to think about what his answer's going to be … he has to labor to get this opinion out because I don't think he believes it.

Abigail concluded that money was important to the speaker she was watching on a videotape because of the softness and carefulness with which he said the word. Based on that combination of speech-quality-perception and the inference from it that money was important, she did not believe him when, later in the interview, he said money was not important to him. Daniel had a template in his mind about how and when a tentative statement is made and how and when a strongly-held one is. He does not merely make an overall conclusion about a statement, but considers its constituent parts, sequence by sequence, weighing the plausibility of each component (note: his process is not unlike that suggested by Criteria-Based Content Analysis [CBCA; Griesel et al., this volume; Steller & Koehnken, 1989] and other language only methods. The added value, however, is the understanding of the uniqueness of the individual producing the words).

Another expert, Julian, is a law professor with a Ph.D. in counseling psychology. He was struck by the use of a particular word, but also checked to see whether it was consistent with other aspects of the truth teller's behavior. In explaining why he thought the man was telling the truth about his opinion he said:

> I'd say the biggest thing on him was that he says, 'hypocritical.' It was a big word, it describes his argument well, he said it so forcefully, and his head supported it.

In these two sentences, Julian illustrates observations consistent with findings from many research areas. He notes the consistency of the word with the argument; he notes that the word is unusual ("a big word"), that it was said forcefully, and that his head movements were consistent with it. He had already commented that the man looked intellectual, so his acceptance of the argument was also based on his assessment of the personality of that individual.

An expert lie detector, Liam, is an internal investigator for the Bureau of Prisons. He illustrates how expert lie detectors not only listen to single words, such as "no" but also observe the accompanying nonverbal behaviors and the consistency over the time of the interview (i.e., 1 min in this instance) of the pairing of verbal and nonverbal behavior. Liam said:

> … the first 'no,' he was sort of, you know, 'Of course I'm not lying,' it was … a downward smirk. And then the last 'no,' it was … an astonishment, his eyes … went up … I'd say he was lying.

What Liam was attending to was the combination of the content of speech with a variety of emotional states or motives (i.e., smirks vs. astonishment) and the inconsistency over the course of the interview of those feelings. The experts almost always commented on their uncertainty, noting that the inconsistency could arise from several factors, only one of which was deception. They would then run through a variety of alternative explanations and, only after ruling those out, would they settle on deception as the more likely explanation.

The scientific method demands replicable operationalization. This assumes that many samples can reasonably be coalesced. This assumption has led to the many valuable findings resulting from the studies briefly reviewed under the first four classifications discussed above: language only, measurement of language and nonlanguage behaviors, accuracy determination under limited information conditions,

and various models of impression formation, including those involving lie detection. The intensive analysis of the decision making processes of expert observers has been used in other arenas as well (Ericcson, 1996). Although most studies of expertise examine physicians, engineers, chess players and the like, Ceci and Liker (1986) examined a horse race handicapper.

Ceci and Liker (1986) provided a lengthy transcript of the thinking processes of a single highly successful handicapper and suggested that his talent combined both quantitative and qualitative aspects. The information from a general data base (e.g., previous record of the horse's speed) was qualified by knowledge of moderating variables—track condition, competition, jockey, weather, etc. They argued that this more individualized database will vary from handicapper to handicapper and will reflect their personalized knowledge. This personalized knowledge is more like the kinds of information possessed by individual therapists and artists. So, while science is necessary to support our understanding of the general processes involved in lying and lie detection, studies of the individualization, the contextualization of the decisions made by particular kinds of lie detectors are also needed.

Some beginning information in support of the above view has been reported originally by O'Sullivan and Ekman (2004) and, in more detail, using more experts by O'Sullivan (2008). The latter report contrasted the means and standard deviations on the three different lie detection scenarios explained above (e.g., opinion about death penalty, crime [e.g., stealing money] and emotion [e.g., feelings while watching pleasant vs. medical film]) obtained by 18 police professionals who had been identified as "truth wizards," as well as the corresponding information on the same three tests by ten expert lie detector therapists or psychologists. These means and standard deviations are listed in Table 10.1 (see above).

All six mean scores are significantly different from chance (50%). There was no significant difference between the police and the therapists on the opinion test since that test was used as the screening measure (note: to be considered for The Truth Wizard Project, potential experts needed to obtain scores of 80 or 90% on the opinion scenario). The opinion lie is the easiest for most examinees, since the items contain many verbal and nonverbal clues to deceit. The pattern of scores on the crime and emotion lies, however, was quite different for the two professional groups. For the police professionals, their accuracy on the crime items was significantly greater than their accuracy on the emotion lie items ($t[17] = 5.52$, $p < 0.000$). It is both not surprising, as well as reassuring, that police observers have significantly greater lie detection accuracy on lies concerning a crime (i.e., theft of a significant amount of money) than they do for lies concerning emotional reactions to films. This accuracy pattern makes sense and is mirrored by a complementary accuracy pattern obtained by the expert lie detectors who are therapists and psychologists. Comparing the therapists' crime and emotion detection accuracy scores with those of the police officers shown in Table 10.1 in a repeated measures ANOVA indicated a significant interaction between profession and test accuracy ($F[1, 26] = 31.407$, $p < 0.000$). Although some police officers and therapists were highly accurate on all three measures, overall, the police were significantly more accurate on the crime scenario and the therapists were significantly more accurate on the emotion scenario. This finding suggests that, while

it is likely that there are some generalized lie detection accuracy processes (e.g., based perhaps on cues such as those discovered in the science-based paradigms discussed earlier), there are also profession-specific, and perhaps, liar specific and expert specific clues to deception. The acquisition of that more specific knowledge base is probably acquired only through individual experience, feedback and motivation to master a particular kind of lie detection.

The idiographic analysis provided by the dissection of the cognitive and emotional processes used by expert lie detectors amplifies rather than contradicts laboratory research findings based on the examination of groups of examinees. All of the verbal, vocal and nonverbal clues to deception reported in the literature were used by one or more of the expert lie detectors. In addition, however, they reported scores of other behaviors that have not been analyzed in the literature: specific kinds of head rotations, nostril flaring, neck blushing, particular patterns of eye movements and, as has been illustrated here, a more nuanced interpretation of language than most language-only methods of deceptive communication ordinarily use.

In addition, the grounding of the "truth wizards'" interpretation of the behavior relevant to lie detection in their understanding of the kind of person showing that behavior is crucial. Heinrich and Borkenau (1998) argued that, in judging deception, most people make overall trait assessments of other people (also see ten Brinke & Porter, this volume). Both the experts and their controls certainly did that. What differentiated them from one another, however, was that the experts had more accurate, complex and far-reaching person perceptions and, although they often started their assessment with a global interpretation of the person, they considered alternative interpretations as they proceeded to watch the interviews (see Griesel et al., this volume). Matched controls, who were not accurate lie detectors, tended not to question their first impression and to ignore the implication of inconsistencies, even when they perceived them.

A difficulty in most experimental research of lie detection accuracy is that we study the pack, rather than the leader; the tribe, rather than Moses. This is appropriate if we are interested in the processes underlying the lie detection of insufficiently accurate observers. If we want to understand how people who actually can detect deception do so, then other subject groups in addition to non-randomly selected college students need to be examined (see a related argument for the study of eyewitness memory by Yuille, this volume). Studies of expertise may lack generalizability (Bond & DePaulo, 2008), but they offer superior guidance in terms of training for improvement and clarification of existing knowledge.

References

Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *The Forensic Examiner, 25*, 6–11.

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

Bond, C. F., Jr., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin, 134*, 477–492.
Bond, C. F., Jr., & Rao, S. R. (2004). Mendacity in a mobile world. In P. A. Granhag & L. Stromwall (Eds.), *The detection of deception in forensic contexts* (pp. 127–147). NY: Cambridge University Press.
Bond, G. D. (2008). Deception detection expertise. *Law and Human Behavior, 4*, 339–351.
Bugental, D., Kaswan, J., & Love, L. (1970). Perceptions of contradictory meanings conveyed by verbal and nonverbal channels. *Journal of Personality and Social Psychology, 16*, 647–655.
Burgoon, J. K., Buller, D. B., White, C. H., Afifi, W., & Buslig, A. L. S. (1999). The role of conversational involvement in deceptive interpersonal interactions. *Personality and Social Psychology Bulletin, 25*(6), 669–685.
Campbell, R. S., & Pennebaker, J. (2003). The secret life of pronouns: Flexibility in writing style and physical health. *Psychological Science, 14*(1), 60–65.
Ceci, S. J., & Liker, J. K. (1986). A day at the races: A study of IQ, expertise, and cognitive complexity. *Journal of Experimental Psychology. General, 115*(3), 255–266.
Chase, W. G., & Simon, H. A. (1988). The mind's eye in chess. In A. M. Collins & E. E. Smith (Eds.), *Readings in cognitive science: A perspective from psychology and artificial intelligence* (pp. 461–494). San Mateo, CA: Morgan Kaufmann.
Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science, 15*(10), 687–693.
DePaulo, B.M. (1998). *Deceiving and detecting deceit: Insights and oversights from the first several hundred studies*. Invited address. Washington, DC: American Psychological Society.
DePaulo, B. M., Lanier, K., & Davis, T. (1983). Detecting the deceit of the motivated liar. *Journal of Personality and Social Psychology, 45*, 1096–1103.
DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74–118.
DePaulo, B. M., Rosenthal, R., Eisenstat, R. A., Rogers, P. L., & Finkelstein, S. (1978). Decoding discrepant nonverbal cues. *Journal of Personality and Social Psychology, 36*(3), 313–323.
Ekman, P. (2001). *Telling lies: Clues to deceit in the marketplace, politics, and marriage* (3rd ed.). New York: W.W. Norton.
Ekman, P. (2003). *Emotions revealed*. New York: Holt.
Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1*, 49–98.
Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.
Ekman, P., Friesen, W. V., & O'Sullivan, M. (1988). Smiles when lying. *Journal of Personality and Social Psychology, 54*(3), 414–420.
Ekman, P., Friesen, W. V., O'Sullivan, M., & Scherer, K. R. (1980). Relative importance of face, body, and speech in judgments of personality and affect. *Journal of Personality and Social Psychology, 38*(2), 270–277.
Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist, 46*(9), 913–920.
Ekman, P., O'Sullivan, M., Friesen, W. V., & Scherer, K. (1991). Invited article: Face, voice and body in detecting deceit. *Journal of Nonverbal Behavior, 15*(2), 125–135.
Ericcson, K. A. (1996). The acquisition of expert performance: An introduction to some of the issues. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 1–50). Hillsdale, NJ: Lawrence Erlbaum.
Ericcson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture and Activity, 5*(3), 178–186.
Feeley, T. H., & deTurck, M. A. (1998). The behavioral correlates of sanctioned and unsanctioned deceptive communication. *Journal of Nonverbal Behavior, 22*(3), 189–204.
Feldman, R. S., & Jenkins, L. (1979). Detection of deception in adults and children via facial expressions. *Child Development, 50*(2), 350–355.

Fiske, S. T. (1992). Thinking is for doing: Portraits of social cognition from Daguerreotype to laserphoto. *Journal of Personality and Social Psychology, 63*(6), 877–889.
Frank, M. G., & Ekman, P. (1997). The ability to detect deceit generalizes across different types of high-stake lies. *Journal of Personality and Social Psychology, 72*(6), 1429–1439.
Funder, D. (1999). *Personality judgment: A realistic approach to person perception*. San Diego: Academic.
Gardner, H. (1993). *Frames of mind: The theory of multiple intelligences*. New York: Perseus.
Granhag, P. A., & Strömwall, L. A. (2002). Repeated interrogations: Verbal and non-verbal cues to deception. *Applied Cognitive Psychology, 16*, 243–257.
Hall, J. A., & Mast, M. S. (2007). Sources of accuracy in the empathic accuracy paradigm. *Emotion, 7*(2), 438–446.
Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2004). Police officers' lie detection accuracy: Interrogating freely versus observing video. *Police Quarterly, 7*(4), 429–436.
Heinrich, C. U., & Borkenau, P. (1998). Deception and deception detection: The role of cross-modal inconsistency. *Journal of Personality, 66*(5), 687–712.
Ickes, W. (1993). Empathic accuracy. *Journal of Personality, 61*, 587–610.
Kraut, R. (1978). Verbal and nonverbal cues in the perception of lying. *Journal of Personality and Social Psychology, 36*(4), 380–391.
Mann, S. A., & Vrij, A. (2006). Police officers' judgments of veracity, tenseness, cognitive load and attempted behavioral control in real-life police interviews. *Psychology, Crime & Law, 12*(3), 307–319.
Masip, J., Sporer, S. L., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law, 11*(1), 99–122.
Mayer, J. D., Salovey, P., & Caruso, D. (2002). *Mayer-Salovey-Caruso emotional intelligence test, user's manual*. Toronto, Canada: Multi-Health Systems.
Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and Human Behavior, 26*(5), 469–480.
Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*(5), 665–675.
O'Sullivan, M. (2003). The fundamental attribution error in detecting deception: The boy-who-cried-wolf effect. *Personality and Social Psychology Bulletin, 29*(10), 1316–1327.
O'Sullivan, M. (2005). Emotional intelligence and detecting deception. Why most people can't "read" others, but a few can. In R. Riggio & R. Feldman (Eds.), *Applications of nonverbal communication* (pp. 215–253). Mahwah, NJ: Erlbaum.
O'Sullivan, M. (2007). Unicorns or Tiger Woods: Are lie detection experts myths or realities? A response to on lie detection wizards by Bond and Uysal. *Law and Human Behavior, 31*, 117–123.
O'Sullivan, M. (2008). Homeruns and humbugs: Comment on Bond and DePaulo (2008). *Psychological Bulletin, 134*, 493–497.
O'Sullivan, M., & Ekman, P. (2004). The wizards of deception detection. In P. A. Granhag & L. Stromwell (Eds.), *Detecting deception in forensic contexts* (pp. 269–286). Cambridge, UK: Cambridge University Press.
O'Sullivan, M., Ekman, P., Friesen, W., & Scherer, K. R. (1985). What you say and how you say it: The contribution of speech content and voice quality to judgments of others. *Journal of Personality and Social Psychology, 48*(1), 54–62.
O'Sullivan, M., & Guilford, J. P. (1975). Six factors of behavioral cognition: Understanding other people. *Journal of Educational Measurement, 12*(4), 255–271.
Park, E. S., Levine, T. R., Harms, C. M., & Ferrara, M. H. (2002). Group and individual accuracy in deception detection. *Communication Research Reports, 19*(2), 99–106.
Park, H. S., Levine, T. R., McCornack, S. A., Morrison, K., & Ferrara, M. (2002). How people really detect lies. *Communication Monographs, 69*(2), 144–157.

Porter, S., Doucette, N. L., Woodworth, M., Earle, J., & MacNeil, B. (2008). Halfe the world knows not how the other halfe lies: Investigation of verbal and non-verbal signs of deception exhibited by criminal offenders and non-offenders. *Legal and Criminological Psychology, 13*, 27–38.

Porter, S., & Yuille, J. C. (1995). Credibility assessment of criminal suspects through statement analysis. *Psychology, Crime & Law, 1*, 1–13.

Rockwell, P., Buller, D. B., & Burgoon, J. K. (1997). The voice of deceit: Refining and expanding cues to deception. *Communication Research Reports, 14*, 451–459.

Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. New York: McGraw-Hill.

Soohoo, T., & O'Sullivan, M. (2001). *Lie detection: Decision reasons and accuracy*. Poster presented at the annual meeting of the Society for Personality and Social Psychology, San Antonio, TX.

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology, 20*, 421–446.

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law, 13*(1), 1–34.

Steller, M., & Koehnken, G. (1989). Criteria-based statement analysis. In D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 217–245). New York, NY: Springer Publishing.

Strömwell, L. A., & Granhag, P. A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges. *Psychology, Crime & Law, 9*, 19–36.

Undeutsch, U. (1982). Statement reality analysis. In A. Trankell (Ed.), *Reconstructing the past: The role of psychologists in criminal trials* (pp. 27–56). Stockholm: Norsted & Sons.

Vrij, A., Akehurst, L., Soukara, S., & Bull, R. (2006). Detecting deceit via analyses of verbal and nonverbal children and adults. *Human Communication Research, 30*, 8–41.

Vrij, A., & Heaven, S. (1999). Vocal and verbal indicators of deception as a function of lie complexity. *Psychology, Crime & Law, 5*(3), 203–215.

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., & Semin, G. R. (1996). Lie experts' beliefs about nonverbal indicators of deception. *Journal of Nonverbal Behavior, 20*(1), 65–80.

Warren, G. (2007). *The development of a deception detection task: The importance of emotion*. Paper presented at the Annual Conference of the Division of Forensic Psychology, British Psychological Society, University of York, UK.

Yuille, J. C. (1989). *Credibility assessment*. The Netherlands: Kluwer.

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